Analysis of reprogramming-associated alterations using an isogenic human stem cell system

DISSERTATION

zur Erlangung des Doktorgrades (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Carolin Haubenreich

aus Kitzingen

Bonn 2020

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Oliver Brüstle

2. Gutachter: Prof. Dr. Walter Witke

Tag der mündlichen Prüfung: 9.3.2021

Erscheinungsjahr: 2022

Contents

1 Introduction 1

1.1 Stem cells 1

1.1.1 Pluripotent stem cells 1

1.1.2 Neural stem cells 4

1.2 Reprogramming to pluripotency 7

1.2.1 Transcriptional and epigenetic changes during reprogramming 9

1.2.2 Aberrations occur during reprogramming to pluripotency 13

1.3 Aim of the study 15

2 Material 16

2.1 Technical equipment 16

2.2 Labority consumables 17

2.3 Chemicals and reagents 18

2.4 Cell culture stock solutions 20

2.5 Cell culture solutions 20

2.6 Cell culture media 22

2.7 Cell lines and strains 23

2.8 Kits 24

2.9 Enzymes 24

2.10 Primers 24

2.11 Plasmids 27

2.12 Molecular biology buffers 27

2.13 Reagents and media for microbiological methods 28

2.14 Antibodies 29

2.15 Immunocytochemistry solutions 29

2.16 Software 30

3 Methods 31

3.1 Cell culture methods 31

3.1.1 Thawing of cells and cell storage 31

3.1.2 Cell number determination 31

3.1.3 Cultivation of pluripotent stem cells 31

3.1.4 Generation of embryoid bodies 32

3.1.5 Generation and cultivation of neural stem cells 32

3.1.6 Differentiation of neural stem cells 33

3.1.7 *In vitro* differentiation of pluripotent cells into three germ layers 33
3.1.8 *In vivo* differentiation of iPSCs into three germ layers 33
3.1.9 Generation of murine embryonic fibroblasts 34
3.1.10 Reprogramming of NSCs to pluripotency 34
3.1.11 Production of viral particles 35
3.1.12 Flow Cytometry 36
3.2 Molecular biological methods 37
3.2.1 Large-scale DNA preparation 37
3.2.2 Endonuclease Restriction 37
3.2.3 DNA gel electrophoresis 37
3.2.4 Purification of total DNA from cells 38
3.2.5 DNA acetate precipitation 38
3.2.6 Purification of total RNA from cells 39
3.2.7 DNA and RNA Quantification 39
3.2.8 DNase digestion and cDNA synthesis 40
3.2.9 Primer design 40
3.2.10 Polymerase chain reaction 40
3.2.11 Quantitative real time PCR 42
3.2.12 Genome wide gene expression analysis 43
3.2.13 Whole genome methylation analysis 44
3.2.14 SNP analysis 44
3.3 Microbiological methods 44
3.3.1 Cultivation and storage of bacterial cells 44
3.3.2 Generation of competent bacteria via $CaCl_2$ 45
3.3.3 Transformation 45
3.4 Cell stainings 46
3.4.1 Fixation of cells 46
3.4.2 Alkaline phosphatase staining 46
3.4.3 Immunocytochemistry 46
4 Results 47
4.1 Generation of a human isogenic stem cell system to study reprogramming-associated aberrations 47
4.1.1 Generation and characterization of iPSCs 47
4.1.2 Human induced pluripotent stem cells exhibit pluripotency properties 49
4.1.3 Derivation of stable NSC lines from pluripotent stem cells 56

4.1.4 iPSC-derived NSC lines resemble their ESC-derived counterparts in marker expression and differentiation potential ... 57
4.2 Analysis of the transcriptional profiles in an isogenic stem cell system ... 59
4.2.1 Differential expression is enriched on the X-chromosome ... 62
4.2.2 Comparison of different reprogramming strategies ... 66
4.3 Enriched differential methylation on the X-Chromosome coincides with differential expression ... 67
4.4 Implications of the reprogramming process on the X-chromosome inactivation status ... 72
5 Discussion ... 76
5.1 Assessment of the isogenic stem cell system ... 76
5.1.1 Induced pluripotent stem cells display bona fide pluripotency ... 79
5.1.2 iPSCs possess comparable neural differentiation potential to ESCs ... 81
5.2 Suitability of the human isogenic stem cell system for the analysis of reprogramming-associated alterations ... 82
5.2.1 An isogenic background is essential for comparative analysis of iPSCs and ESCs ... 82
5.2.2 Profiling of somatic cells is essential to approve suitability of iPSCs for research and therapy ... 84
5.3 iPSC- and ESC-derived NSCs display a highly similar transcriptome and DNA methylation pattern ... 85
5.4 Consequences of reprogramming on genomic imprinting ... 86
5.5 Consequences of reprogramming on X-chromosomal gene expression and methylation ... 89
5.6 Dosage compensation in iPSC-derived neural stem cells ... 91
5.7 Limitations of the applied isogenic stem cell system ... 98
5.8 Outlook ... 99
5.9 General Conclusion ... 101
6 Abbreviations ... 104
7 Abstract ... 106
8 Zusammenfassung ... 107
9 References ... 109
10 Danksagung ... 136
11 Publications ... 138

1 Introduction

1.1 Stem cells

Stem cells are undifferentiated cells owing the ability for self-renewal and the capacity to differentiate into at least one specialized cell type (Weissman, 2000). During development of a mammalian organism, after fusion of the female oocyte and a male sperm, a so-called zygote is formed. The zygote divides numerous times and gives rise to further differentiated cells that in the end constitute the whole embryo as well as extra-embryonic tissue. The capacity to generate the whole organism is constricted to the zygote and the cells of the first three cell divisions and is called totipotency. After the 8-cell-state, which is considered as morula, the blastocyst is formed. This blastocyst consists of an outer layer of trophoblast cells and an inner layer of cells, the inner cell mass (ICM). Cells of the ICM have the capacity to generate all cells of the three germ layers including germ cells and are therefore considered as pluripotent. During further development, stem cells with a more restricted potency occur, which can also be found in a variety of tissues in an adult organism, such as in bone marrow, intestine, pancreas, liver and skin (Hall and Watt, 1989; Potten and Loeffler, 1990; Weissman, 2000). Many of these cells have a multipotent capacity, allowing them to form multiple cell types of their lineage and are classified according to their origin, like hematopoietic and neural stem cells. Hematopoietic stem cells, for example, give rise to lymphoid and myeloid progenitor cells, which finally generate all different blood cells. Finally, there are also unipotent stem cells that only possess the potency to differentiate into one specialized cell type like spermatogonial stem cells.

1.1.1 Pluripotent stem cells

Pluripotent stem cells can differentiate into cells of all three germ layers including the germ cells. There are several ways to capture this pluripotency *in vitro* (Figure 1). During early embryonic development, the inner cell mass of a blastocyst consists of pluripotent stem cells. This pluripotent capacity can be captured in culture, when cells of the ICM are isolated and cultivated in appropriate media to generate embryonic stem cell lines (Evans and Kaufman, 1981; Martin, 1981; Thomson et al., 1998). Similarly, another type of pluripotent cells can be generated via parthenogenetic development of oocytes. These so called parthenogenetic embryonic stem cells can be generated when unfertilized oocytes, arrested

in the metaphase of the second meiotic division (MII), are chemically activated to initiate embryonic development (Cibelli et al., 2002; Lin et al., 2007).

Furthermore, the pluripotent state can also be induced in somatic cells by somatic cell nuclear transfer (SCNT), cell fusion with embryonic stem cells or direct reprogramming to pluripotency. In the case of SCNT, the nucleus of a somatic cell is inserted into a previously enucleated unfertilized oocyte. This leads to a reprogramming of the nucleus by the cytoplasmic environment in the host oocyte. Upon chemical stimulation, the division of the fused cell is triggered and normal development as in the case of a fertilized egg is initiated (Gurdon, 1962; Campbell et al., 1996; Wilmut et al., 1997). At the blastocyst state, the cells of the ICM can be transferred to the appropriate media as it is done for embryonic stem cells and kept as pluripotent stem cells in culture. This procedure is called therapeutic cloning, as it is possible to create patient-specific stem cell lines, that contain only the genetic information of the somatic cell initially used, thus enabling their use for therapeutic purposes in personalized medicine (Yamada et al., 2014). SCNT can also be used to generate a whole organism with the genetic information of only one individual, which is referred as reproductive cloning, as for example in the case of Dolly the sheep, which was the first reported cloning of a mammal (Wilmut et al., 1997).

In addition to exploiting the oocytic development potential, somatic cells can be epigenetically reprogrammed by fusion with already established embryonic stem cells. The hybridization of both cell types is achieved by electrical fusion, resulting in stable pluripotent hybrid cells that retain a tetraploid set of chromosomes (Tada et al., 2001; Cowan et al., 2005).

Finally, somatic cells can be directly reprogrammed to pluripotency. First described by Takahashi and Yamanaka in 2006, the ectopic expression of four transcription factors in somatic cells can lead to a stable reprogramming of cell identity if maintained under appropriate culture conditions (Takahashi and Yamanaka, 2006). This method has the advantage that at no stage an oocyte is required for the production of pluripotent cells.

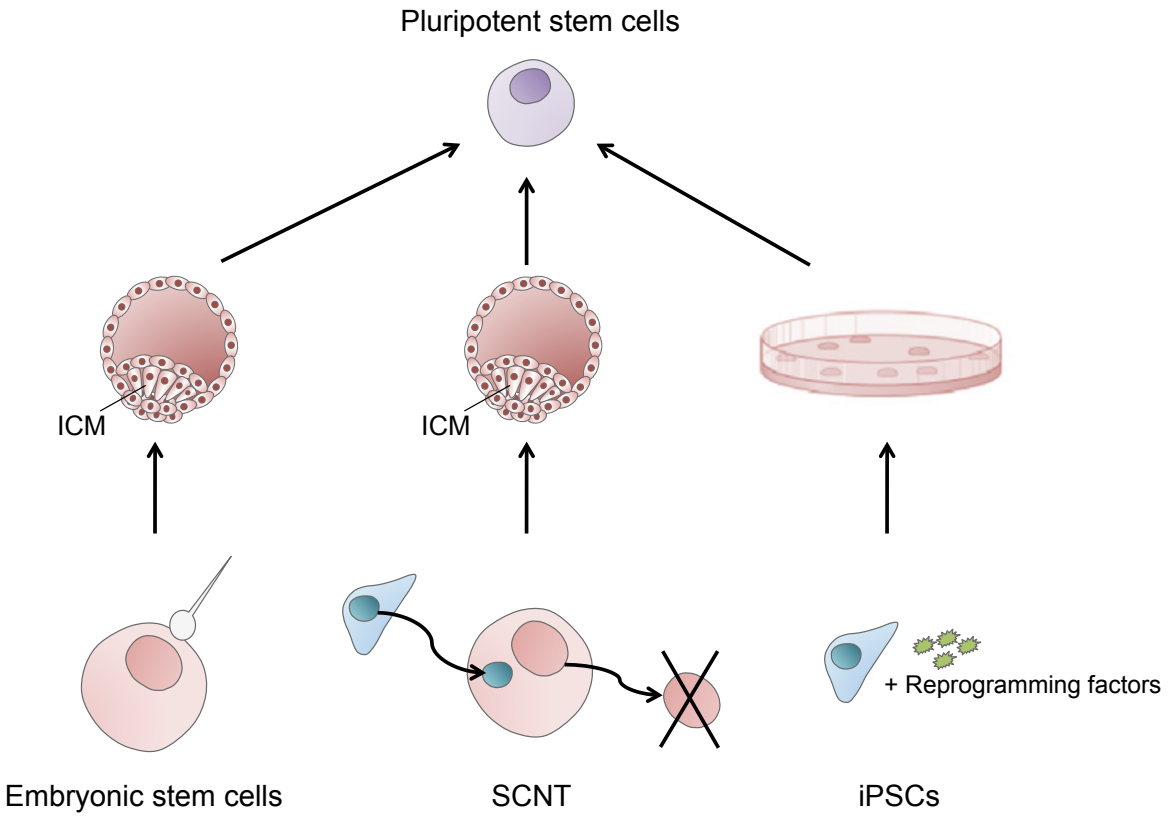


Figure 1: Classical ways to capture the pluripotent state *in vitro*. Pluripotent stem cells can be generated by extraction and cultivation of cells of the inner cell mass (ICM) at the blastocyst stage after fertilization or somatic cell nuclear transfer into an enucleated oocyte and by direct reprogramming of somatic cells via forced expression of certain factors.

In rodents, two different types of pluripotent stem cells can be obtained at two distinct stages of early development. On the one hand, naïve mouse embryonic stem cells (mESCs) can be derived from the ICM of pre-implantation blastocysts and on the other hand, so called primed epiblast stem cells (mEpiSCs) can be generated from the early post-implantation embryo (Tesar et al., 2007; Rossant, 2008). Both cell types are considered as pluripotent, but represent distinct pluripotent states that differ on the molecular and epigenetic level. Naïve embryonic stem cells efficiently contribute to chimeric embryos when injected into blastocysts, maintain both X-chromosomes in an active state (XaXa), differentiate into primordial germ cells (PGC) *in vitro*, show a high clonality, express Oct4 from its distal enhancer, show a global decrease in DNA methylation and contain H3K27me3 repressive chromatin deposition on lineage regulatory genes (Marks et al., 2012). In contrast, neither freshly isolated post-implantation epiblast cells nor primed mEpiSCs are capable to efficiently repopulate a host blastocyst (Rossant et al., 1978; Tesar et al., 2007). Furthermore, EpiSCs retain predominantly an inactive X-chromosome (XaXi), are poised for PGC differentiation (Hayashi and Surani, 2009), show a low cloning efficiency when

dissociated during passaging, express Oct4 from the proximal enhancer and show increased global DNA methylation and bivalent chromatin marks on developmental regulatory gene promoters. These specific molecular differences are also reflected in distinct culture conditions for both cell types. Self-renewal of naïve mESCs is dependent on the cytokine leukemia inhibitory factor (LIF) and either serum, bone morphogenic protein 4 or inhibition of GSK3 and MEK/ERK (2i; Wray et al., 2010). mESCs differentiate upon bFGF/ERK signaling, while EpiSCs are maintained by bFGF and Activin. When compared to the human system, human ESCs (hESCs) resemble primed EpiSCs in their properties and culture conditions and differentiate or die when cultured in 2i-LIF medium (Hanna et al., 2010). So far, this was considered as insignificant species-specific characteristic, but there is increasing evidence that ground state naïve pluripotency is also possible in the human system. First studies indicate a shift of conventional human primed pluripotent cells to naïve ground state pluripotency, when cultured in 2i-LIF medium with the addition of further small molecules or simultaneous ectopic expression of pluripotency factors (Guo et al., 2009, 2016; Hanna et al., 2010; Gafni et al., 2013; Takashima et al., 2014; Theunissen et al., 2014; Duggal et al., 2015).

1.1.2 Neural stem cells

Neural stem cells (NSCs) are defined as a multipotent cell population with the potential for proliferation and the capacity for self-renewal. Present during embryonic development as well as in the adult mammalian brain, these cells generate neurons, astrocytes and oligodendrocytes in the developing nervous system and are crucial for the neurogenesis potential of the adult brain. *In vivo*, NSCs are located in a multicellular, specialized environment to sustain self-renewal and to regulate their differentiation, the so-called niche. This microenvironment provides factors balancing between symmetrical divisions for stem cell self-renewal and asymmetrical divisions to generate fate-committed progeny. *In vitro*, NSCs are cultured and expanded with specific growth factors, mimicking this niche, as free-floating aggregates or as adherent cultures (Conti and Cattaneo, 2010).

After gastrulation during early vertebrate neural development the epiblast forms the ectodermal germ layer, which in turn forms a so-called neural plate, from where the entire nervous system emerges. The neuroepithelial cells of the neural plate fold into a neural tube

and are regionalized to create specific classes of neural progenitors of the forebrain, midbrain, hindbrain and spinal cord (Pankratz et al., 2007).

Starting from pluripotent stem cells, there are various protocols for the generation of distinct neural stem cells that attempt to capture these various stages of neural development *in vitro* (Figure 2). First studies in mouse suggest the existence of FGF2/LIF responsive primitive neuroepithelial progenitors (NEPs). Some of these cells still exhibit ESC characteristics such as Oct4 expression, which might indicate an incomplete neural specification. In addition, murine primitive neuroepithelial progenitors can only be transiently kept in culture, as they spontaneously convert toward a more committed precursor stage with a more restricted differentiation potential (Tropepe et al., 2001).

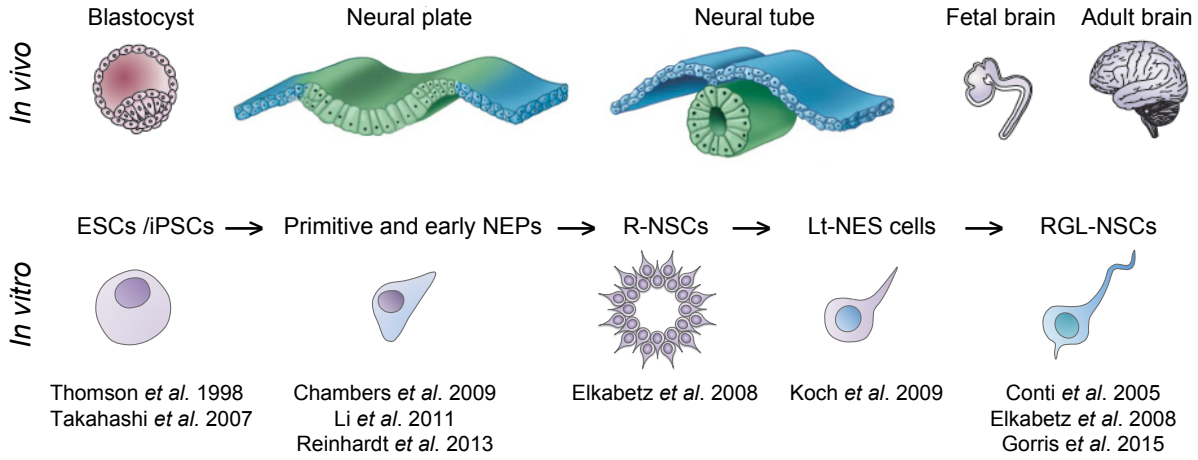


Figure 2: Different human neural stem cell types that can be captured *in vitro*. During embryonic development different neural stem cell populations are generated. These distinct developmental stages presumably can also be captured *in vitro*. Obtained human neural stem cell populations differ in their morphology, marker expression, cell-cycle characteristics and developmental potential. Illustration is adapted from Conti and Cattaneo 2010.

One of the earliest precursor population in human neural development can be induced in pluripotent cells by simultaneous inhibition of GSK3, TGFβ and Notch-signaling via small molecules. These human primitive neuroepithelial progenitor cells show a high neurogenic differentiation potential, responsiveness to regionalization cues allowing the differentiation towards neuronal subpopulations like dopaminercic neurons and motorneurons and, in contrast to murine primitive NEPs, maintain stable self-renewal in the presence human LIF (Li et al., 2011).

Another NSC-differentiation protocol from human pluripotent cells is based on dual inhibition of SMAD signaling by using the BMP inhibitor noggin and SB431542, an inhibitor of

the Lefty/Activin branch of the TGFβ pathway (Chambers et al., 2009). Notably, comparable results are suggested by expression of inhibitory SMAD7, which represses both branches of TGFβ signaling (Ozair et al., 2013). Treatment of hESCs with noggin and SB431542 results in fully neuralized, early NEPs expressing early neural markers like SOX1, PAX6, OTX2 and FOXG1. NSCs derived by this approach are highly responsive to patterning morphogenes, for the derivation of region-specific neuronal fates like TH positive neurons and HB9 positive motor neurons. However, early NEPs seem to be a transient *in vitro* stage as passaging shifts them to a more committed NEP population that grows in rosette-like clusters (Chambers et al., 2009). These radially aligned epithelial cells or rosette-NSCs (R-NSCs) can be kept in culture for a few passages in the precence of sonic hedgehog (SHH) and Notch ligands and can also be directly isolated from default neuralized human ESCs via cell sorting for the anterior marker Forse1 and neural cadherin (N-cadherin; Elkabetz et al., 2008). The homogenous cell population of R-NSCs shows responsiveness to regionalization cues that promote differentiation toward subtype-specific neural identities of the central nervous system (CNS) and peripheral nervous system (PNS). However, the observation of tumorigenic potential for overgrowth following transplantation in adult rodent brains seems to be a critical point for their application in cell therapy. More recently, a neural precursor cell population lacking the rosette-like organization but with a similar developmental stage as R-NSCs could be generated by combined modulation of WNT and SHH signaling (Reinhardt et al., 2013). Upon neural induction with the small molecules dorsomorphin and SB431542, these small molecule neural precursor cells (smNPCs) can be stably propagated with the GSK3 inhibitor CHIR to stimulate WNT signaling and purmorphamine for SHH pathway stimulation and can be differentiated in response to instructive cues into midbrain dopaminergic neurons, motor neurons as well as peripheral neurons. Another human ESC-derived rosette-like patterned neural stem cell population, named lt-hESNSCs (or lt-NES cells), could be generated and stably maintained over 150 *in vitro* passages in the presence of EFG and FGF2 (Koch et al., 2009; Falk et al., 2012). This robust, homogenous and pure NSC-population represents a more posteriorized regional identity than R-NSCs, but still retains expression of Forse1 and N-cadherin and responds to instructive cues that direct induction of specific neuronal identities, like motor neurons and midbrain dopaminergic neurons, even at higher passages. A later cell population of neural development, the radial glia (RG) cells, can also be derived from pluripotent cells in various ways, resulting in distinct

propagation and differentiation potentials (Nat et al., 2007; Glaser et al., 2007; Gorris et al., 2015). A remarkably stable population can be derived by cultivation of SOX1-positive cells in the presence of EGF and FGF2, resulting in elongated, bipolar shaped nestin-, PAX6- and BLBP-positive RG-like NSCs. These RG-like NSCs can be stably maintained for over 100 passages and retain even at late passages their differential potential into neurons and astrocytes (Conti et al., 2005; Conti and Cattaneo, 2010). Moreover, also R-NSCs can be converted into RG-like NSCs with similar properties, when cultured in the presence of the commonly used mitogens EGF and FGF2. In comparison to neuroepithelial cells, RG-like NSCs obtain a higher gliogenic potential and are more limited in their susceptibility to regionalization cues and capacity for neuronal subtype formation (Elkabetz et al., 2008; Ostermann et al., 2019).

1.2 Reprogramming to pluripotency

Embryonic stem cells (ESCs), derived from the inner cell mass of mammalian blastocysts, have the ability to grow indefinitely while maintaining pluripotency and the ability to differentiate into cells of all three germ layers (Evans and Kaufman, 1981; Martin, 1981). These properties have led to the expectations that human ESCs might be useful to understand disease mechanisms, to screen for effective and safe drugs and to treat a host of diseases, such as Parkinson's disease, spinal cord injury and diabetes (Thomson et al., 1998). However, the use of human embryos is limited and faces ethical controversies that hamper the application of human ESCs, and in addition there is the problem of potential tissue rejection following transplantation in patients. One way to circumvent these issues in research and therapy is the generation of pluripotent cells directly from the patients´ own somatic cells via direct reprogramming to pluripotency (Figure 3).

Diverse mouse and human somatic cells can be directly reprogrammed by ectopic expression of different sets of defined factors to yield induced pluripotent stem cells (iPSCs). These iPSCs resemble ESCs in their capacity to form chimeric embryos and contribute to all three germ layers (Takahashi and Yamanaka, 2006; Wernig et al., 2007; Okita et al., 2007). Reprogramming of mouse and human fibroblasts was originally achieved by ectopic expression of POU-class 5 homeobox 1 (POU5F1, also known as OCT4), SRY-box containing gene 2 (SOX2), Krüppel-like factor 4 (KLF4) and myelocytomatosis oncogene (MYC; c-Myc), which are also called Yamanaka factors (Takahashi and Yamanaka, 2006; Takahashi et al.,

2007). Other studies showed that the pluripotent state could also be induced by replacing each of these Yamanaka factors with other pluripotency factors resulting in various combinations. For example, somatic cells can successfully be reprogrammed by replacing KLF4 and c-Myc with the pluripotency factors NANOG and LIN28 in human cells (Yu et al., 2007), or with Esrrb in the mouse (Feng et al., 2009) and OCT4 can be replaced with epithelial cadherin (E-cadherin; Redmer et al., 2011).

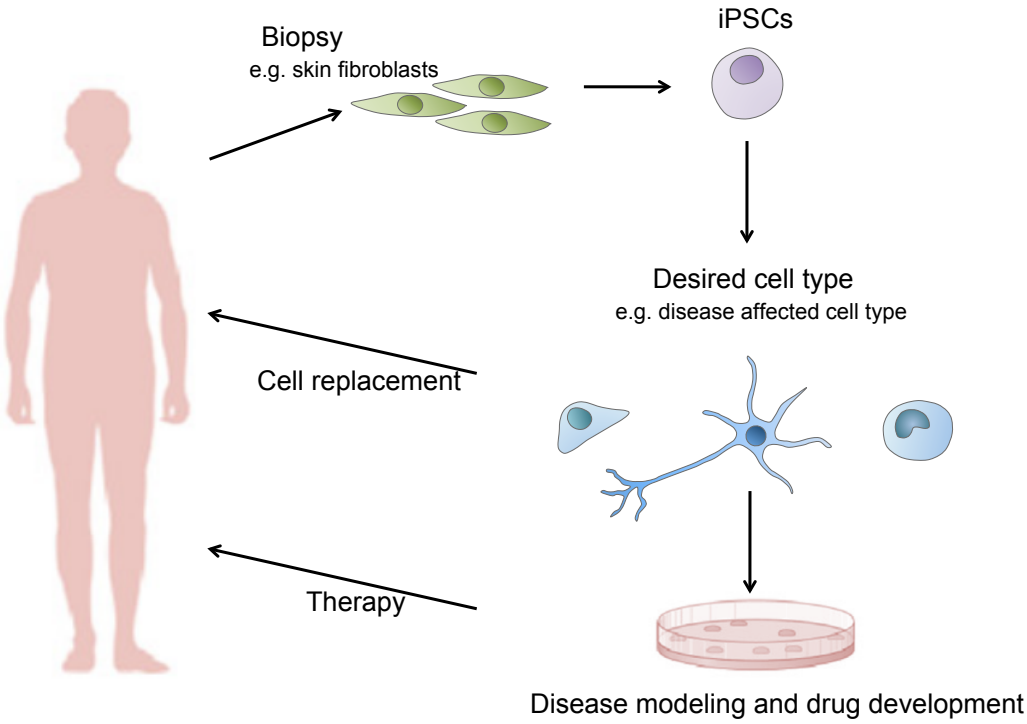


Figure 3: Possible applications of iPSCs in research and therapy. iPSCs represent a valuable tool for cell replacement therapies and basic research. Somatic cells of diseased and healthy donors can be reprogrammed to iPSCs, which can then be differentiated into the desired (e.g. disease affected) cell type. Patient specific cells can either be used in disease modeling and drug discovery or after modification (e.g. gene correction) for autologous cell therapy

Individual factors can also be omitted in the reprogramming-cocktail. For generation of iPSCs from mouse and human fibroblasts c-Myc is dispensable, suggesting that endogenous c-Myc expression could have a role in the reprogramming process (Nakagawa et al., 2008; Wernig et al., 2008). Similarly, due to endogenous expression of Sox2 in neural stem cells, these cells can be reprogrammed with only two factors (Oct4 and Klf4 or c-Myc; Kim et al., 2008) or even with Oct4 only (Kim et al., 2009a; c).

Epigenetic modifiers could also be used to improve efficiency of reprogramming or replace factors of the reprogramming cocktail, like the histone deacetylase inhibitor (HDAC), valproic acid (VPA), sodium butyrate and other small molecules (Shi et al., 2008; Huangfu et al.,

2008b; Mali et al., 2010) as well as microRNAs (Anokye-Danso et al., 2011; Miyoshi et al., 2011; Subramanyam et al., 2011).

Numerous studies could demonstrate the robustness of reprogramming, considering that not only fibroblasts and neural stem cells, but also various cell types can be reprogrammed. Successful reprogramming was demonstrated for renal epithelial cells from urine (Zhou et al., 2012), cord blood cells (Haase et al., 2009), liver and stomach cells (Aoi et al., 2008), pancreatic beta cells (Stadtfeld et al., 2008a), mesenchymal stem cells, keratinocytes (Wernig et al., 2008) and even post mitotic cells like neurons (Kim et al., 2011a) and terminally differentiated mature B lymphocytes (Hanna et al., 2008). Moreover, apart from human and mouse, this technique is applicable to a diversity of species like marmoset (Wu et al., 2010), pig (West et al., 2010), prairie vole (Manoli et al., 2012) and horse (Breton et al., 2013).

Furthermore, pluripotency can be established via several techniques. Initially, reprogramming factors were introduced into the somatic cells by genome-integrating retroviral and lentiviral delivery (tetracycline inducible or CRE-excisable polycistronic vectors) or transposon-based systems (piggyBac; Woltjen et al., 2009; Kaji et al., 2009; Yusa et al., 2009), but also non-integrating adenovirus (Stadtfeld et al., 2008b) or Sendai virus (Fusaki et al., 2009), direct delivery of plasmids (Okita et al., 2008; Chou et al., 2011), recombinant proteins (Zhou et al., 2009; Kim et al., 2009b), synthetic modified mRNAs (Warren et al., 2010), mature microRNAs (Miyoshi et al., 2011) and small molecules (Hou et al., 2013) have been used for induction of pluripotency.

Though reprogramming to pluripotency is robust and largely reproducible, it is a rather complex process, which is highly inefficient at the population level and requires a considerable latency before autonomous pluripotency is achieved (Pour et al., 2015). There are several factors that can compromise its efficiency and quality like the cell type of origin and its number of passages in culture, the age of the cell donor, composition of factors and their stoichiometry, type of delivery as well as culture conditions like medium type and oxygen level.

1.2.1 Transcriptional and epigenetic changes during reprogramming

In order to reach pluripotency during reprogramming, several rearrangements must take place in a somatic cell. The expression profile has to be adjusted from the differentiated

state to a pluripotent transcriptional status, by shutting down lineage specific genes and activating pluripotency circuits. This is achieved by global epigenetic changes due to DNA methylation and acetylation as well as histone-based chromatin remodeling. In the case of reprogramming postmitotic cell types, even cell cycling has to be reestablished.

These readjustments involve distinct steps that can be classified as an early stochastic and a later deterministic phase. When reprogramming mouse fibroblasts with the Yamanaka cocktail, as a primary response to transgene induction, almost exclusively genes located in accessible H3K4me3 chromatin are transcriptionally affected (Koche et al., 2011). In this open chromatin regions OKSM can bind to the promoter regions of their downstream targets and up or down regulate their expression. In this first wave of differential expression, more sites are transcriptionally activated than necessary, for example keratinocyte/epidermis-related genes are upregulated, which are not expressed in fibroblasts or iPSCs (O'Malley et al., 2013). Those genes are transiently expressed and again downregulated when full pluripotency is reached (Soufi et al., 2012; Polo et al., 2012; Takahashi et al., 2014). Most routes may lead to a dead end, but some cells undergo a so called mesenchymal to epithelial transition (MET) when using OKSM factors. During this transition, Snail and Thy1 genes are silenced, TGFβ signaling is suppressed and E-cadherin expression is gained (Li et al., 2010; Samavarchi-Tehrani et al., 2010; O'Malley et al., 2013). Besides these expressional changes, MET induces dramatic cytoskeletal reorganizations and morphological changes that lead to a compact epithelial phenotype (Thiery et al., 2009). However, MET seems not necessary for other factor combinations as other routes can also lead to pluripotency by balancing different lineages (Shu et al., 2013; Montserrat et al., 2013). Right after MET, early onset pluripotency genes are upregulated like endogenous Oct4, Fgf4 and Fbxo15, but these are not strictly predictive of successful reprogramming (Buganim et al., 2012).

Furthermore, OSK but not c-Myc can bind as so-called pioneer factors at heterochromatin with exception of H3K9me3 marked regions (Soufi et al., 2012). The factors bind in nucleosome rich regions predominantly synergistically at distal elements of silent genes that will not become reactivated until later stages of reprogramming. It seems that those regions cannot yet be activated because of lacking promoter binding, but are getting prepared for later reactivation, including pluripotency genes like Esrrb and Sall4 (Soufi et al., 2012). Because c-Myc alone stringently binds to open chromatin regions, it is dependent on OSK co-

binding to target heterochromatic regions. By this co-targeting of all four factors, reprogramming promoting genes like Glis1 (Maekawa et al., 2011), the microRNA mir-302/367 cluster (Anokye-Danso et al., 2011) and other genes involved in MET are activated (Li et al., 2010). Expression of miR-302b posttranscriptionally stabilizes the transition to an epithelial phenotype via inhibition of the TGFβ pathway through binding to the mRNA of its receptor Tgfbr2 and thereby preventing the reversal process of an epithelial to mesenchymal transition (EMT) (Subramanyam et al., 2011).

However, pioneer factors cannot bind to H3K9me3 marked heterochromatin regions (Soufi et al., 2012). Interestingly, many late pluripotency genes such as Nanog, Dppa4, Sox2, Gdf3 and Prdm14 are located in these regions (Buganim et al., 2012). It seems that H3K9me3 regions represent a barrier for the transition to a deterministic phase of reprogramming, since once obtained, endogenous expression of Nanog and Sox2 initiates a whole cascade of further gene activations, such as Dppa2, Dppa3 and Gdf3, to stabilize pluripotency (Ng and Surani, 2011; Buganim et al., 2012; Golipour et al., 2012; Polo et al., 2012; Soufi, 2014). Bone morphogenic proteins (BMPs) in serum could be identified to arrest the cells in an intermediate pre-iPSC state by activating H3K9 methyltransferases, whereas addition of vitamin C weakens this effect by regulating corresponding demethyltransferases (Chen et al., 2013a). Furthermore, the knock down of histone H3K9 methyltransferase SUV39H1 enhances efficiency and kinetics of reprogramming (Onder et al., 2012; Sridharan et al., 2013).

Among the initially activated genes a bias for c-Myc targets can be detected (Koche et al., 2011). Those sites are located in regions that either retain the open chromatin mark H3K4me2 if it is already present in the starting fibroblasts, or gain this mark de novo during early reprogramming (Koche et al., 2011; Soufi, 2014). c-Myc also co-binds to OSK helping to activate bound genes by opening the chromatin structure as it is thought to recruit histone acetylase complexes (Adhikary and Eilers, 2005) and RNA polymerase II (Nie et al., 2012). It also plays a role in amplification of gene expression, cell cycle progression and apoptosis (Rahl et al., 2010; Lin et al., 2012; Nie et al., 2012). In general, it appears that c-Myc accelerates reprogramming by opening up the chromatin, promoting cell cycle and survival, thus increasing chances for stochastic events to occur (Knoepfler, 2008; Hanna et al., 2009).

Opening chromatin seems to be also a reason why small molecules, like the histone deacetylase inhibitor (HDAC) valproic acid (VPA) and sodium butyrate, facilitate this first

stochastic step of reprogramming by randomly destabilizing the chromatin (Huangfu et al., 2008a; Mali et al., 2010). Blocking apoptosis and senescence, which is the ultimate fate of the majority of cells, seems also to enhance the stochastic process as it holds the cells in a pending trial and error phase until they might find the right route to pluripotency.

Overcoming the barrier of H3K9me3 leads to a more sequential or hierarchical phase in reprogramming. Activation of Sox2 and Nanog initiates a series of consecutive steps leading to the pluripotency state by activation of other late onset pluripotency genes including Esrrb, Utf1, Lin28, and Dppa2. Those genes are, in contrast to early pluripotency associated genes, predictive for complete reprogramming (Buganim et al., 2012).

In general, the chromatin of pluripotent cells is predominantly in an “open” and hyper-dynamic state when compared to somatic cells (Meshorer et al., 2006; Gaspar-Maia et al., 2011; Kobayashi and Kikyo, 2015). Heterochromatin gets predominantly hypomethylated at satellite repeats (Maherali et al., 2007) and the chromatin state is preserved by balancing the exchange of histone and non-histone proteins, like hetrochromatin protein 1 (HP1), linker histone H1°, core histones H2B and H3 as well as antagonistic processes like addition and removal of methylation marks (Shipony et al., 2014). The open nature of iPSC/ESC chromatin is also reflected in a global transcriptional hyperactivity (Efroni et al., 2008). In case of differentiation-related genes, pluripotent cells establish H3K27me3 repressive or bivalent histone modifications, resulting in a silenced expression with the possibility for rapid activation (Bernstein et al., 2006). The predominately open chromatin includes also telomeric and subtelomeric regions that are enriched in somatic cells in repressive histone trimethylation of H3K9 and H4K20, binding of HP1 and DNA methylation. Loss of this epigenetic marks as well as the enhanced transcriptional activity of telomerase leads to a telomere elongation when reaching pluripotency (García-Cao et al., 2004; Gonzalo et al., 2006; Marion et al., 2009).

Not only expression and chromatin reorganization occurs in waves during reprogramming, but also cell proliferation, metabolism, DNA repair and RNA processing (Hansson et al., 2012; Polo et al., 2012). On the protein level, epithelial proteins like Meis1, Nfact and Cript get transiently up-regulated during reprogramming in order to reach MET before later on a network of proteins regulating DNA replication and repair, chromatin homeostasis, gene expression and cell adhesion is up-regulated (Huang et al., 2012; Hansson et al., 2012). All those changes manifest in a shrinking cell size and a rapid shift in cell division (Smith et al.,

2010). Considering these enormous, multilevel rearrangements that occur during reprogramming to pluripotency, it is reasonable that aberrations could take place, which are described in more detail in the next chapter.

1.2.2 Aberrations occur during reprogramming to pluripotency

Several studies in the mouse and human system investigated the differences between ESCs and iPSCs. Although iPSCs and ESCs are very similar in their expression of pluripotency genes, their ability to differentiate into all three germ layers and their typical morphologic characteristics, there are indications that those two populations can differ in their transcriptional profile, epigenetic signatures and differentiation potential. Global transcriptional analysis of several iPSC and ESC lines revealed deviations in hundreds to thousands of genes (Chin et al., 2009; Marchetto et al., 2009). Mounting evidence suggested that in part this was due to a somatic cell memory specific of their tissue of origin. Comparing iPSCs derived from different cell sources revealed differences in their transcriptional profile as well as their epigenetic pattern that could be traced back to their parental cells (Chin et al., 2009; Ohi et al., 2011; Bar-Nur et al., 2011; Kim et al., 2011b; Roost et al., 2017). This so-called epigenetic memory relates to the somatic lineage but also the age of the cell of origin and seems to be transient and more prominent in early passages of induced pluripotent cells (Marchetto et al., 2009; Ghosh et al., 2010; Nishino et al., 2011; Lo Sardo et al., 2017). Somatic epigenetic memory leads also to altered differentiation propensities of the iPSCs by promoting their specification along lineages related to the donor cell, while limiting alternate cell fates (Feng et al., 2010; Polo et al., 2010; Kim et al., 2010). Human fibroblast-derived iPSCs, for instance, show a reduced efficiency and increased variability in their neural differentiation capacity when compared to ESCs (Hu et al., 2010), and human pancreatic islet beta cell-derived iPSCs demonstrated an increased ability to differentiate into insulin-producing cells *in vitro* and *in vivo* (Bar-Nur et al., 2011). Furthermore, analyzing DNA methylation pattern, iPSCs show differentially methylated regions (DMRs) that are dissimilar to both the respective cells of origin and ESCs whereas the majority was hypomethylated, which may indicate a deficiency in reestablishing DNA methylation during reprogramming (Lister et al., 2011; Huang et al., 2013; Nishizawa et al., 2016). The posttranscriptional regulation of gene expression is another epigenetic mechanism that can be affected during reprogramming. For example, the microRNA miR-

371/372/373 cluster was identified to be differentially expressed in human ESCs and iPSCs (Wilson et al., 2009).

Whole genome expression profiles of iPSCs and ESCs also indicated instability of imprinted gene expression in some iPSC clones (Pick et al., 2009; Stadtfeld et al., 2010; Teichroeb et al., 2011). Genomic imprinting is an epigenetic phenomenon whereby genes of certain clusters are expressed in a parent-of-origin dependent, monoallelic manner (Barlow and Bartolomei, 2014). This process is essential for normal development in vertebrates and abnormal biallelic expression of imprinted genes is correlated with several human diseases and malignancies (Cerrato et al., 2008; Chak et al., 2017). In mouse iPSCs a few transcripts of the imprinted Dlk1-Dio3 cluster were identified to be aberrantly expressed which led to an altered developmental propensity as shown by poor contribution to chimeras and the disability to generate entirely iPSC-derived animals (Stadtfeld et al., 2010). Moreover, human female iPSCs predominantly do not undergo X-chromosome reactivation during reprogramming and retain the inactive X-chromosome of the somatic cell of origin at low passages (Tchieu et al., 2010; Pomp et al., 2011). This X-chromosome-inactivation (XCI) may be partially erased over time in culture and cannot be reestablished by either differentiation or further reprogramming (Mekhoubad et al., 2012). On protein level, mass spectrometry revealed that differences on transcript level in part are also transferred to the proteome and phosphoproteome (Phanstiel et al., 2011; Munoz et al., 2011).

The reprogramming process also has the potential to compromise the genomic integrity of iPSCs (Ji et al., 2012; Bhutani et al., 2016; Rouhani et al., 2016; Yoshihara et al., 2017). Genetic variations occur most prominent at initial stages (Hussein et al., 2011; Sugiura et al., 2014) and can result in deletions of tumor suppressor genes (Mayshar et al., 2010; Laurent et al., 2011; Gore et al., 2011). At last, transplantation assays in mice might indicate substantial degrees of immunogenicity and tumorigenicity of iPSCs. On the one hand, autologous iPSCs seem to be less immune-tolerated by the recipient than ESCs and consequently less efficient in teratoma formation, as the majority of teratomas formed by iPSCs show a high degree of T cells infiltration (Zhao et al., 2011). On the other hand, an elevated tumor formation propensity could be detected in iPSC-chimera progeny, due to c-Myc transgene reactivation, and in transplatation assays of iPSC-derived neural progenitors due to higher degree of persisting undifferentiated cells (Okita et al., 2007; Miura et al., 2009).

1.3 Aim of the study

Human induced pluripotent stem cells (iPSCs) could provide a continuous donor source for the generation of specific somatic cell types for disease modeling or neural cell therapy. Considering the harsh reprogramming procedure and the requirement of epigenetic rearrangement for proper reprogramming, a critical question to be addressed is whether and to what extent alterations caused by the reprogramming process can confound readout parameters in disease modeling or influence the clinical safety of human induced pluripotent stem cell (iPSC)-derived cell populations. Several studies have so far mostly investigated the pluripotent state via comparison of iPSCs with embryonic stem cells. In the mouse system a few studies went into more detail analyzing also the influence of reprogramming on iPSC-derived somatic cell populations. However, it remains unclear whether those discoveries hold also true in the human system. Moreover, many studies on human cells are based on cells with different genetic backgrounds, yet, as research in the mouse system has already shown, an isogenic background is substantial to detect also minor alterations. The main objective of this study is therefore to generate a human isogenic system to investigate reprogramming-associated aberrations in an iPSCs-derived somatic cell population. In this study, human embryonic stem cell (ESC)-derived neural stem cells (ESC-NSCs) are used as a well-characterized, highly standardized, stable somatic cell population. This well-defined stem cell population is reprogrammed into iPSCs, which are subsequently re-differentiated into the very same somatic cell type. After successful characterization of all cell populations on specific marker expression on mRNA and protein level, further analysis of their differentiation aims to dissect any major differences in their developmental potential.

Having established a stable isogenic cell culture model, global expression and methylation profiling performed on both neural stem cell populations and their parental, pluripotent cells of origin are aimed to elucidate any transcriptional or epigenetic alteration that may influence their biomedical application. By also analyzing the somatic iPSC-derived cell populations a further insight should be gained whether alterations that might occur in the pluripotent state are inherited to their progeny.

2 Material

2.1 Technical equipment

Device	Name	Manufacturer
Autoclave	D-150	Systec
Balance	BL610	Sartorius
Balance	LA310S	Sartorius
Block heater	Thermomixer compact	Eppendorf
Centrifuge (cell culture)	Megafuge 1.0R	Sorvall /Kendro
Centrifuge (table top)	5415D	Eppendorf
Concentrator	Speed Vac 5301	Eppendorf
Counting chamber	Fuchs-Rosenthal	Faust
Digital camera	C 5050 Zoom	Olympus Optical
Digital camera	Power Shot G5	Canon
FACS® analyzer	Analyzer FACS® Calibur	BD Biosciences
Flourescence microscope	Axioskop 2	Carl Zeiss
Freezer -80°C	HERAfreeze	Kendro
Gel electrophoresis chamber	Agagel	Biometra
Horizontal flow hood	HERAsecure	Kendro
Imaging system	Geldoc EZ	Bio-Rad
Incubator	HERAcell	Kendro
Inverse light microscope	Axiovert 25	Carl Zeiss
LED light source	Colibri 2	Carl Zeiss
Liquid nitrogen store	MVE 611	Chart Industries
Microscope	Axiovert 40 CFL	Carl Zeiss
Microscope	Axiovert 200M	Carl Zeiss
Microscope	Axioskop 2	Carl Zeiss
Microscope	IX 81 Confocal	Olympus
Microscope camera	Axiocam MRM	Carl Zeiss
Microscope camera	ProgRes C14	Jena Optic
Microscope laser	Laser Mells Griot	Griot Lasergroup
Micro-Specrophotometer	Nanodrop ND-1000	Thermo Scientific
PCR cycler	T3000 Termocycler	Biometra
pH-meter	CG840	Schott

Pipetteboy	Accu-Jet	Brand
Power supply for elecrophoresis	Standard Power Pack P25	Biometra
Real-time qPCR mashine	Mastercycler realplex	Eppendorf
Refrigerator	G 2013 Comfort	Liebherr
Shaker	Bühler WS 10	Johanna Otto
Sterile laminar flow hood	HERAsafe	Kendro
Stereo microscope	STEMI 2000-C	Carl Zeiss
Thermocycler	T3 Thermocycler	Biometra
Table centrifuge	Centrifuge 5415R	Eppendorf
Ultracentrifuge	Discovery 90SE	Sorvall
UV-Vis Spectrophotometer	BioPhotometer	Eppendorf
Vacuum pump	Vacuubrand	Brand
Vortexer	Vortex Genie 2	Scientific Industries
Water bath	1008	GFL
Water filter	Milipak 40	Millipore

2.2 Labority consumables

Item	Manufacturer
4-well culture dishes	Nunc
6-well culture dishes	Nunc
12-well culture dishes	Nunc
24-well culture dishes	Nunc
Cryovials 1 ml	Nunc
Cryovials 1.8 ml	Nunc
PCR strip tubes 0.2 ml	peqLab
Petri dishes ⌀ 3.5 cm	BD Biosciences
Petri dishes ⌀ 6 cm	BD Biosciences
Petri dishes ⌀ 10 cm	PAA
Serological Pipettes 1 ml	BD Biosciences
Serological Pipettes 2 ml	BD Biosciences
Serological Pipettes 5 ml	BD Biosciences
Serological Pipettes 10 ml	BD Biosciences
Serological Pipettes 25 ml	BD Biosciences

Syringes 20 ml	BD Biosciences
Syringe filter 0.2µm	PALL
Syringe filter 0.45 µm	PALL
Tissue culture dishes ⌀ 3.5 cm	PAA
Tissue culture dishes ⌀ 6 cm	PAA
Tissue culture dishes ⌀ 10 cm	PAA
Tissue culture dishes ⌀ 15 cm	PAA
Tubes 0.5 ml	Greiner Bio-One
Tubes 1.5 ml	Greiner Bio-One
Tubes 2.0 ml	Greiner Bio-One
Tubes 15 ml	Greiner Bio-One
Tubes 50 ml	Greiner Bio-One

2.3 Chemicals and reagents

Chemicals	Manufacturer
2-Mercaptoethanol	Thermo Fisher Scientific
2-Phospho-L-ascorbic	Sigma-Aldrich
8-Cl-cAMP	Enzo
Agar	Sigma-Aldrich
Agarose	PeqLab
Ampiciline	Sigma-Aldrich
Ascorbic acid	Sigma-Aldrich
B-27 supplement	Thermo Fisher Scientific
Bromphenol blue	Sigma-Aldrich
BSA solution (7.5%)	Sigma-Aldrich
CaCl2	Sigma-Aldrich
cAMP	Sigma-Aldrich
Chloroform	Sigma-Aldrich
Chloroquin	Sigma-Aldrich
Collagenase type IV	Thermo Fisher Scientific
CytoconTM buffer II (Cyto-buffer)	Evotec
DAPI	Sigma-Aldrich
DMEM/F12	Thermo Fisher Scientific

DMEM high glucose	Thermo Fisher Scientific
DMSO	Sigma-Aldrich
DNA ladder (100 bp)	PeqLab
DNA ladder (1 kb)	PeqLab
dNTPs	PeqLab
Doxycycline	Sigma-Aldrich
D-PBS	Thermo Fisher Scientific
EDTA	Sigma-Aldrich
EGF	R&D Systems
Ethanol	Merk
Ethidium bromide	Sigma-Aldrich
FCS	Thermo Fisher Scientific
FGF2 (ESC medium)	Thermo Fisher Scientific
FGF2 (NSC medium)	R&D Systems
G418 solution	Sigma-Aldrich
Gelatin	Thermo Fisher Scientific
Glucose	Sigma-Aldrich
Glycerol	Sigma-Aldrich
Glycine	Sigma-Aldrich
HCl	Sigma-Aldrich
HEPES	Sigma-Aldrich
Insulin	Sigma-Aldrich
Isopropanol	Sigma-Aldrich
Knockout-DMEM	Thermo Fisher Scientific
Laminin	Sigma-Aldrich
L-Glutamine	Thermo Fisher Scientific
Matrigel	BD Biosciences
Methanol	Carl Roth
Mowiol	Sigma-Aldrich
Myo-Inositol	Sigma-Aldrich
N2-supplement	Thermo Fisher Scientific
N2-supplement	PAA
NaCl	Sigma-Aldrich

NaHCO3	Sigma-Aldrich
NaOH	Sigma-Aldrich
Neurobasal medium	Thermo Fisher Scientific
Non-essential amino acids	Thermo Fisher Scientific
Opti-MEM basal media	Thermo Fisher Scientific
Paraformaldehyde (PFA)	Sigma-Aldrich
Penicillin-Streptomycin	Thermo Fisher Scientific
Phenol	Sigma-Aldrich
Poly-L-ornithin	Sigma-Aldrich
Polyvinylalcohol	Sigma-Aldrich
Serum replacement	Thermo Fisher Scientific
Sodium azide	Sigma-Aldrich
Sodium pyruvate	Thermo Fisher Scientific
Sucrose (60%)	Carl Roth
Transferrin	Millipore
Tris	Merk
Triton-X100	Sigma-Aldrich
Trypan blue	Thermo Fisher Scientific
Trypsin-EDTA (10x)	Thermo Fisher Scientific
Yeast extract	Sigma-Aldrich

2.4 Cell culture stock solutions

Reagent	Concentration	Solvent
EGF	10 µg/ml	0.1 M Acedic acid, 0.1% BSA
FGF2	10 µg/ml	PBS, 0,1% BSA
G418	50 mg/ml	H2O

2.5 Cell culture solutions

Poly-L-ornithine (PO) coating solution	%
H2O	99
Poly-L-ornithine (1.5 mg/ml stock)	1
Mixed, sterile filtered and stored at 4°C	

Laminin (Ln) coating solution	%
H2O	100
Laminin	1 µg/ml

Matrigel (MG) coating solution	%
DMEM/F12	97
Matrigel	3

0.1 % Gelatin	%
H2O	99.9
Gelatin	0.1
Mixed, autoclaved and stored at 4°C	

2x HBS buffer	
NaCl	8 g
KCl	0.38 g
Na2HPO4	0.1 g
HEPES	5 g
Glucose	1 g
H2O	Add to 500 ml
Adjust to pH 7.05, sterile-filter and store at -20°C	

Cyto buffer	
Myo-Inositol	43.25 g
H2O	800 ml
PBS	200 ml
Polyvinylalcohol	5 g in small portions and continous stirring, if necessary heat up to max. 40°C
sterile-filter and store at 4°C for max. 4 months	

1x Trypsin/EDTA (TE)	%
PBS	90
Trypsin/EDTA (10x)	10

Trypsin inhibitor (TI)	%
PBS	100
Trypsin inhibitor	0.25mg/ml (=700 units/mg)
Mixed, sterile filtered and stored at 4°C	

2.6 Cell culture media

Freezing media	DMSO	Additive(s)
NSC freezing medium	10 %	Knockout-Serum replacement 70%, Cyto buffer 20%
iPSC/ESC freezing medium	10 %	Knockout-Serum replacement 90%, 5 µM ROCK inhibitor
MEF freezing medium	10 %	FCS (heat inactivated) 90%
Virus freezing medium (VFM)	-	HEPES buffered solution (HBS) 92%, BSA solution (7.5%) 8%

MEF medium	%
DMEM high glucose	86
FCS (heat inactivated)	10
Sodium pyruvate	1
L-Glutamine	1
Non-essential amino acids	1
Pen/ Strep	1

Neural stem cell (N2) medium	%
DMEM/F12	97.5
N2 suplement	1
Pen/ Strep	1
D-Glucose	1.5 mg/ml

Neuronal generation (NGMC) medium	%
Neural stem cell (N2) medium	50
Neurobasal medium	48
B27 supplement	2
Pen/ Strep	0.5
cAMP	100 ng/ml

ESC medium	%
Knockout-DMEM	78
Knockout-Serum replacement	18.6
Non-essential amino acids	1
L-Glutamine	1
2-Mercaptoethanol	0.4
Pen/ Strep	1
FGF2	4 ng/ml

EB medium	%
Knockout-DMEM	78
Knockout-Serum replacement	19
Non-essential amino acids	1
L-Glutamine	1
Pen/ Strep	1

2.7 Cell lines and strains

Cell line, bacterial strain or mouse strain	Source
DH5a E.coli	Thermo Fisher Scientific
CD1 mouse strain (primary embryonic fibroblasts)	Charles River
SCID/beige mouse strain	Charles River
Phoenix-GP	ATCC
HEK-293FT	Life Technologies
H9.2 hESC line	Haifa, Israel (Amit et al., 2000)
I3 hESC line	Haifa, Israel (Amit et al., 2000)
I6 hESC line	Haifa, Israel (Amit et al., 2000)
SAi5 primary NSCs	Smith lab, Cambridge, UK (Sun et al., 2008)

2.8 Kits

Name	Manufacturer
Alkaline Phosphatase Substate Kit III	Vector laboratories
CytoTune-iPS Sendai Reprogramming Kit	Thermo Fisher Scientific
DNeasy Blood and Tissue Kit	Qiagen
DNA Clean & Concentrator Kit	Zymo Research
RNeasy Mini Kit	Qiagen
iScipt cDNA Synthesis Kit	Bio-Rad
PeqGOLD Plasmid Miniprep Kit I	Peqlab
HiPure Plasmid Filter Maxiprep Kit	Thermo Fisher Scientific

2.9 Enzymes

Enzyme name	Manufacturer
Alkaline Phosphatase	New England Biolabs
DNase (cell culture)	Cell Systems
DNase I	Thermo Fisher Scientific
GoTaq® Flexi DNA Polymerase	Promega
Phusion High Fidelity Polymerase	Finnzymes
Taq DNA Polymerase, recombinant	Thermo Fisher Scientific
NdeI restriction endonuclease	New England Biolabs
NotI restriction endonuclease	New England Biolabs
SpeI restriction endonuclease	New England Biolabs

2.10 Primers

PCR primers		
Target	Forward	Reverse
GAPDH	ATGACCCCTTCATTGACCTCAACT	ATGGGTGTGAACCATGAGAAGTAT
BRIX	CACCACGGTATCATCCCAAAAGCCAACC	ACGCCGATGCATGTTTGGTGACTGGTAG
DNMT3b	TGCTGCTCACAGGGCCCGATACTTC	TCCTTTCGAGCTCAGTGCACCACAAAAC
FGF4	CTACAACGCCTACGAGTCCTACA	GTTGCACCAGAAAAGTCAGAGTTG
GDF3	CTTATGCTACGTAAAGGAGCTGGG	GTGCCAACCCAGGTCCCGGAAGTT
LEFTB	CTTGGGGACTATGGAGCTCAGGGCGAC	CATGGGCAGCGAGTCAGTCTCCGAGG
NANOG	CAGCCCCGATTCTTCCACCAGTCCC	CGGAAGATTCCCAGTCGGGTTCACC

NODAL	GGGCAAGAGGCACCGTCGACATCA	GGGACTCGGTGGGGCTGGTAACGTTTC
REX1	CAGATCCTAAACAGCTCGCAGAAT	GCGTACGCAAATTAAAGTCCAGA
TERT	CCTGCTCAAGCTGACTCGACACCGTG	GGAAAAGCTGGCCCTGGGGTGGAGC
TDGF	CTGCTGCCTGAATGGGGGAACCTGC	GCCACGAGGTGCTCATCCATCACAAGG
UTF1	CCGTCGCTGAACACCGCCCTGCTG	CGCGCTGCCCAGAATGAAGCCCAC
SRY	AGAGTGAAGCGACCCATGAA	TGTTGTCCAGTTGCACTTCG
OCT4 endo	GACAGGGGGAGGGGAGGAGCTAG	GTTCCCTCCAACCAGTTGCCCCAAAC
OCT4 trans	GTACTCCTCGGTCCCTTTCC	TGACACCAGACCAACTGGTAA
SOX2 endo	GTATCAGGAGTTGTCAAGGCAGAG	TCCTAGTCTTAAAGAGGCAGCAAAC
SOX2 trans	CATGTCCCAGCACTACCAGA	TGACACCAGACCAACTGGTAA
cMYC endo	TTCGGGTAGTGGAAAACCAG	CCTCCTCGTCGCAGTAGAAA
cMYC trans	GGAAACGACGAGAACAGTTGA	TGACACCAGACCAACTGGTAA
KLF4 endo	GACCAGGCACTACCGTAAACA	CTGGCAGTGTGGGTCATATC
KLF4 trans	GACCACCTCGCCTTACACAT	TGACACCAGACCAACTGGTAA
SeV trans	GGATCACTAGGTGATATCGAGC	ACCAGACAAGAGTTTAAGAGATATGTATC

qPCR primers		
Target	Forward	Reverse
18S rRNA	ATTCTTGGACCGGCGCAA	CCGACCGGCGATGCGGC
GAPDH	CTGCTTTTAACTCTGGTAAAGT	TGGGGCGATGCTGGCGC
AFF2	ACCTGAAGCTGAGCAGTGAT	AACAGCCTGGTAGAGCTCAG
ANO5	GGCTGTCCTTTCTGTTGCAA	GCTGTGTGGCATTTGTTGAG
ATP1A2	TGGAAAGATGGGGAGCTCTG	CTCTCCCCAAATGTTGCCAC
CDH1	GAACGCATTGCCACATACAC	ATTCGGGCTTGTTGTCATTC
CSAG3	AGACTGTTGAGAGACGCTGG	CTCTTTGGTGTTGGTGGGTG
CXORF57	CCCACACGCTAATCCAGTTG	CATGCCCGGTTTAATTCCCT
EGR2	TCTGACAACATCTACCCGGT	GTTCATCTGGTCAAAGGGGC
GABRG2	ACTTCCGGAGATTATGTGGTCA	CACCCAGGATAGGACGACAA
H1F0	ACCAAGAAACCCAAAGCCAC	TTGGGTTTCACTGGTTTGGC
HEPH	CGTCCCTATACCATCCACCC	GGGGAACAGAGTCATCAGCT
HIST1H3I	CTGGTGGGGCTATTTGAGGA	CGCAAGCTGGATGTCTTTAGG
KLHL13	GCCTCTCATCCCATTTGCAG	GGTCAAAGCCCTGTAACACC
LGI1	TGGCTAGTGGAATGGCTTGG	GCGCTTCTTGTATTCTGGGG

MEGEA2B	GAGTCATCATGCCTCTTGAGC	CTCTCCTCGGGCCTCAAG
NANOG	TACCTCAGCCTCCAGCAGAT	TGCGTCACACCATTGCTATT
PDCH15	TCAGCTTTGAGTGAACTTTGACA	AGAGAGAGCCCAGGATGATC
PLIN2	GAGCTCCACGTATGACCTCA	TGCCATCTCACACACAGACT
PNPLA4	CCAACGCCAAAACCAGAGAA	TCCACTAGCTTCAGTCCTGC
TAAR3	CAAACAGCATGCTCGAGTCA	CGCTGCTTTCCTGTCCTTTT
TERT	TGTGCACCAACATCTACAAG	GCGTTCTTGGCTTTCAGGAT
TMEM255A	CACCTGCCACCATATTCTGC	CAGACTGGGGCTCATCAGAA
TOM1L1	ACTTGGTCACAGGGCTTCC	CGCCTTTCTTAACCAGGTCG
TSIX	GAGGGACAGAAGTTGGGACA	ATTCACTCCAAGGGCCATCA
SEPP1	CCTGGAGCATAAGAGATCAAGA	ACAGGTATCAGCTGGCTTGA
SHISA3	ATCGTCGGCTCCATCTTCAT	CTGATAGCTGCGGAGTGAGA
SYT11	CTCAAAGCCAGACACTTGCC	GTGCACTTCTTCACATGGGT
XIST 001	TTGGATTTGGCCTGCTGTTC	CAGGGACAGGCACAGAAAAG
ZDHHC15	GAAGATGGCTCTGTCTGGGG	CCAGGCAGAGTTCAAAGACG
ZNF215	CTGGAGAAGAATCATCCCATGG	AGGCATCTGAAGAAGGGTGT
8162183	CATAGTCAGGAGGCCAGGAG	ACAGGATGATGGCCCATAGG
FUW-OCT4-trans	TTTCCCCCTGTCTCTGTCAC	GGCATTAAAGCAGCGTATCC
FUW-KLF4-trans	GACCACCTCGCCTTACACAT	GGATTAAAGCAGCGTATCC
FUW-cMYC-trans	ACAGCTACGGAACTCTTGTGC	GGCATTAAAGCAGCGTATCC

Primers for allele-specific expression analysis				
Gene	SNP		Forward	Reverse
HEPH	rs806607	T/C	ACCAAGCTGCAAGAATCTACT	CCAGTGGCCAGACAGTAGTA
	rs809363	A/G	ACCAAGCTGCAAGAATCTACT	CCAGTGGCCAGACAGTAGTA
TMEM255A	rs5957345	T/C	GCCCTATCTTTCCCTACCCC	ACATCCTGTTCCCCACTACTG
TSIX	rs5981565	C/T	GCTGGTGGGAGAAAGTATGG	AAGGCAAGATCAGCTAATACCA
XIST	rs1620574	A/G	GTGTCTTGGGTAGCAGAAGAAAA	CCGAGCCCCACAGAAAGTAA

2.11 Plasmids

Plasmid name	Source
pMXs-OCT4	Addgene
pMXs-KLF4	Addgene
pMXs-cMYC	Addgene
pMD2.G	Addgene
psPAX2	Addgene
pGFP	Addgene
FUW-tetO-hOCT4	Addgene
FUW-tetO-hKLF4	Addgene
FUW-tetO-hMYC	Addgene
pLVX-Eto	Modified by Jerome Mertens from pLVX-Tet-ON-Advanced, Clonetech

2.12 Molecular biology buffers

DNA loading buffer (6x)	
0.5 M EDTA (pH8)	2 ml
Sucrose	6 g
2 % Bromphenol-blue solution	0.2 ml
2 % Xylene-cyanol solution	0.2 ml
Ficoll	0.2 g
H2O	3.8 ml
Mixed, aliquoted in 1ml and stored at 4°C	

50x Tris-acetate-EDTA-buffer (TAE)	
Tris	242 g
0.5 M EDTA (pH 8)	100 ml
Water-free acedic acid	57.1 ml
H2O	Add to 1000 ml
For use dilute with H2O 1:50	

TE buffer	
Tris-HCl	1.58 g
EDTA	0.29 g
H2O	Add to 1000 ml
Adjust to pH 8 with solid NaOH	

2.13 Reagents and media for microbiological methods

LB agar	
Tryptone	10 g
Yeast extract	5 g
NaCl	5 g
Agar	7 g
H2O	Add to 1000 ml
Adjust to pH 7.0 if necessary, autoclave and store at 4°C	

LB medium	
Tryptone	10 g
Yeast extract	5 g
NaCl	5 g
1 M NaOH	1 ml
H2O	Add to 1000 ml
Adjust to pH 7.0 if necessary, autoclave and store at 4°C	

SOB medium	
Tryptone	20 g
Yeast extract	5 g
NaCl	0.5 g
KCl	0.19 g
2 M MgCl2	5 ml
H2O	Add to 1000 ml
Adjust to pH 7.0 with 5 M NaOH, autoclave and store at 4°C	

2.14 Antibodies

Primary antibody	species	Isotype	Dilution	Source
AFP	rb	IgG	1:100	DAKO
b-III-Tubulin	ms	IgG	1:1500	Covance
b-III-Tubulin	rb	IgG	1:2500	Covance
DACH1	rb	IgG	1:50	ProteinTech
GABA	rb	IgG	1:600	Sigma
GFAP	rb	IgG	1:250	DAKO
H3K27me3	rb	IgG	1:1000	Millipore
MAP2ab	ms	IgG	1:250	Chemicon
Nestin	ms	IgG	1:300	R&D Systems
OCT3/4	ms	IgG	1:400	Santa Cruz
PLZF	ms	IgG	1:20	Santa Cruz
SMA	ms	IgG	1:100	DAKO
SOX2	ms	IgG	1:500	R&D Systems
TRA1-60	ms	IgM	1:500	Thermo Fisher Scientific
TRA1-81	ms	IgM	1:500	Thermo Fisher Scientific
ZO1	rb	IgG	1:50	Zymed

Secondary antibody	Dilution	Source
Alexa Flour 488 labeled anti-rabbit IgG	1:1000	Thermo Fisher Scientific
Alexa Flour 488 labeled anti-mouse IgG	1:1000	Thermo Fisher Scientific
Alexa Flour 555 labeled anti-rabbit IgG	1:1000	Thermo Fisher Scientific
Alexa Flour 555 labeled anti-mouse IgG	1:1000	Thermo Fisher Scientific
Alexa Flour 555 labeled anti-mouse IgM	1:1000	Thermo Fisher Scientific

2.15 Immunocytochemistry solutions

PFA fixation solution (4%)	
PFA	40 g
H2O	1000 ml
Heated until PFA dissolves completely, pH adjusted to 7.4, sterile filtered and stored at -20°C	

Immuno-blocking solution	%
PBS	89.9
FCS	10
Triton X100 (for intracellular epitopes)	0.1

Moviol	
Tris solution (0.2 M, pH8.5)	12 ml
H2O	6 ml
Glycerol	6 g
Moviol	2.6 g
DABCO	0.1 g

2.16 Software

Name	Manufacturer
ApE – AplasmidEditor 2.0.36	M. Wayne Davis
AxioVision 4.5	Carl Zeiss MicroImaging
FlowJo 6.8	Tree Star
Illumina BeadStudio Genome Studio 2011	Illumina
ImageJ 64	NIH
Microsoft Office 2008	Microsoft
Photoshop CS3	Adobe
Primer 3 v4.1.0	Rozen and Skaletsky (2000)
Prism 5	Graphpad
Quantity One	Bio-Rad

3 Methods

3.1 Cell culture methods

Generally, all cell culture experiments were performed under sterile conditions using a sterile hood with sterile media, glass and plastics. Cells were cultivated in an incubator at 37°C, 5% CO_2 and 95% humidity.

3.1.1 Thawing of cells and cell storage

Cryopreserved cells were thawed in a water bath at 37°C by gentle agitation until only a small ice crystal remained. Cells were immediately mixed with 10 ml pre-warmed culture medium and centrifuged at 212 x g for 5 minutes (pluripotent aggregates at 110 x g for 3 minutes) at 4°C. The supernatant was removed and the pelleted cells were carefully resuspended in an appropriate volume of growth medium (in case of pluripotent cells, 5 µM ROCK inhibitor was added) and seeded on an appropriate cell culture dish.

For cryopreservation, pelleted cells were gently resuspended in 1 ml of the according freezing medium, transferred to cryo-vials and frozen in two steps. First, the cryo-vials were placed in a freezing container (Nalgene) filled with isopropanol at -80°C to ensure slow freezing in a rate of 1°C/min. After 24h the tubes were transferred to liquid nitrogen for long time storage.

3.1.2 Cell number determination

20 µl of cell suspension was mixed 1:1 with trypan blue stain (GIBCO). This allows distinguishing between dead and living cells, as dead cells will take up the trypan blue and will appear blue under the microscope. In contrast, living cells don't take up the blue color and appear brightly light. Approximately 20 µl of the mixture was applied to a Fuchs-Rosenthal chamber via capillary forces between chamber and cover plate. Cells of at least four big squares were counted to determine an average value, which was filled in the following equation (1). If too many cells were present, the cell suspension was further diluted.

Cells per ml of suspension = average cell number per square x dilution factor x 10^4 (1)

3.1.3 Cultivation of pluripotent stem cells

Human pluripotent stem cells were cultivated with or without supporting fibroblasts, so-

called feeder cells. For propagation on feeder cells, one day prior usage irradiated mouse embryonic fibroblasts (MEF) were plated at a density of 1.2 x 10^6 cells/ml per gelatin (0.1%, 20 min) coated 6-well dish.

When cultivated on a layer of irradiated mouse embryonic fibroblasts (MEFs), serum-free ESC-medium was changed on a daily basis and cells were passaged every 4 to 5 days with collagenase IV (1 mg/ml, Thermo Fisher Scientific). After incubation with collagenase for 45 to 60 minutes the detached colonies were rinsed off the plates and pelleted at 800 rpm for 3 minutes at 4°C. While resuspending, the aggregates were triturated in small fragments and replated on new MEF cells in a ratio of 1:2 - 1:4.

When cultivated without feeder cells, as it was done prior to RNA and DNA isolation, pluripotent colonies were plated on matrigel-coated (MG) dishes in E8-medium, which was replaced on a daily basis. For passaging, the medium was removed and the plates were washed two times with pre-warmed PBS before applying PBS-EDTA. After 7 minutes the supernatant was removed and colonies washed off with E8-medium and directly plated on new MG-coated dishes.

Pluripotent cells could be frozen as single cells as well as colonies, when incubating the cells 1 hour before splitting with ROCK inhibitor (5µM) and resuspension in ESC-freezing medium.

3.1.4 Generation of embryoid bodies

For formation of embryoid bodies (EBs), pluripotent colonies were detached from MEF-coated plates with collagenase. In order to remove any remaining fibroblast cells, floating aggregates were washed three times. EBs were allowed to sink in a 15 ml polypropylene tube and washing media was subsequently removed. To avoid attachment to the culture dish, aggregates were cultivated on non-adherent plates and EB-medium was changed every other day by sedimentation.

3.1.5 Generation and cultivation of neural stem cells

In vitro differentiation of pluripotent stem cells into long term proliferating pluripotent stem cell-derived neuroepithelial stem (lt-NES) cells was performed as described previously (Koch et al., 2009). To induce differentiation, pluripotent cells were cultivated as embryoid bodies. After 4 days floating aggregates were allowed to adhere on polyornithine/laminin-coated (PO/Ln; Sigma Aldrich/Thermo Fisher Scientific) plates. Two days after attachment of the

EBs, the cultivation media was changed to N2 containing DMEM/F12 supplemented with FGF2 (10 ng/ml, R&D Systems) and changed every other day. After 10 days neural tube-like structures were mechanically separated from the EB-outgrowth and isolated for propagation as neurospheres in NSC-media. After 2 to 3 days floating spheres were incubated with trypsin EDTA for 5 minutes and mechanically triturated to single cells after addition of trypsin inhibitor (TI). Cells were pelleted at 1000 rpm for 5 min at 4°C and seeded and cultivated as a monolayer on Po/Ln-coated dishes in N2 supplemented medium containing FGF2, EGF (both 10 ng/ml, R&D Systems) and 1 µl/ml B27 supplement. Upon confluence, NSCs were passaged 1:2 or 1:3 via incubation for 2-5 min at 37°C with trypsin EDTA. Reactions were stopped by addition of TI and cells were centrifuged at 1000 rpm for 5 minutes, resuspended in culture medium and plated on new PO/Ln-coated dishes. Medium was changed on a daily basis for the first 5 passages. Thereafter, it was alternately either supplemented with EGF/FGF2/B27 or changed. NSCs could be cryopreserved when resuspended in neural stem cell freezing medium.

3.1.6 Differentiation of neural stem cells

To induce terminal differentiation, NSCs were cultured in NGMC medium without EGF and FGF2 on MG-coated dishes. Medium was changed every other day. After 6 weeks of differentiation cells were analyzed.

3.1.7 *In vitro* differentiation of pluripotent cells into three germ layers

Embryoid bodies were plated on gelatin-coated dishes after propagation for at least 4 days in EB medium. Aggregates were allowed to adhere and cultivation media was changed to MEF medium. Following 10 days of differentiation, outgrowth of diverse cell types could be detected.

3.1.8 *In vivo* differentiation of iPSCs into three germ layers

After detaching iPSC colonies with collagenase, aggregates were rinsed three times to remove any remaining feeder cells and cultivated on MG-coated dishes in E8-medium. After one passage, cells were detached as single cells with alphazym or accutase (Thermo Fisher Scientific). Around 1 x 10^6 single cells were resuspended in 250 µl of PBS and injected in the testis of immune-compromised SCID/beige mice. In the following one to three months

tumor formation could be detected and teratomas were removed. Transplantation of cells and isolation of teratomas was performed by Anke Leinhaas at the Haus für Experimentelle Therapie (HET), University of Bonn. Isolated teratomas were fixed in 4% PFA, embedded in paraffin, cut into sections and processed to histology examination using standard haematoxylin/eosin staining in cooperation with the Institute for Neuropathology, University of Bonn.

3.1.9 Generation of murine embryonic fibroblasts

Pluripotent stem cells can be cultured on a layer of murine embryonic fibroblasts (MEFs), also called feeder cells, which provide nutritional and cellular support for pluripotent cells. Prior to co-cultivation, MEFs were expanded and mitotically inactivated. Primary fibroblasts were isolated from CD1 mouse embryos (E13.5-E14.5) by Anke Leinhaas at the Haus für Experimentelle Therapie (HET), University of Bonn. For expansion, an aliquot of those primary cells was plated on three 15 cm dishes in MEF medium. Upon confluence, cells were rinsed with PBS and detached with trypsin/EDTA at 37°C. Adding FCS-containing MEF-medium stopped the trypsin activity and cells were centrifuged at 1000 rpm for 5 minutes. The pelleted cells were resuspended and plated in a ratio of 1:3 on new dishes. Following the third passage, cells were gamma-irradiated using a RS 2000 X-Ray biological irradiator (Rad Source) in a total volume of 30 ml MEF-medium in a T125 flask with a single dose of 15Gy. Thereafter, aliquots of 2.4x10^6 cells were cryopreserved in MEF freezing medium for later use.

3.1.10 Reprogramming of NSCs to pluripotency

NSCs were reprogrammed by retroviral, lentiviral or Sendai-viral transfection.

For retroviral reprogramming, the plasmids pMXs-OCT3/4, pMXs-KLF4, pMXs-cMYC, and pMXs-GFP (Addgene) were used to generate infectious particles. NSCs were incubated with either a freshly generated or a frozen cocktail of retroviruses supplemented with polybrene (4µg/ml) for 16 h. Medium was replaced every day. Two days after transduction, NSCs were detached with trypsin EDTA/trypsin inhibitor and plated at different densities on irradiated feeder cells, with supplementation of ROCK inhibitor (5µM). Medium was changed to ESC-medium in dilution steps of 50%, 75% and 100%. Afterwards, medium was changed every other day until colonies were mechanically isolated. As a transduction control a separate

well of NSCs were infected with GFP-virus at the same time and analyzed by flow cytometry two days post transfection.

For inducible expression via the lentiviral TetON-system, first stable cell lines harboring the reprogramming factors were generated. NSCs were transduced with the pLVXTP-Tet-On (Clontech) lentivirus, when full confluency was reached. Transduceded cells were subsequently selected with G418 (200 µg/ml, PAA) for several days to generate a stable rtTAAdv-expressing cell line. NSCs were then transduced with FUW-tetO-hOCT4 (Addgene) viral particles and selected with puromycin (1 µg/ml, Sigma Aldrich) to generate a stable inducible OCT4-NSC line. This line was subsequently infected with FUW-tetO-hKLF4 (Addgene). After several days of stable cultivation, doxycycline (1 µg/ml, Sigma-Aldrich) was added to the medium to induce transgene expression. After 24 hours, cells were detached with trypsin EDTA/trypsin inhibitor and plated at different densities on feeder-coated plates with supplementation of ROCK inhibitor (5µM) and doxycycline. Medium was changed to ESC-medium in dilution steps of 50%, 75% and 100% and changed every other day. Doxycycline was withdrawn upon colony formation after 2-3 weeks. Doxycycline-independent colonies were mechanically isolated and expanded.

In case of integration-free reprogramming with Sendai-virus, NSCs were infected with equal amounts of Sendai-viral particles of OCT4, KLF4 and c-Myc at 80% confluency by spinfection at 32°C; 1500 x g, 30 min with soft acceleration and deceleration settings followed by incubation for 8-14 hours. Medium was replaced every day. Two days post transduction, NSCs were detached with trypsin EDTA/trypsin inhibitor and plated at different densities on feeder-coated plates with supplementation of ROCK inhibitor (5µM). Medium was changed to ESC-medium in dilution steps of 50%, 75% and 100% and changed every other day until colonies were mechanically isolated. As a transfection control a separate well of NSCs was infected with GFP-Sendai virus at the same time and analyzed by flow cytometry two days post transfection.

3.1.11 Production of viral particles

For production of lentiviral particles using the FUW-tetON-plasmid-system, HEK-293FT cells were cultured in MEF medium and transfected with viral plasmids by calcium phosphate precipitation. 293-FT cells were plated on polyornithine/gelatin-coated 15 cm dishes at a density of 14 x 10^6 cells one day prior transfection. One hour before transfection, medium

was replaced with 16 ml pre-warmed MEF-medium. 1.4 ml DNA-mix (15 µg respective transfer vector, 3.5 µg envelope plasmid pMD2.G, 7.5 µg packaging plasmid psPAX in H2O ddest) was added to 178 µl of ice-cold 2.5 M CaCl2 solution and subsequently 1.4 ml of 2 x HBS buffer was slowly added under constant mixing and incubated for 45 minutes. In the meantime, 8 µl of 25 mM chloroquine (5 min before addition of the HBS-DNA mix) was added to the 293 cells. HBS-DNA mix was resuspended and transferred to the cells by gentl dropping. Medium was changed after 12 hours. After 48 h, virus particles were harvested the first time by collecting the supernatant and medium was replaced with pre-warmed MEF-medium. Supernatant was stored at 4°C and pooled with the supernatant of the second collection another 24 hours later. Supernatant was filtered through a 0.45µm Cellulose-Acetate-filter prior to virus concentration.

Concentration of viral particles was either performed by ultracentrifugation at 44,000 rcf and 4°C for 120 min, or by PEG-precipitation. For concentration of viral particles via PEG-precipitation, 10 ml 50% PEG 6000 solution was added to 40 ml filtered vector-containing cell culture supernatant. Furthermore, 4.25 ml of 4 M NaCl, and 4.6 ml PBS was added. During incubation at 4°C for 1.5 hours, samples were mixed every 20-30 minutes. Afterwards, samples were centrifuged at 7000 x g for 10 minutes at 4°C. The concentrated viral pellet was resuspended in virus freezing medium (VFM) and stored at -80°C. For the production of retroviral particles using the pMXs-vectors, PhoenixGP cells were used as producing cell line and the packaging plasmid psPAX was omitted.

3.1.12 Flow Cytometry

In parallel to the non-selectable reprogramming approaches, GFP-containing viral particles were applied to a control population of NSCs and analyzed by flow cytometry. In brief, GFP-transfected NSCs and non-transfected NSCs were trypsinized to generate a single cell suspension and resuspended in PBS. When applied to the FACSCalibur cytometer (BD) at least 30,000 events for each condition were analyzed with respect to forward and sideward light scatter and their EGFP fluorescence intensity. For visualization and analysis, the obtained data was transferred to FlowJo software.

3.2 Molecular biological methods

3.2.1 Large-scale DNA preparation

Plasmid DNA was purified using the HiPure Plasmid Filter Maxiprep Kit (Thermo Fisher Scientific). For plasmid purification an appropriate volume of LB medium (500 ml) supplemented with the appropriate antibiotic was inoculated either with a single colony picked from a recently streaked LB-agar plate or with a freshly prepared transformation assay. The culture was grown overnight (12-16 h) at 37°C and 120 rpm. For harvesting, the bacterial cells were centrifuged at 6000 x g for 15 minutes at 4°C. The supernatant was carefully discarded and purification of the plasmid DNA was performed according to the manufacturer's protocol for maxi preparation.

3.2.2 Endonuclease Restriction

Restriction endonucleases recognize short DNA sequences and cut double-stranded DNA at certain locations within or outside of these recognition sites. Recognition sites are normally palindromic DNA sequences with a length of 4 to 8 base pairs in length. Dependent on the used restriction enzyme, either blunt ends or, due to shifted cutting sites, complementary single-stranded ends -so-called sticky ends- can be revealed. The function of restriction enzymes is dependent on the concentration of NaCl/KCl in the reaction. Because of that, they should be used with the supplied restriction buffer. In the case of using two restriction enzymes in one reaction, which need different buffers, a condition was chosen in which at least 80 % enzyme activity was given according to the manufacturer's data.

For test-restriction of purified plasmid DNA 6 units restriction enzyme(s) (New England Biolabs) and 2.5 µl appropriate restriction buffer (10x) were added to about 600 ng DNA and filled up to 25 µl with deionized water. The mixture was incubated at 37°C for 90 minutes.

3.2.3 DNA gel electrophoresis

Gel electrophoresis enables the separation of charged molecules according to their size and conformation. DNA molecules are negatively charged due to their sugar-phosphate backbone and migrate towards the anode in an electrical field. The mobility of the molecules is dependent on the applied voltage, the agarose concentration in the gel and the temperature. By variation of the conditions, an optimal DNA separation can be achieved. The migration of the DNA molecule is indirect proportional to the logarithm of the fragment

size (Helling et al., 1974) and sizing can be carried out using DNA fragments of known sizes (DNA marker). For visualization of the DNA, ethidium bromide (EtBr) was added to the agarose solution. Ethidium bromide intercalates between the bases of nucleic acids and forms a stable complex that is fluorescent under UV light. Depending on the size of the PCR product, different percentages of peqGOLD Universal Agarose (Peqlab) were used ranging between 1% and 3% (w/v). Agarose was dissolved in 1xTAE buffer (1x, 0.04 M Tris-acetate, 0.001 M EDTA) by boiling the mixture in a microwave. After cooling down to about 60°C, EtBr was added in a ratio 1:10,000 (10 mg/ml, Roth) and poured into a horizontal gel electrophoresis chamber assembled with a comb. After solidifying of the agarose, the gel was covered with 1xTAE buffer and the comb was removed. In the formed slots, the samples (pre-mixed with sample buffer 6x; 0.25% bromphenole blue, 0.25% xylencyanol, 30% glycerin) were applied. Additionally, a DNA marker (peqGOLD 100bp/1kb DNA Ladder, peqlab) was run separately in order to serve as migration standard. Electrophoresis was carried out at constant electric current of 100 V (Biometra Standard Power Pack P25 power supply) to obtain an optimal fragment separation. After 20-30 minutes, the agarose gel could be analyzed under UV light using the Gel doc system (BioRad Laboratories).

3.2.4 Purification of total DNA from cells

Purification of genomic DNA was carried out with the DNeasy Blood & Tissue Kit (Qiagen). Cells were washed once with ice-cold PBS and then scraped off the tissue culture plate. The suspension was centrifuged at 1,000 x g for 5 minutes at 4°C, supernatant removed and cell pellets were snap-frozen and stored at -80°C or immediately used for DNA purification following the manufacturer's instructions. DNA was eluted in 75 µl of deionized water. Elution was repeated 3 times to maximize the yield of DNA.

3.2.5 DNA acetate precipitation

For DNA precipitation, 15 µl 3 M sodium acetate was added to 150 µl purified DNA and mixed briefly by vortexing. Thereafter, 400 µl ice-cold ethanol (100 %) were added, mixed by vortexing and subsequently incubated on dry ice for 5 minutes. Afterwards, the mixture was centrifuged at 20000 rcf for 30 minutes. The supernatant was removed and 1 ml room-tempered ethanol 70 % was added. After inverting the tube several times, it was centrifuged

again (10 min, 20000 rcf) and the supernatant was removed. The pellet was dried for a few minutes and finally dissolved in 30 µl distilled water.

3.2.6 Purification of total RNA from cells

As long as not specifically mentioned, isolation of total RNA from cells was carried out using the RNeasy Mini Kit (Qiagen) following the manufacturer's instructions. Briefly, cells were rinsed once with ice-cold PBS and directly lysed with RLT buffer supplemented with beta-mercaptoethanol. The lysate was either directly used for RNA purification or stored at -80°C for later extraction. For array analysis, an on-column digest with RNase-free DNase (Qiagen) was performed following the manufacturer's instructions. RNA was eluted in 30 µl RNase-free water and stored at -80°C.

For isolation of total RNA by peqGOLD TriFast (Peqlab), cells were rinsed once with ice-cold PBS and directly lysed with 1 ml of Trifast. Extraction of RNA was carried out according to the instructions given by the manufacturer. RNA was resuspended in 80 µl RNAse-free water and stored at -80°C.

3.2.7 DNA and RNA Quantification

DNA and RNA quantity and quality were analyzed with a Nanodrop ND-1000 Spectrophotometer (PeqLab). Concentration is determined by the absorbance of nucleic acids when exposed to ultraviolet light at a wavelength of 260 nm. The ratio of absorptions at 260/280 nm wavelengths was used to estimate contaminations of the nucleic acids with protein or phenol, whereas a ratio at 260/280 nm >1.8 for DNA and >2.0 for RNA is regarded as pure. Additional analysis of the ratio at 260/230 nm gives information about other organic compounds, where a ratio of optical densities at 260/230 nm lower than 2.0 may indicate contaminants that absorb at 230 nm.

Prior to DNA or RNA measurement, 1 µl of deionized water was used as internal blank of the instrument and an additional 1 µl of the buffer/water in which the nucleic acid is dissolved as blank of the solution. All solutions were applied to the Nanodrop without bubbles and after each measurement the instrument was cleaned with a lint-free paper.

3.2.8 DNase digestion and cDNA synthesis

If not already applied during RNA purification, 1 µg RNA was incubated with DNase I (Thermo Fisher Scientific) prior to cDNA synthesis, following the manufacturer's instructions. The DNase was inactivated by adding 1 µl EDTA (25 mM) and incubation at 65°C for 10 minutes. First strand cDNA synthesis was performed using the iScript cDNA Synthesis Kit (BioRad Laboratories). 10 µl of DNase digestion were directly applied for cDNA synthesis. Reverse transcription was carried out according to the manufacturer's protocol.

3.2.9 Primer design

Primer sequences were either obtained from the publications cited or designed with Primer3 (https://primer3.ut.ee/) using data from the Ensembl genome browser (http://ensembl.org) and NCBI geo (http://www.ncbi.nlm.nhi.gov/geo). PCR primers should range in length from 15 to 30 bases with a GC-content of 40-60%, while sequences that might generate internal secondary structures should be avoided. In the case of primers used for quantitative Real-time PCR, a product size of 100 to 200 bp was considered as optimal for primer design. If applicable, intron-spanning primers were chosen as a further quality control. In case of remaining genomic DNA, an additional product is amplified besides the expected one. All primers were manufactured by Thermo Fisher Scientific and sequences are shown in the material section.

3.2.10 Polymerase chain reaction

Polymerase chain reaction (PCR) is a fast method for amplifying nucleic acids *in vitro* (Saiki et al., 1985). PCR utilizes the ability of DNA polymerases to replicate DNA via employing a single DNA strand as a template for the synthesis of a complementary strand. For replication initiation the polymerase requires a short piece of double stranded DNA, which can be replaced *in vitro* by a synthetic oligonucleotide with a length of about 20 bp. If two oligonucleotides are chosen to flank a desired nucleotide sequence on sense and antisense strand this nucleotide sequence can be selectively amplified.

For amplification the cycling of three successive temperature steps are required: First, a brief heat treatment is required to separate double stranded DNA (denaturation). In the second step, the oligonucleotides (primer) attach to the complementary single-stranded DNA (annealing). Due to this, the temperature in this step is adjusted to the melting

temperature of the applied oligonucleotides. In the third stride, the DNA polymerase elongates the double stranded DNA from 5´ to 3´ (extension). In this last step the temperature is chosen at the activity maximum of the utilized DNA polymerase.

In each cycle of DNA synthesis the newly generated fragments serve as templates in their turn, which then results in an exponential amplification. Because of the brief heating in each cycle the technique requires a thermo-stable DNA polymerase, isolated from a thermophilic, marine bacterium like *thermus aquaticus* (Taq DNA polymerase).

For semi-quantitative reverse transcriptase PCR (RT-PCR) analysis, the purified RNA was treaded with DNase and transcribed to cDNA, which afterwards was diluted 1:5 and used as template DNA in the reaction. To compare the gene expression in different samples, the intensities of the corresponding bands after gel electrophoresis were analyzed and normalized to GAPDH (25 cycles). For all sets of primers a non-template control was used as negative control. PCR conditions and cycle numbers were optimized for each primer set to obtain a specific and optimal amplification of PCR products and to guarantee analysis in the exponential phase of the reaction.

The following components were added to the reaction:

Component	Volume in µl
Template DNA	1
dNTP´s (2 mM)	2
MgCl2 (50 mM)	0.6
10x Buffer	2
Primer Mix Forward and Reverse (5 µM)	2
Taq polymerase (5 U/µl)	0.1
H_2O	Ad 20

PCR was performed in a T3 Thermocycler (Biometra) according to the following program:

Step	Process	Temperature	Time
I	Hot start	94°C	hold
II	Initial denaturation	94°C	5 min
III	Denaturation	94°C	30 sec
IV	Annealing	60°C	30 sec
V	Extension	72°C	1 min
	Repetition of step III to V for 30 - 35 cycles		
VI	Final extension	72°C	10 min
VII	Hold	4°C	hold

3.2.11 Quantitative real time PCR

The quantitative real time polymerase chain reaction (qRT-PCR or qPCR) is a technique for amplification and simultaneous quantification of a target DNA molecule. The method is based on the conventional polymerase chain reaction, but with the addition that the amplified DNA product is quantified in real time via fluorescence measurement after each cycle. The increase in fluorescence is proportional to the amplified DNA, because the used fluorescent dye SYBR® Green (Thermo Fisher Scientific) intercalates in double-stranded DNA and a fluorescent complex is formed. This complex can be excited by blue light with a wavelength of 498 nm and emits green light at a wavelength of 522 nm. Due to the fact that the fluorescence dye intercalates in all double-stranded products, the specificity of the reaction has to be verified by melting curve analysis. At the end of the last elongation phase, the temperature is slowly and continuously enhanced from 55°C to 95°C leading to a melting of the amplified products. At a specific melting temperature, the double-stranded DNA splits in single strands and by this the fluorescence dye SYBR® Green is released and the fluorescence declines. Thereby, it can be examined whether unspecific products have been formed. In addition to this, the products are also applied to a 3% agarose gel for a second revision of unspecific results. All real time PCR reactions were carried out with the realplex Mastercycler (Eppendorf), whereas the following components were added to the reaction:

Component	Volume [µl]
2x qPCR Master-Mix	12.5
Primer-mix forward and Reverse (5 µM)	1
Template cDNA (diluted 1:5)	1
GoTaq polymerase	0.15
H_2O	Ad 25

The following program was performed in the real time PCR:

Step	Process	Temperature	Time
I	Initial denaturation	95°C	3 min
II	Denaturation	95°C	15 sec
V	Annealing	60°C	20 sec
IV	Extension	72°C	30 sec
	Repetition of step II to IV for 40 cycles		
V	Final extension	72°C	1 min
VI	Melting Curve	55°C to 95°C	20min
VII	Hold	4°C	hold

Results of specific amplifications were evaluated with delta-delta C_T method ($\Delta\Delta C_T$) according to Livak and Schmittgen (Livak and Schmittgen, 2001). The calculation of the relative gene expression of a specific target gene is based on the difference in expression of this gene in the sample and the control relatively to the expression of the reference gene. The C_T value (threshold cycle) declares the cycle in which the amount of amplified fragment crosses a defined threshold. In the common $\Delta\Delta C_T$ method an efficiency of 100% (E = 2) for the PCR reaction and a duplication of each potential template is assumed. Furthermore, the results of the target gene were normalized with a reference gene. As reference a gene with almost unregulated mRNA levels should be chosen, a so-called “housekeeping gene”. In the presented study GAPDH or 18S was selected as reference gene. The obtained C_T value in the sample minus the obtained C_T value for the reference gene results in the ΔC_T (certain threshold level). When these were inserted in equation (2) the expression ratio can be calculated.

$$\text{Ratio} = (E_{target})^{\Delta CTtarget(control - sample)} / (E_{ref})^{\Delta CTref(control - sample)} \quad (2)$$

3.2.12 Genome wide gene expression analysis

Genome-wide gene expression analysis was performed at the Institute for Biomedical Engineering, RWTH Aachen University and analyzed in cooperation with Michael Lenz, Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University.

Whole transcript profiles were generated using HumanGene1.0stv1 microarrays (Affymetrix) enabling the quantification of gene expression levels of > 41,000 probes covering protein coding and long intergenic non-coding transcripts. The integrity of isolated total RNA was examined using BioAnalyzer 2100 (Agilent Technologies) following the instructions of the manufacturer’s protocol. All RNA samples used in this study showed intact 28S and 18S ribosomal RNA signals and RNA integrity number (RIN) >9.5. Data was preprocessed with the Robust Multichip Average (RMA) method using apt-probeset-summarize from the Affymetrix Power tools software suite and further analyzed using the Multi Experiment Viewer (MeV, part of TM4 Microarray Software Suite). Hierarchical clustering was performed with Pearson correlation. Significance analysis for microarrays (SAM) was used to search for differences between ESC- and iPSC-derived NSCs. For evaluation of pluripotency based on gene expression, profiles PluriTest was applied with a reference dataset consisting of 98

pluripotent and 1,028 non-pluripotent samples. A previously published dataset on gene expression profiles of human fibroblasts (Shao et al., 2013), ESCs and iPSCs, which were analyzed on the same platform were used as references. These data were quantile normalized and a selection of pluripotency markers was used for heatmap presentation.

3.2.13 Whole genome methylation analysis

Genome-wide DNA methylation analysis was performed at the Institute for Human Genetics, Bonn University and at the Institute for Genetics and Epigenetics, Saarbrücken University and analyzed in cooperation with Michael Lenz, Aachen Institute for Advanced Study in Computational Engineering Science (AICES), RWTH Aachen University.

Whole genome DNA methylation profiles were analyzed using Infinium HumanMethylation 450K beadchip assay (Illumina) enabling the measurement of DNA methylation levels at >450,000 CpGs covering 99% of all reference sequence (RefSeq) genes. The array was scanned on an iScan (Illumina) scanner and analyzed with GenomeStudio software (Illumina).

3.2.14 SNP analysis

Whole genome single nucleotide polymorphism (SNP) genotyping was performed at the Institute for Human Genetics, Bonn University, and was used for assessment of copy number variation and identificaton of heterozygous SNPs for allele-specific expression analysis. 500 ng of genomic DNA diluted in TE4 buffer was processed according to the manufacturer's protocol and hybridized to a HumanCytoSNP-12 bead chip (Illumina) interrogating >200,000 SNPs across the human genome. The array was scanned on an iScan (Illumina) scanner and analyzed with GenomeStudio (Illumina) via assessing B allele frequency and log R ratios.

3.3 Microbiological methods

3.3.1 Cultivation and storage of bacterial cells

Bacterial cells were cultivated in common LB media, supplemented with an appropriate antibiotic, at 37°C in a shaking incubator (GFL 3033).

For storing bacterial cells 830 µl of an overnight culture and 170 µl of sterile glycerol 87 % (autoclaved) were filled in cryo-polypropylene tubes, immediately frozen in liquid nitrogen and stored at -80°C.

3.3.2 Generation of competent bacteria via $CaCl_2$

The competence signifies the ability of bacterial cells to internalize DNA. The natural competence of bacteria alters with the physiological state of the cell and reaches a peak in the medial exponential period of growth before it declines rapidly to a minimum. After induction with $CaCl_2$ the ability of bacterial cells to assimilate loose DNA from the ambient medium can be intensely increased. This elevated ability is maintained for several months if stored at -80°C (Mandel and Higa, 1970). From the glycerol stock a droplet of *E.coli* was seeded on LB agar and cultivated overnight at 37°C. A single colony of *E.coli* was inoculated into 50 ml LB medium supplemented with the appropriate antibiotic and grown overnight at 37°C with shaking (200 rpm). 10 ml of the culture were inoculated into 500 ml LB medium (plus antibiotic) in a 2-liter flask and grown at 37°C, with shaking (200 rpm) until an OD_{600} of about 0.6 was reached. The culture was left on ice for 5 minutes and centrifuged at 2500 x g for another 5 minutes at 4°C. The pellet was gently resuspended in 20 ml ice-cold 50 mM $CaCl_2$ solution (autoclaved). After incubation on ice for 30 minutes, the mixture was again centrifuged at 2500 x g for 5 minutes at 4°C. After resuspension in 5 ml cold 50 mM $CaCl_2$ with 1 ml glycerol 87%, the pellet was incubated on ice for 3 hours. Afterwards, the mixture was divided into 100 µl aliquots and placed in pre-chilled, sterile polypropylene tubes. These were immediately frozen in liquid nitrogen and stored at -80°C.

3.3.3 Transformation

Using the following method, DNA for amplification was internalized into chemical competent bacteria via a heat shock treatment.

One aliquot of chemical-competent bacterial cells was thawed on ice. After 10 minutes, 100 ng of the desired plasmid was added and mixed by gently flicking of the tube. After 20 additional minutes of incubation on ice, the bacterial cells were heat-shocked at 42°C for 40 seconds using a water bath and immediately cooled down on ice for 2 minutes. For regeneration of the bacterial cells, 1 ml pre-warmed SOC or LB medium was added to the mixture and incubated at 200 rpm for 1 hour at 37°C. The mixture was used in case of retransformation to inoculate 500 ml of LB medium containing 100 µg/ml ampicillin (or appropriate antibiotic dependic on the plasmid used) and cultivated overnight at 37°C at 200 rpm.

3.4 Cell stainings

3.4.1 Fixation of cells

Culture medium was removed and cells were washed once with PBS before being fixed with 4% PFA (supplemented with 1:500 glutaraldehyde for GABA staining only) for 20 minutes at RT. Subsequently, cells were rinsed three times with PBS and either stored at 4°C in PBS for later staining or used immediately.

3.4.2 Alkaline phosphatase staining

Enzymatic assay of alkaline phosphatase (AP) activity was performed with Alkaline Phosphatase Substate Kit III (Vector laboratories) following the manufacturer's guidelines on fixed cells. In brief, staining solution was freshly prepared by consecutive addition of the solutions A, B, and C to 100 mM Tris-HCl-buffer during shaking. After incubation for at least 30 minutes, blue colored colonies became visible. Cells were rinsed two times with PBS, covered with moviol and a glass coverslip.

3.4.3 Immunocytochemistry

In order to analyze protein expression immunocytochemistry was conducted. First, unspecific binding sites were blocked via incubating the cell for 30 minutes at RT with blocking solution (10% FCS in PBS) which was supplemented with 0.1 % Triton X-100 for permeabilization to detect intracellular proteins. Then, the primary antibody was applied in blocking solution overnight at 4°C or for two hours at RT. After three washing steps with PBS, plates were incubated with the appropriated secondary antibody in blocking solution for 1 hour at RT and in darkness. Cell nuclei were visualized by subsequent counterstaining with DAPI (1:10000 in PBS) for 3 minutes at RT. Finally, cells were rinsed two times with PBS, covered with moviol and a glass coverslip and analyzed via fluorescence microscopy.

4 Results

4.1 Generation of a human isogenic stem cell system to study reprogramming-associated aberrations

Human pluripotent stem cells (iPSCs) represent a fascinating and virtually unlimited source for the generation of different cell types for research and therapy. However, it is questionable whether and to what extent alterations caused by the reprogramming process can affect readout parameters in disease modeling or clinical safety of human iPSC-derived cell populations. The presented work is aimed at using human iPSCs and neural stem cells to generate a human isogenic stem cell system suitable to explore on a genome-wide level transcriptional and DNA methylation changes associated with the reprogramming process. This human isogenic system of pluripotent and somatic cell populations was generated by initial differentiation of widely available human ESC lines into neural stem cells (NSCs) (Figure 4). This cell type was chosen, as it represents a well-defined, highly standardized and homogenous stem cell population, which can easily be characterized by its typical rosette-like structure and marker expression. Those ESC-derived NSCs were reprogrammed via three different methods and two different factor combinations. The resulting pluripotent cell lines were characterized for pluripotency-related markers and their differentiation potential in all three germ layers, before subsequent differentiation into the very same type of NSCs (iPSC-NSC). The different cell populations in this human isogenic cell system were then analyzed for their expression and methylation profiles.

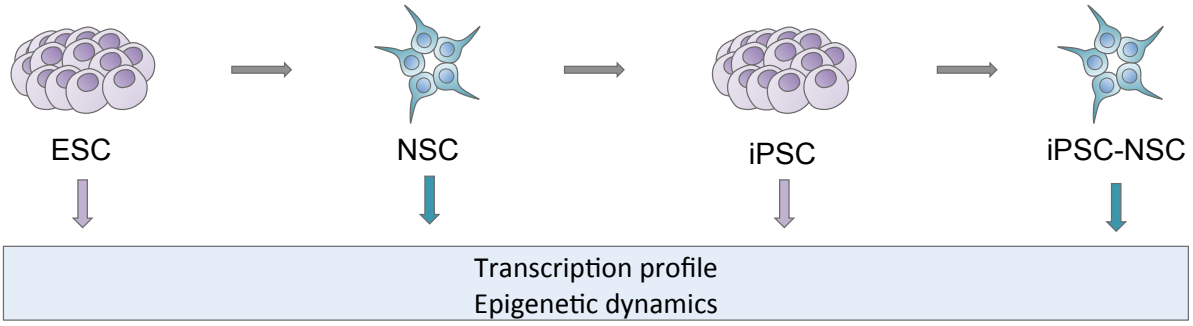


Figure 4: Schematic illustration of the generation of a human isogenic stem cell system. Human embryonic stem cells were differentiated to neural stem cells, which were subsequently reprogrammed to iPSCs and again differentiated into the very same neural stem cell population.

4.1.1 Generation and characterization of iPSCs

To gain insight into reprogramming-associated alterations, human ESC-derived neural stem cells of three different genetic backgrounds and three different reprogramming methods

were applied (Table 1). In detail, NSCs derived from the human ESC-line I3 were either transduced with retroviral vectors encoding OCT4, KLF4 and c-Myc, lentiviral vectors encoding OCT4 and KLF4 under control of tetracycline-inducible EF1alpha promoter or non-integrating Sendai-viral vectors encoding OCT4, KLF4 and c-Myc. In addition, H9.2 ESC-derived neural stem cells were reprogrammed by the aforementioned lentiviral system expressing OCT4 and KLF4, and I6 ESC-derived neural stem cells were reprogrammed by retroviral transduction of OCT4, KLF4 and c-MYC.

Table1: Generated iPSC lines

Name	Background	Sex	Method	Factors	Experiment
I3OKM#3R	I3 NSCs	female	Retrovirus	OCT4, KLF4, c-Myc	1
I3OKM#8R	I3 NSCs	female	Retrovirus	OCT4, KLF4, c-Myc	1
I6OKM#1R*	I6 NSCs	male	Retrovirus	OCT4, KLF4, c-Myc	-
I6OKM#2R*	I6 NSCs	male	Retrovirus	OCT4, KLF4, c-Myc	-
I6OKM#3R*	I6 NSCs	male	Retrovirus	OCT4, KLF4, c-Myc	-
I6OKM#4R*	I6 NSCs	male	Retrovirus	OCT4, KLF4, c-Myc	-
I3OK#2	I3 NSCs	female	TetOn-Lentivirus	OCT4, KLF4	2
I3OK#3	I3 NSCs	female	TetOn-Lentivirus	OCT4, KLF4	2
I3OK#4	I3 NSCs	female	TetOn-Lentivirus	OCT4, KLF4	2
H9.2OK#11	H9.2 NSCs	female	TetOn-Lentivirus	OCT4, KLF4	3
H9.2OK#14	H9.2 NSCs	female	TetOn-Lentivirus	OCT4, KLF4	3
H9.2OK#15	H9.2 NSCs	female	TetOn-Lentivirus	OCT4, KLF4	3
I3OKM#1	I3 NSCs	female	Sendai-virus	OCT4, KLF4, c-Myc	4
I3OKM#3	I3 NSCs	female	Sendai-virus	OCT4, KLF4, c-Myc	4
I3OKM#5	I3 NSCs	female	Sendai-virus	OCT4, KLF4, c-Myc	4

*incomplete transgene silencing

For retroviral or non-integrating Sendai-viral reprogramming of ESC-derived NSCs, cells were transduced with viral particles for transient expression of the transcription factors OCT4, KLF4 and c-Myc. Two days after transduction and cultivation in NSC medium, transduced NSCs were transferred to a layer of irradiated mouse embryonic fibroblasts and cultured in iPSC medium until colony formation. Colonies usually appeared three to four weeks after transduction. Mouse embryonic fibroblasts, mitotically inactivated using gamma irradiation, were used as so-called “feeder cells”, which are necessary to structurally support pluripotent cells and secrete needed factors into the culture medium. Routinely, the efficiency of non-

selectable retro-viral and Sendai-viral transduction was estimated via simultaneous transduction of GFP-virus to a control NSC population and subsequent quantification of GFP-positive cells by flow cytometry.

In case of reprogramming via the lentiviral TetON-system, stable transgenic NSC lines, harboring the tetracycline-inducible constructs of the reprogramming factors OCT4 and KLF4, were generated via virus transduction followed by a multistep chemoselection process that was developed to enhance reprogramming efficiency by minimizing the derivation of transgenic polyclonal cultures of NSCs. First NSCs were transduced with the Tet-On-lentivirus and subsequently selected with G418 for several days to generate a stable rtTAAdv-expressing cell line. These NSCs were then transduced with OCT4 viral particles and selected with puromycin resulting in a stable inducible OCT4-NSC line, which was additionally transduced with virus coding for KLF4. After several days of stable cultivation, doxycycline was added to the medium to induce expression of the reprogramming factors. One day following induction, NSCs were transferred to a layer of irradiated murine fibroblasts and cultured in iPSC medium containing doxycycline. To assure that only doxycycline-independent colonies were further propagated, doxycycline was withdrawn upon colony formation after 2-3 weeks.

Emerging human iPSCs could be distinguished morphologically by its clear-zoned colony-shaped growth on the ambient mouse embryonic fibroblasts and their compact epithelial cell pattern. Upon visual examination single clones were mechanically isolated, propagated as clonal cell lines and extensively characterized.

4.1.2 Human induced pluripotent stem cells exhibit pluripotency properties

First, the generated iPSC lines were examined for their transgene independence in order to exclude that residual transgene expression confounds a subsequent differentiation and/or analysis of reprogramming-associated alterations. Therefore, the expression status of transgenic factors was analyzed via transgene-specific PCR or qPCR (Figure 5 A-C).

The used genome-integrating retroviral vectors should be silenced upon pluripotency, while integrated doxycycline-inducible constructs should no longer be expressed upon withdrawal of doxycycline. Furthermore, non-integrating Sendai-viral particles should get lost over time in culture, as they are incapable of transmissible-virion production due to a deletion of the *envelope fusion gene* (*f-gene*). If necessary, this spontaneous loss may be enhanced by

cultivation at 39°C for several days, as further mutations cause an elevated temperature-sensitivity of the Sendai-viral particles (Inoue et al., 2003; Fusaki et al., 2009; Ban et al., 2011).

I3 NSC-derived iPSCs obtained via retroviral transduction of reprogramming factors (I3OKM#3R, I3OKM#8R) showed proper silencing of the transgenes, while I6 NSC-derived clones show an insufficient downregulation of the KLF4 transgene and were therefore excluded from further experiments (Figure 5 A). Furthermore, iPSCs generated with the tetracycline-inducible system showed only minimal leakiness in some lines (Figure 5 B) and all iPSCs obtained via Sendai-viral transduction were negative for viral constructs (Figure 5 C).

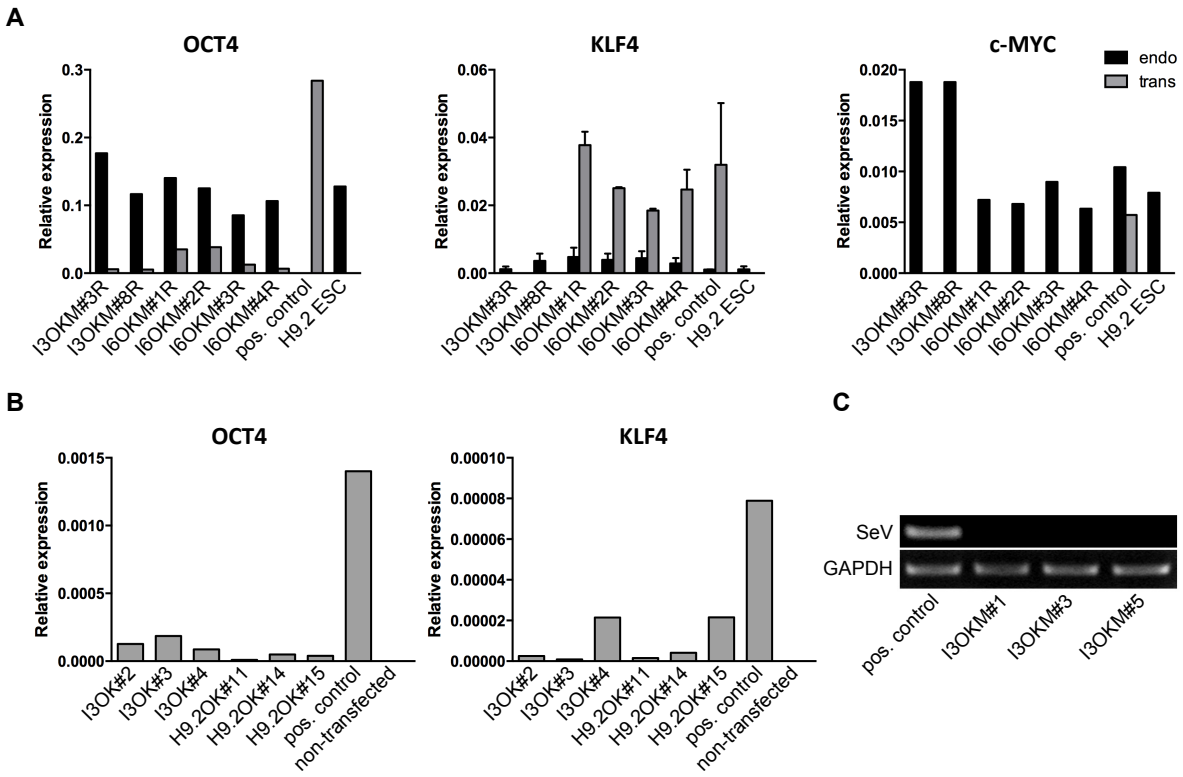


Figure 5: Quality control of induced pluripotent stem cells to ensure transgene independency. (A) Transgene silencing of retroviral constructs. Retroviral transcript-specific qPCR confirmed effective silencing of the used retroviruses for the lines I3OKM#3R and I3OKM#8R, while I6 NSC-derived iPSC lines show an insufficient downregulation of the KLF4 transgene; pos. control: HEK-cells 24h after transfection **(B)** qPCR analysis of the rigor of the tetracycline-inducible system; pos. control: transgene expression of a stable I3OKM line after 24 hours of doxycycline incubation **(C)** Sendai-viral specific RT-PCR indicates loss of Sendai-viral particles at passage 8 and no cultivation at 39°C was required. Pos. control: freshly transfected cells.

The remaining human iPSC lines were positive for the pluripotency-associated markers such as alkaline phosphatase (AP), OCT4, tumor-related antigen TRA1-60 and TRA1-81 on protein level (Figure 6).

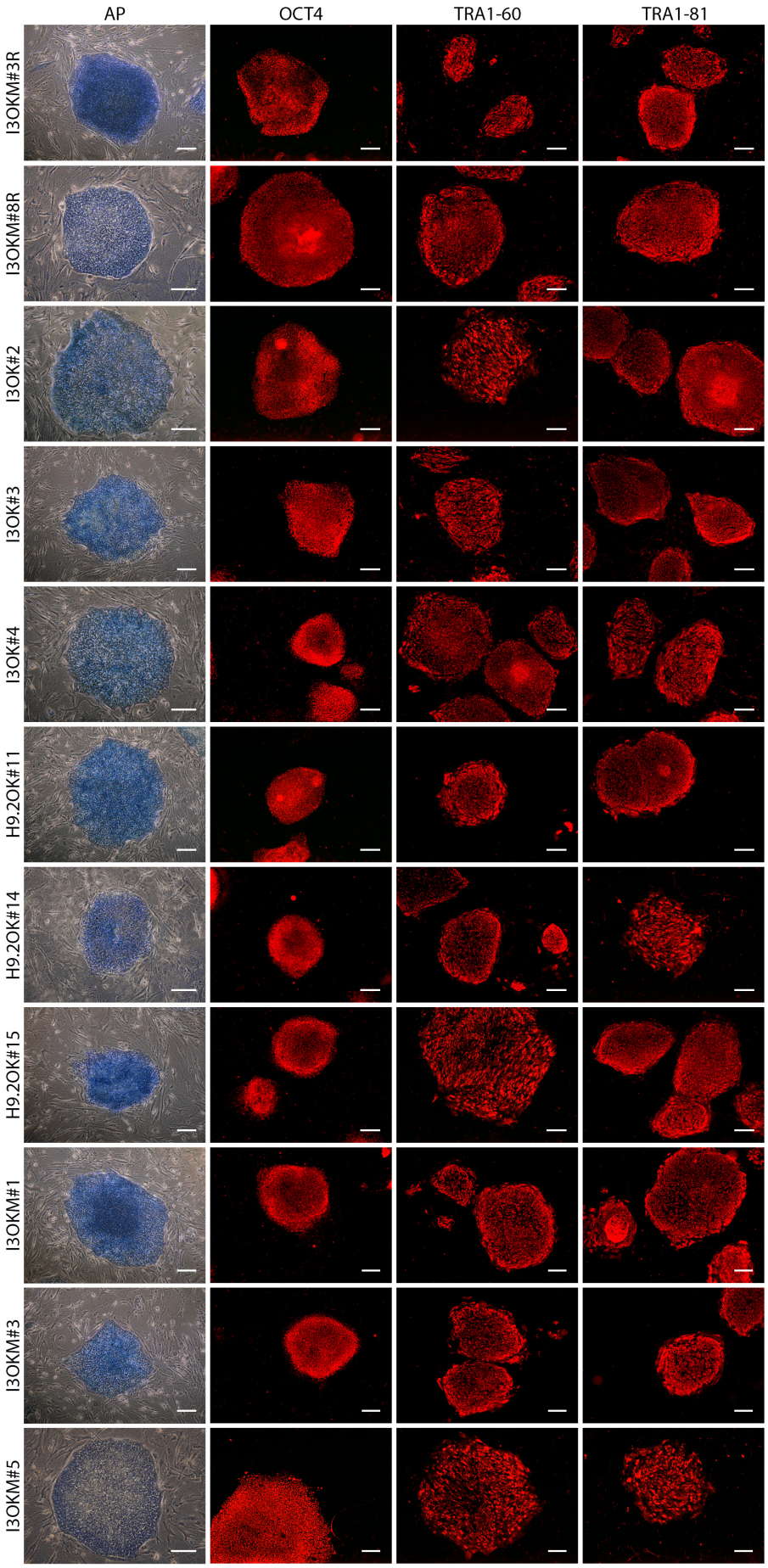


Figure 6:
Human iPSCs express pluripotency-associated proteins.
iPSCs derived from human ESC-NSCs are positive for alkaline phosphatase (AP) and express pluripotency-associated proteins OCT4, TRA 1-60 and TRA 1-81 as shown by immunocytochemical stainings. Scale bars: 250µm.

While alkaline phosphatase (AP) is expressed in most cell types, it is highly elevated in pluripotent cells like ESCs and iPSCs. AP is responsible for hydrolytic dephosphorylation of several molecules including alkaloids, nucleotides and proteins under alkaline conditions. Upon alkaline phosphatase staining, a blue reaction product is formed in the presence of AP enzyme and undifferentiated pluripotent cells appear blue while differentiated cells like the surrounding MEFs appear colorless. The transcription factor octamer-binding protein 3/4 (OCT4, also known as POU5F1) is expressed in the nucleus and contains a POU homeodomain that plays a key role in embryonic development and stem cell pluripotency. The phenotypic markers, TRA1-81 and TRA1-60, which are different epitopes on the same sialylated keratan sulfate proteoglycan podocalyxin, are expressed on the cell surface (Badcock et al., 1999; Adewumi et al., 2007). Podocalyxin is a high molecular weight transmembrane glycoprotein, which can be found on the surface of cells of the inner cell mass (but not morula or trophoblast) and human embryonic stem cells. The expression of both carbohydrate epitopes is stage-specific and gets lost upon cell differentiation (Schopperle and DeWolf, 2007).

Exemplarily, iPSC lines generated via retroviral transduction of reprogramming factors (I3OKM#3R and I3OKM#8R) were further analyzed for additional pluripotency marker expression (Adewumi et al., 2007). iPSCs clones I3OKM#3R and I3OKM#8R displayed similar levels of undifferentiated ESC-marker gene expression as ESCs, such as *nanog homeobox* (*NANOG*), *DNA (cytosine-5-)-methyltransferase 3 beta* (*DNMT3B*), *fibroblast growth factor 4* (*FGF4*), *growth and differentiation factor 3* (*GDF3*), *biogenesis of ribosomes* (*BRIX*), *undifferentiated embryonic cell transcription factor 1* (*UTF1*), *nodal growth differentiation factor* (*NODAL*), *teratocarcinoma-derived growth factor 1* (*TDGF1*), *reduced expression 1* (*REX1*) and *left-right determination factor 1* (*LEFTB*), as indicated via RT-PCR, while the excluded I6 NSC-derived iPSCs lack some of these markers (Figure 7).

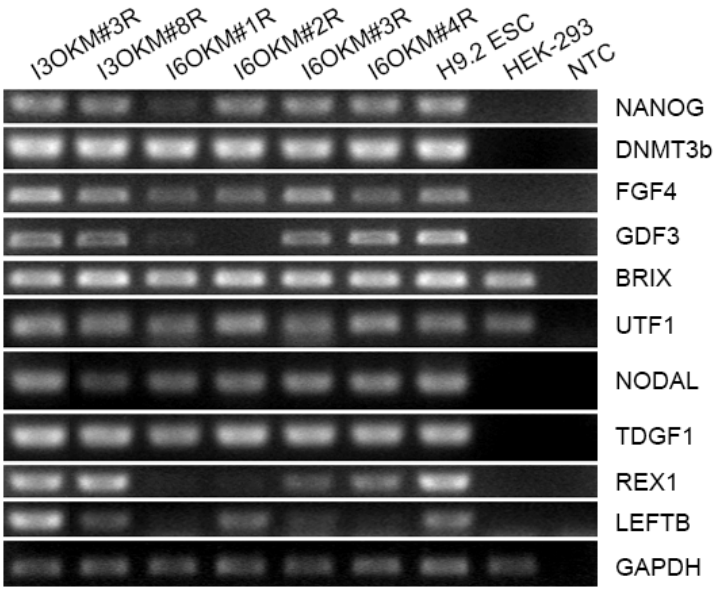
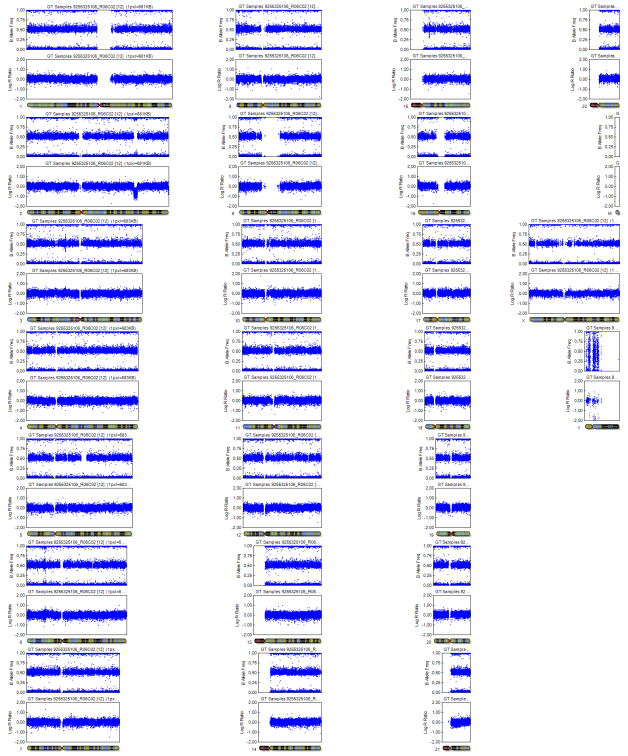


Figure 7: iPSCs derived by retroviral transduction expressed pluripotency-associated marker genes on mRNA level.
Semi-quantitative RT-PCR analysis of several pluripotency markers: *nanog homeobox (NANOG)*, *DNA (cytosine-5-)-methyltransferase 3 beta (DNMT3B), fibroblast growth factor 4 (FGF4), growth and differentiation factor 3 (GDF3), biogenesis of ribosomes (BRIX), undifferentiated embryonic cell transcription factor 1 (UTF1), nodal growth differentiation factor (NODAL), teratocarcinoma-derived growth factor 1 (TDGF1), reduced expression 1 (REX1)* and *left-right determination factor 1* (*LEFTB*). NTC: no template control

Furthermore, the genomic integrity of the different iPSC lines was analyzed. To this end, genome-wide high-resolution single nucleotide polymorphism (SNP) analysis was performed. This method allows to analyze the relative DNA content (log R ratio, LRR) and the frequency of SNPs or alleles (B allele frequency, BAF) present in a sample. The log R ratio is the log2 ratio of the observed to the expected signal intensity and should be around 0 for samples with a normal chromosome content, while duplications and deletions lead to increased or decreased values, respectively. The B allele frequency visualizes the proportional occurrence of the so-called reference allele (B allele). In a normal diploid chromosome set, BAF has a value of 0 or 1 in the case of a homozygous SNP (A/A, B/B) or a value of 0.5 in the case of a heterozygous SNP (A/B), and the duplication or deletion of a SNP/allele changes this ratio accordingly. In combination these two characteristics can therefore be used to identify copy number variation (CNV) in the genome (Figure 8 A). Sendai-reprogrammed iPSC lines acquired no aberrations when compared to their ESC-derived NSC line of origin using standard settings of the CNV partition plugin v3.2.0 that highlights chromosomal copy number changes of regions ≥ 1 Mb containing ≥ 50 continuous SNP probes. All distinctions that could not be detected via SNP analysis in NSC lines of origin are listed in Figure 8 B.

A



B

Cell line	Chromosome	Region	Type
I3OKM#3R	17	16287657-17665514	Gain
	17	11356404-16287482	Loss
	20	29267954-62219097	Gain
I3OKM#8R	17	16287482-17665514	Gain
	17	11356404-16292487	Loss
	20	29267954-62207762	Gain
	X	1018727-154475150	Loss
I3OK#2	7	11170396-43522684	Loss
I3OK#3	7	11154487-43522684	Loss
I3OK#4	7	11170396-43529304	Loss
H9.2OK#11	1	143343508-249075535	Gain
	17	42835129-81051007	Gain
H9.2OK#14	1	143343508-249075535	Gain
	17	42835129-81051007	Gain
H9.2OK#15	1	143343508-249218992	Gain
	17	42835129-81051007	Gain

Figure 8: Test for chromosomal alterations. (A) Representative genome-wide single nucleotide polymorphism analysis of iPSC line I3OKM#1 with no additionally acquired abnormalities. Chromosomal aberration on chromosome 2 was already present in I3 NSC line of origin. BAF is shown in the upper panels and LRR in the lower panels of each chromosome. **(B)** Table of detected aberrations that could not be detected via SNP analysis in starting populations.

4.1.2.1 iPSCs exhibit differentiation potential into all three germ layers

Pluripotent cells are characterized by their potential to differentiate into all three germ layers. To validate full differentiation potential *in vitro*, iPSC-derived embryoid bodies were cultivated in the absence of bFGF and plated on gelatin-coated dishes. Spontaneous differentiation was analyzed after 10 to 14 days by immunocytochemistry for alpha-fetoprotein (AFP, endoderm), alpha-2-smooth muscle actin (SMA, mesoderm) and beta-III-tubulin (TUBB3, ectoderm). For all analyzed iPSCs lines, cells of the three germ layers could be found *in vitro* (Figure 9).

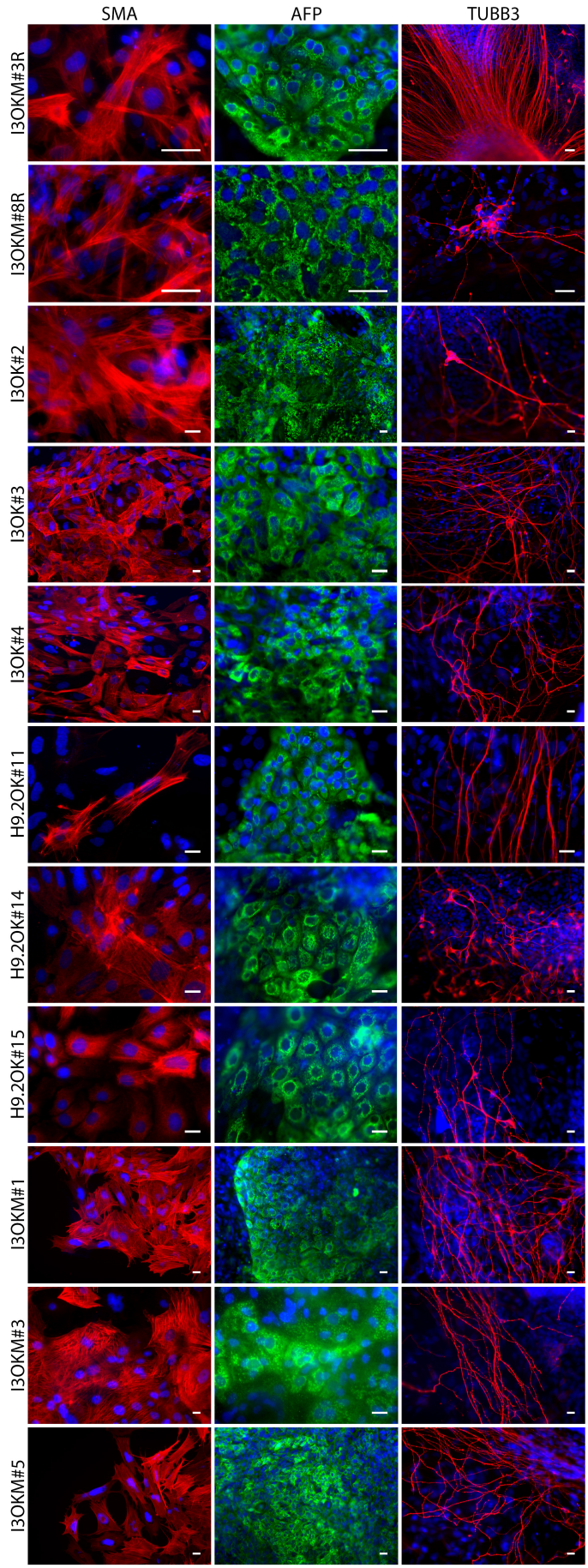


Figure 9: *In vitro* differentiation potential of human iPSC. All generated iPSC lines possess the capacity to differentiate into all three germ layers as indicated by immunocytochemical analysis for SMA, AFP and TUBB3. Nuclei were counterstained with DAPI (blue). Scale bars: 50µm.

Exemplarily, retroviral iPSC clones (I3OKM#3R and I3OKM#8R) were further analyzed for their *in vivo* differentiation potential by teratoma formation assay in immune-deficient mice (Wesselschmidt, 2011). A teratoma is a tumor, which exhibits differentiated derivates of all three embryonic germ layers. To exclude a possible immune rejection of human cells *in vivo*, genetically immune-compromised SCID/beige mice were used. Due to mutations, these mice are defective in generating mature immune-competent B- and T-lymphocytes as well as natural killer cells (Boermans et al., 1992). Therefore, xenographic cells cannot be identified and eliminated. iPSCs were injected into testicles of immune-deficient mice and teratomas were analyzed histologically 6 to 8 weeks post-transplantation via hematoxylin and eosin (H&E) staining. The isolated teratomas showed derivates of all three germ layers such as structures of adenoid tissue (endoderm), primitive neuroepithelium (ectoderm) and cartilage (mesoderm) (Figure 10). In summary, it can be said that all iPSC clones show the cardinal properties of ESCs.

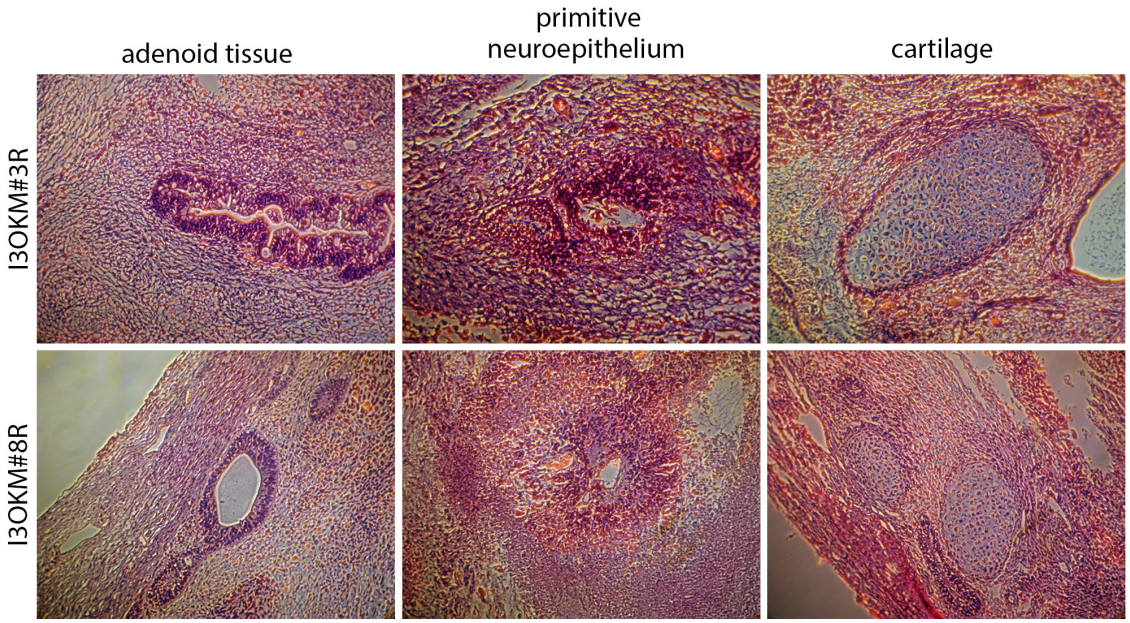


Figure 10: *In vivo* differentiation potential of iPSC lines derived by retroviral transduction. Hematoxylin and eosin (H&E) staining of formed teratomas revealed the potential of iPSCs to differentiate into all three germ layers *in vivo* as indicated by adenoid tissue (endoderm), primitive neuroepithelium (ectoderm) and cartilage (mesoderm) formation.

4.1.3 Derivation of stable NSC lines from pluripotent stem cells

In order to analyze reprogramming-associated alterations in a somatic cell type, induced pluripotent stem cells were differentiated into the very same NSC population as their embryonic counterparts. As reproducibility and comparability between single experiments

and cell lines are key requirements in attempts of that kind, a stable and reproducible generation of a well-defined and homogenous neuronal stem cell population is an important prerequisite. The NSCs used in this study, known as long term proliferating pluripotent stem cell-derived neuroepithelial stem (lt-NES) cells, represent a remarkably stable stem cell population with the ability of homogenous self-renewal over more than 100 passages while maintaining doubling times and avoiding senescence (Koch et al., 2009; Falk et al., 2012). Furthermore, lt-NES cells represent a multipotent intermediate stem cell population between pluripotent stem cells and differentiated neuronal cultures, which stably gives rise to mature cultures of human neurons. *In vitro* differentiation of ESCs and iPSCs into lt-NES cells was performed as described previously (Koch et al., 2009).

To induce differentiation, pluripotent cells were detached with collagenase and cultivated as embryoid bodies (EBs). After 4 to 5 days, floating aggregates were allowed to adhere and emerging neural rosette structures were mechanically isolated and propagated as neurospheres in NSC-media for additional three days. Floating spheres were dissociated to single cells and cultivated as a homogenous monolayer on polyornithine/laminin (Po/Ln)-coated dishes in the presence of FGF2, EGF and B27 supplement (Figure 11).

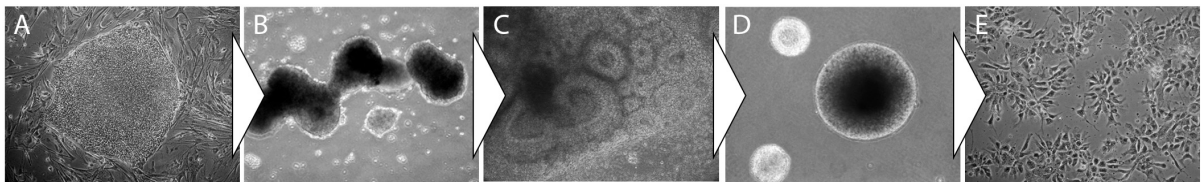


Figure 11: Schematic illustration of the experimental workflow for generation of human ESC- and iPSC-derived NSCs. Pluripotent cells **(A)** cultured on fibroblasts were dissociated by collagenase treatment, cultured as free-floating embryoid bodies **(B)** for 4 to 5 days, plated on PO/Ln-coated dishes and differentiated for 10 days in N2 medium containing FGF2. Neural tube-like structures **(C)** were mechanically isolated and expanded as floating aggregates **(D)** for 3 days. Subsequently, aggregates were triturated to single cells and cultured as a monolayer **(E)** on PO/Ln-coated dishes. Phase contrast images of (A) iPSC colony on feeder cells, (B) EBs of detached colonies, (C) neural rosette structures, (D) free-floating neurospheres and (E) established NSCs.

4.1.4 iPSC-derived NSC lines resemble their ESC-derived counterparts in marker expression and differentiation potential

Neural stem cells derived from different iPSC lines and ESCs were analyzed for the maintenance of neural stem cell characteristics and their differentiation potential. Human iPSC-derived NSCs grow as a homogenous population with typical rosette-like structures comparable to ESC-derived NSCs. As shown by immunocytochemical analysis, both cell populations display nuclear expression of the neural stem cell markers SRY-box 2 (SOX2),

promyelocytic leukemia zinc finger (PLZF) and dachshund family transcription factor 1 (DACH1). Furthermore, ESC-NSCs and iPSC-NSCs were positive for the intermediate filament nestin (NES) and tight junction protein 1 (ZO1), which is characteristic for this neural stem cell population. ZO1 showed a distinct localization in the center of the rosette-like structures independent of the ESC or iPSC origin (Figure 12).

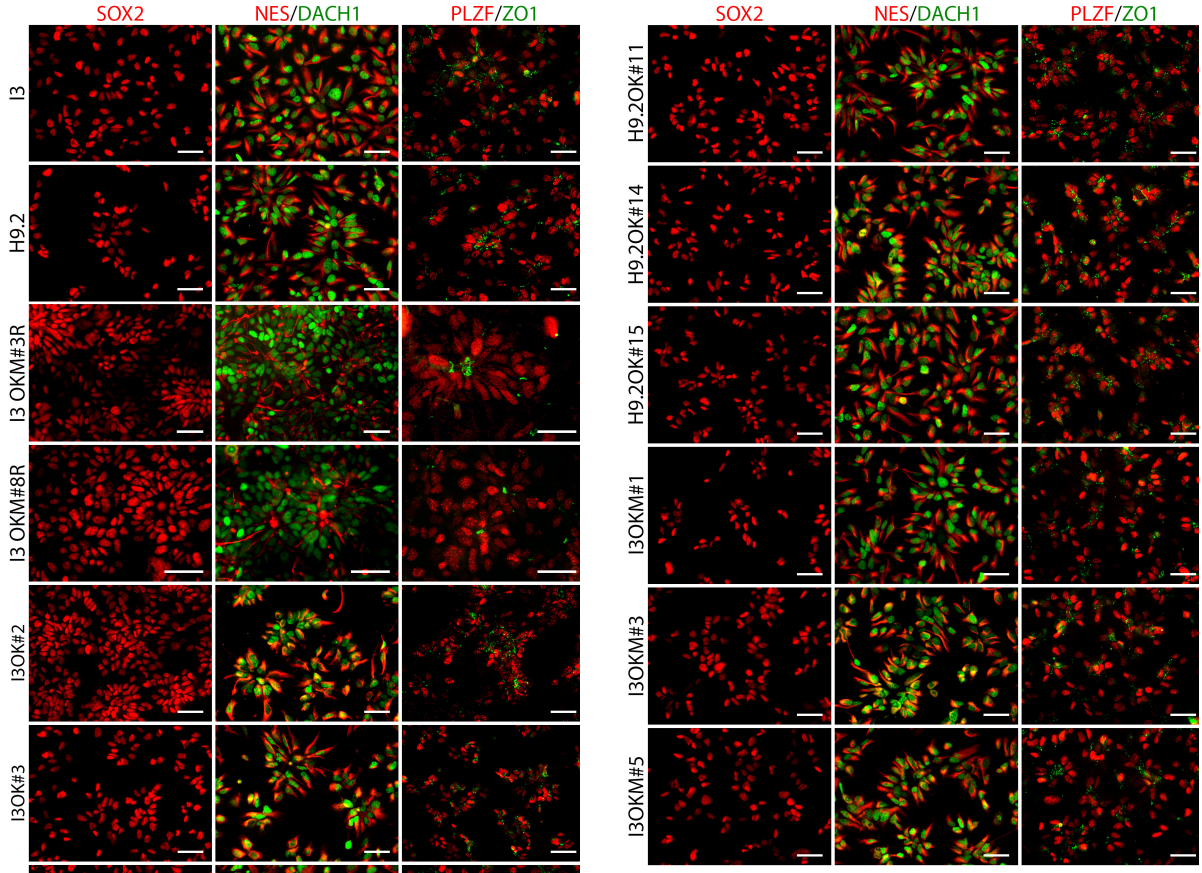


Figure 12: iPSC-derived NSCs resemble ESC-derived NSCs in their marker expression.
Immunocytochemical analysis for SOX2, NES, DACH1, PLZF and ZO1 indicated a similar neural stem cell identity of ESC- and iPSC-NSCs. Scale bars: 50 µm.

Analogous to ESC-derived NSCs, iPSC-derived NSCs can be cultured in the presence of growth factors for more than 60 passages as a stable stem cell population and preserve their neuronal differentiation potential. iPSC- and ESC-derived NSCs were analyzed after 6 weeks of undirected differentiation upon growth factor withdrawal. Upon differentiation, a major fraction of β-III-tubulin (TUBB3) and microtubule-associated protein 2 (MAP2AB) positive neurons with the typical neuronal morphology and a minor fraction of glial fibrillary acidic protein (GFAP) positive astrocytes could be observed in both, iPSC- and ESC-derived,

populations. In addition, both neuronal populations showed cells positive for a gamma-aminobutyric acid (GABA) neuronal subtype (Koch et al., 2009) (Figure 13). A GABAergic neurotransmitter phenotype of iPSC-NSC- and ESC-NSC-derived neuronal cultures may indicate a similar ventral-anterior hindbrain identity of both populations, a region that gives rise to inhibitory hindbrain interneurons during human development.

The expression of neural stem cell markers and the neurogenic potential of iPSC-derived NSCs could be confirmed for all lines analyzed.

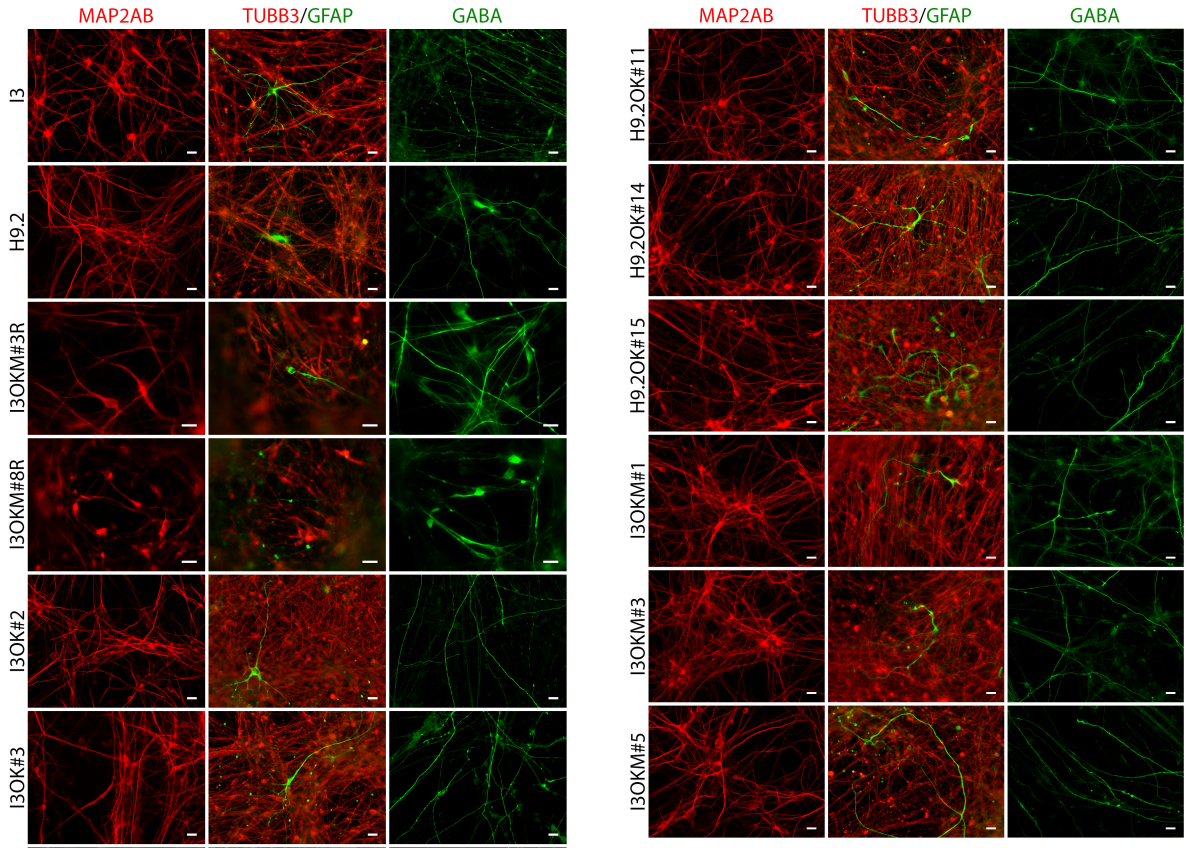


Figure 13: iPSC-derived NSCs resemble ESC-derived NSCs in their differentiation potential.
Immunocytochemical analysis demonstrated similar gliogenic (GFAP) and neurogenic differentiation potential (MAP2AB, TUBB3, GABA). Scale bars: 20µm.

4.2 Analysis of the transcriptional profiles in an isogenic stem cell system

Having established a stable isogenic cell culture model, whole genome expression analysis was performed on both neural stem cell populations as well as their parental, pluripotent cells of origin to elucidate any transcriptional or epigenetic alteration that may influence their biomedical application. Total RNA from three H9.2 iPSC-derived NSC lines, eight I3 iPSC-

derived NSC lines, five iPSC lines (I3) and their corresponding ESC(-derived) counterparts were labeled and hybridized (in triplicates) to an Affymetrix HumanGene1.0stv1 whole genome microarray. This assay enables the quantification of gene expression of more than 41,000 probes covering protein coding and long intergenic non-coding transcripts. The integrity of isolated total RNA was examined using BioAnalyzer 2100 (Agilent Technologies), showing intact 28S and 18S ribosomal RNA signals and RNA integrity number (RIN) >9.5 for all samples.

To get an overview of the samples and to detect global trends, large-scale pattern based analyses of the expression data was performed. The success of the reprogramming and differentiation procedures was evaluated by PluriTest and PhysioSpace analyses, respectively and principal component analysis and hierarchical clustering were utilized to check for outliers and to detect (dis)similarities between samples in an unsupervised manner.

Comparing the transcriptional data with a reference dataset consisting of 98 pluripotent and 1,028 non-pluripotent samples (Shao et al., 2013), pluriTest analysis (Müller et al., 2011) showed clustering of ESCs and iPSCs with other pluripotent samples (Figure 14 A) and clustering of NSC lines in close vicinity on the path from pluripotent to neuronal cells (pluripotency score vs. neurality score) independent of their cellular origin (Figure 14 B). This confirmed successful reprogramming of the cells as well as differentiation of NSCs towards the neural linage.

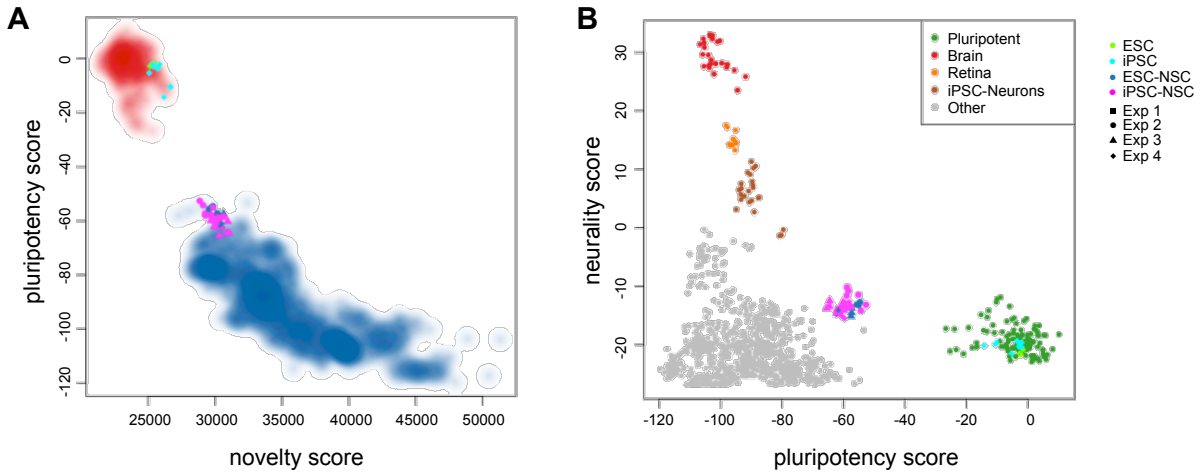


Figure 14: Analyzed samples show correct clustering when compared with reference datasets. (A) PluriTest analysis of ESCs, iPSC and NSC lines. Red and blue clouds indicate pluripotent and non-pluripotent reference samples, respectively. **(B)** Neurality-pluripotency analysis of ESCs, iPSC and NSC lines. ESC and iPSC lines cluster with public domain control pluripotent cell lines, whereas NSC samples cluster between neural tissue/iPSC-derived neurons and pluripotent cells. Control fibroblasts show no neurality or pluripotency.

However, PhysioSpace analysis (Lenz et al., 2013), a method to analyze differentiation in a broader context via detecting the shift of differentiated cells (compared to undifferentiated ESCs) towards signatures representing different human tissues, indicated a stronger bias to fetal brain and other tissues in the retroviral iPSC-NSCs (I3OKM#3R, I3OKM#8R) than for all other NSC lines (Figure 15 A).

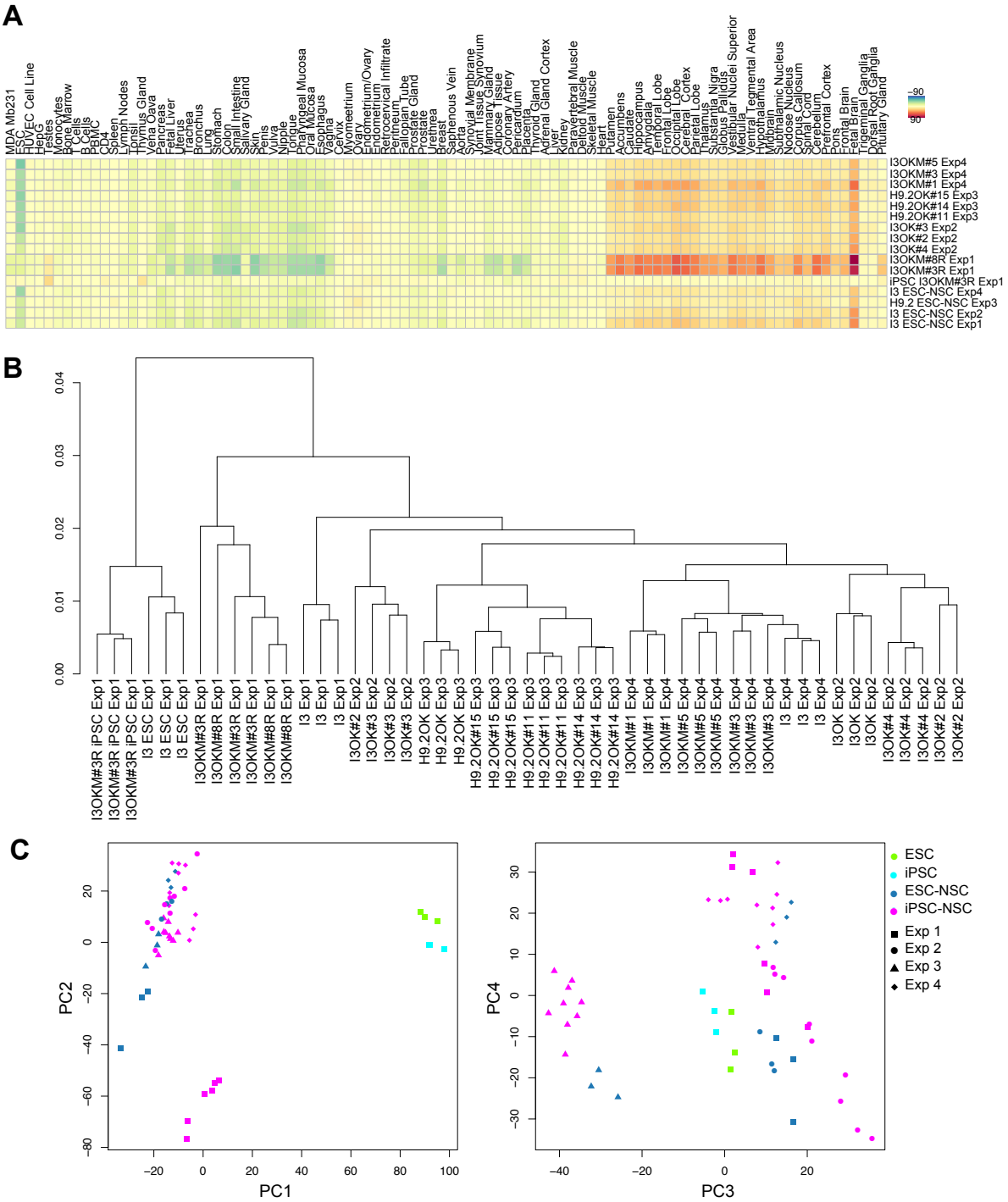


Figure 15: Comparability of the datasets. (A) PhysioSpace analysis, **(B)** Global hierarchical clustering of gene expression data and **(C)** Principal component analysis of ESCs, iPSC and NSC lines. The first principal component PC1 indicates a separate clustering of pluripotent samples and NSCs, PC2 indicates separate clustering of NSC samples of the first experiment (Exp1) and all other NSC samples (Exp2, Exp3, Exp4) and PC3 indicates separate clustering of samples of the two different genetic backgrounds H9.2 (Exp3) and I3 (Exp1, Exp2, Exp4).

Furthermore, hypothesis-free examination via hierarchical clustering and principal component analysis (PCA) revealed larger differences between the first set of NSCs and all other NSC lines analyzed than between different backgrounds of H9.2- and I3-derived cells (Figure 15 B, C). Samples of the first experiment clustered separately in principal component 2 (PC2) from all other neural stem cell samples (Figure 15 C). As this lack of comparability may hamper later data interpretation, the first data set was excluded from further analysis.

4.2.1 Differential expression is enriched on the X-chromosome

The remaining datasets, nine iPSC-NSC lines and their matched ESC-derived counterparts (three ESC-NSC lines) generated in three independent experiments all in biological triplicates and one parental ESC line (I3; triplicates) and three iPSC lines derived thereof (experiment 4, each in triplicates) were analyzed in more detail. Hierarchical clustering indicated a correct separate clustering of pluripotent and neural stem cells and revealed larger differences between H9.2- and I3-derived cells than between isogenic ESC- and iPSC-derived NSCs. Even distinct experiments showed larger differences than isogenic ESC- and iPSC-derived NSCs (Figure 16 A). Furthermore, also principal component analysis indicated that iPSC and ESC-derived NSCs are closer to each other than NSCs with different genetic backgrounds (Figure 16 B), pointing to a high similarity between ESC- and iPSC-derived NSCs.

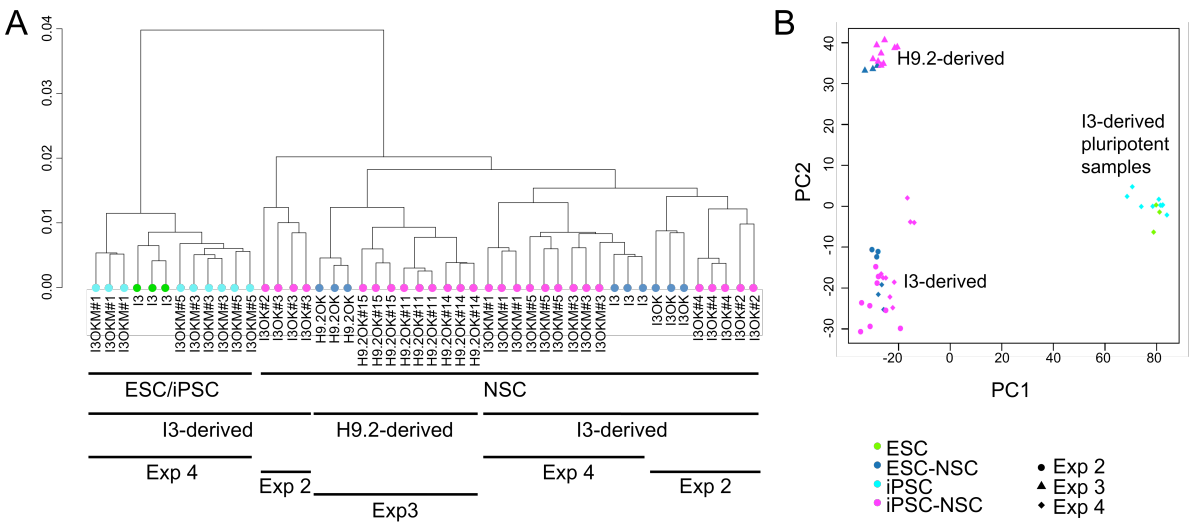
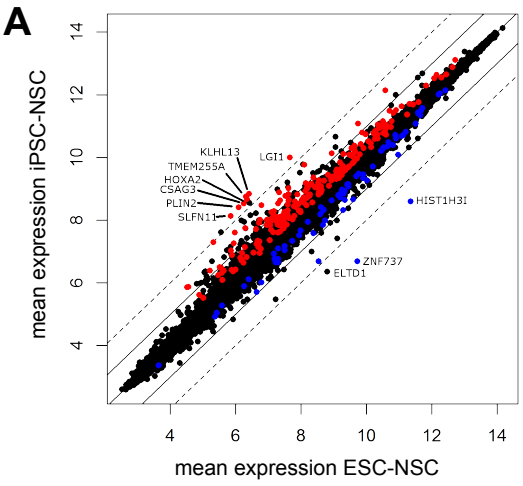


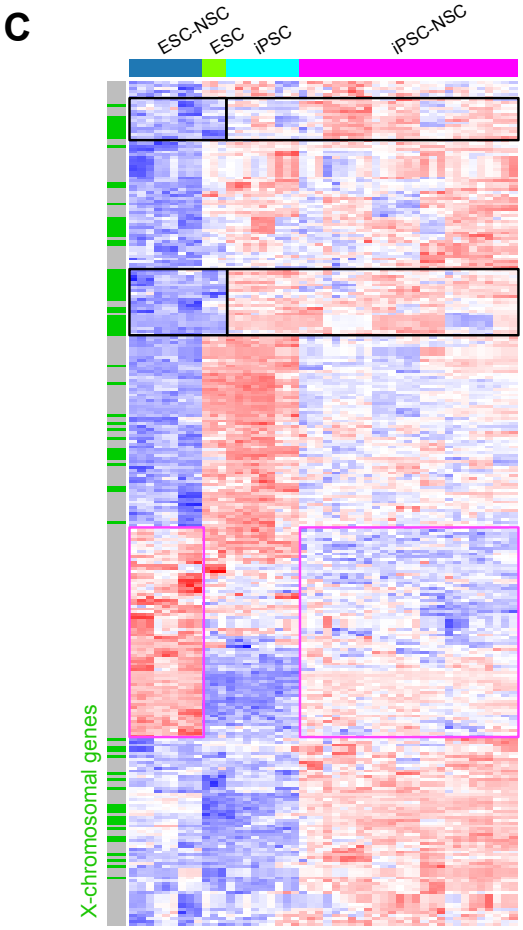
Figure 16: Samples of different genetic backgrounds show distinct clustering. (A) Global hierarchical clustering and **(B)** Principal component analysis of gene expression data shows separate clustering for NSC samples derived from I3 and H9.2 genetic background.

Overall gene expression analysis revealed only minor differences between ESC- and iPSC-derived lines (Figure 17 A) (as few as 36 genes deviate at a 2-fold difference, adjusted p-value <0.01; Figure 17 B).



B

Gene	Chromosome	Overall log2 fold change	Overall adj. p-value
ZNF737	19	-3,04	1,01E-04
HIST1H3I	6	-2,76	1,13E-03
PECAM1	17	-1,85	6,26E-03
LRCH2	X	1,00	1,25E-04
SLC9A7	X	1,02	2,07E-04
EPB41L3	18	1,03	5,70E-04
CDYL2	16	1,04	6,75E-11
STPG1	1	1,08	3,34E-03
C21orf90	21	1,11	1,37E-04
ASTN1	1	1,15	2,80E-03
ANO5	11	1,16	7,48E-03
SLITRK3	3	1,18	6,01E-05
TGFB2	1	1,19	6,91E-05
KIF1A	2	1,20	6,91E-05
CXorf57	X	1,21	7,39E-12
STX3	11	1,26	3,07E-04
SOHLH2	13	1,26	5,20E-03
MAGEA2B	X	1,25	5,77E-03
RBM11	21	1,31	6,98E-03
ZNF215	11	1,31	4,23E-04
SYT11	1	1,34	9,85E-06
LINC00461	5	1,35	2,27E-03
TOM1L1	17	1,39	8,73E-03
SYT14	1	1,45	1,11E-03
COL1A2	7	1,56	6,35E-03
H1F0	22	1,56	4,04E-08
NFIA	1	1,62	7,52E-03
STMN2	8	1,67	3,77E-03
SLFN5	17	1,68	7,36E-07
ZDHHC15	X	1,98	1,08E-04
SLFN11	17	2,27	1,74E-03
CSAG3	X	2,29	2,99E-03
PLIN2	9	2,31	2,38E-05
LGI1	10	2,34	5,04E-03
KLHL13	X	2,42	7,39E-12
TMEM255A	X	2,44	8,66E-08



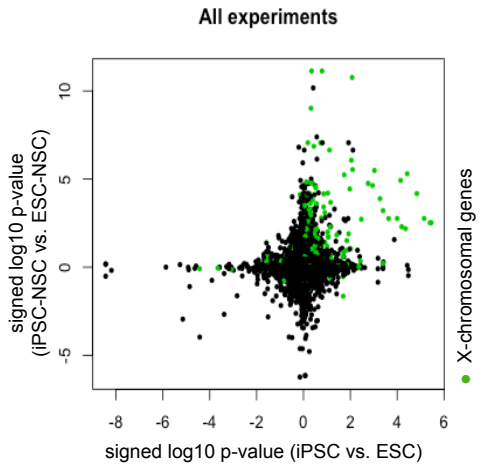


Figure 17: Differential expression is enriched on the X-Chromosome. (A) Scatter plot of global gene expression patterns comparing iPSC-NSCs with ESC-NSCs. Dotted and solid lines indicate fold change >4 and >2 in gene expression levels, respectively. Up- and down-regulated fractions are indicated in red and blue, respectively. **(B)** List of differentially expressed genes (log2-fold change>1, adjusted p-value <0.01) **(C)** Transcriptome heatmap of differentially expressed genes (p-value <0.01) comparing iPSC-NSCs with ESC-NSCs. Black boxes indicate joint up-regulation in iPSCs and iPSC-NSCs compared to ESCs and ESC-NSCs. Pink boxes indicate down-regulated genes when comparing iPSC-NSCs with ESC-NSCs. **(D)** Scatter plot of global gene expression comparing iPSC-NSCs with ESC-NSCs and iPSCs with ESCs.

Amongst the differentially expressed genes (DEGs) the predicted imprinted zinc finger gene ZNF215 was found, but no enrichment for other imprinted genes could be detected. In general, no gene ontology was enriched in the differentially expressed genes between ESC- and iPSC-derived NSCs (p-value <0.05; DAVID; Huang et al., 2009a; b). However, gene set enrichment analysis (Enrichr; Chen et al., 2013b; Kuleshov et al., 2016) revealed that 13 of the 36 differentially expressed genes (adjusted p-value = 0.00004058; LINC00461, KLHL13, CDYL2, ZDHHC15, LRCH2, SLFN11, EPB41L3, SLITRK3, CXORF57, ZNF215, KIF1A, ASTN1, RBM11) were associated with the histone modification gene set H3K27me3 in SK-N-SH cells (H3K27me3_SK-N-SH_hg19, containing more than 2200 genes).

Furthermore, many of the differentially expressed genes such as TMEM255A, KLHL13, CSAG3 or ZDHHC15 are located on the X-chromosome and several differentially expressed X-chromosomal genes displayed a joint up-regulation in iPSCs and iPSC-derived NSCs, whereas just a few autosomal genes showed a joint up- or down-regulation in pluripotent and neural cells, indicating that these alterations are maintained during differentiation (Figure 17 C, D).

To verify whether the obtained array-based results are reliable, the expression differences of several identified differentially expressed genes were validated via qPCR analysis (Figure 18 A, B). The qPCR-data could confirm the change in expression for all samples with at least 2-fold difference and revealed comparable trends, yet with slightly higher differences between ESC- and iPSC-derived NSCs.

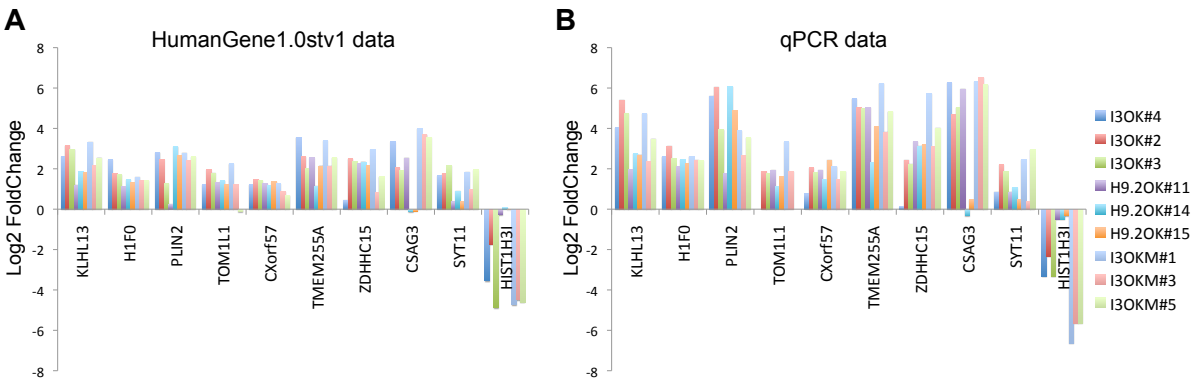


Figure 18: Validation of expression data via qPCR. Fold change expression (Log2) of representative genes for pairwise comparisons of iPSC-NSCs with their corresponding ESC-derived counterparts, obtained by the HumanGene1.0stv1 assay **(A)** and qPCR **(B)**.

To obtain a clearer picture, the transcriptional data of ESC- and iPSC-NSCs was separately analyzed for autosomal genes and X-chromosomal genes.

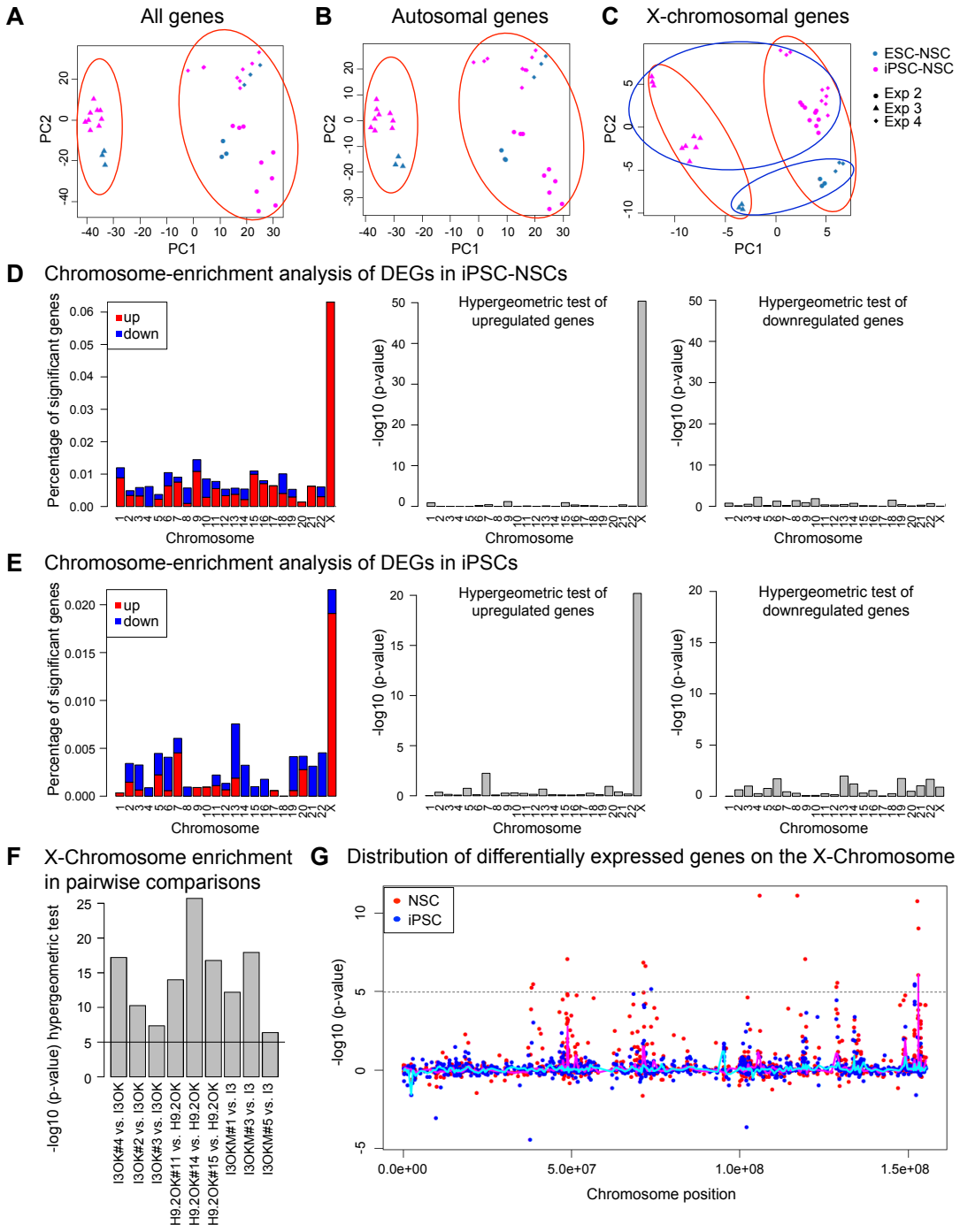


Figure 19: Differential expression is enriched on the X-Chromosome. Principal component analysis of gene expression data (NSC samples only) of **(A)** all analyzed genes, **(B)** autosomal genes and **(C)** X-chromosomal genes. PC2 shows a clear overall separate clustering of iPSC- and ESC-derived NSC samples when analyzing X-chromosomal genes. Red and blue circles indicate separate clustering of different genetic backgrounds (I3 vs. H9.2) and ESC- and iPSC-derived samples, respectively. **(D)** Chromosome-enrichment analysis of DEGs when comparing iPSC- with ESC-NSCs. Graphical representation of significant genes (adj. p-value <0.01) in percentages of measured genes on a given chromosome as well as hypergeometric test. Up- and down-regulated fractions are indicated in red and blue, respectively. **(E)** Chromosome-enrichment analysis of DEGs when comparing iPSCs with ESCs. Graphical representation of significant genes (adj. p-value <0.01; FC>2) in percentages of measured genes on a given chromosome as well as hypergeometric test. Up- and down-regulated fractions are indicated in red and blue, respectively. **(F)** X-Chromosome enrichment for all pairwise comparisons (p-value < 10^{-5}) P-value cutoff for individual genes: 0.05 (t-test, Benjamini-Hochberg-adjusted) **(G)** Distribution of DEGs over the X-chromosome. Pink and cyan lines indicate local enrichment in NSCs and pluripotent cells, respectively.

Irrespective of whether all genes, autosomal genes or X-chromosomal genes were analyzed, a separation according to the genetic background was observed in principle component 1 (PC1; Figure 19 A-C, red circles). Interestingly, with the exception of the third experiment, no clear overall separation of ESC- and iPSC-NSCs was observed in principle component 1 or 2 when analyzing all genes or autosomal genes only, whereas when investigating X-chromosomal genes only, a separation of ESC- and iPSC-derived cells became evident (blue circles). This is also reflected in the Pearson correlation coefficient, with about 0.98 and 0.97, in autosomal genes and X-chromosomal genes respectively.

To visualize a chromosome-specific enrichment of DEGs, the genomic localization of differentially expressed genes (adjusted p-value <0.01) was investigated and their proportion was plotted against chromosome number (Figure 19 D). Indeed, this revealed a significant overrepresentation of differentially expressed genes on the X-chromosome (75 out of 292 genes; 8 out of the 36 genes with a > 2-fold deviation), which was also true for the pluripotent cell lines (Figure 19 E). Interestingly, essentially all differentially expressed genes on the X-chromosome showed an up-regulation in the iPSC-NSCs whereas no such enrichment was detectable on autosomal chromosomes (Figure 19 D-E).

X-chromosome enrichment of differentially expressed genes between ESC- and iPSC-derived NSCs was also significant for all individual pairwise comparisons (p-value < 10^{-5}, P-value cutoff for individual genes: 0.05, t-test, Benjamini-Hochberg-adjusted) regardless of their genetic background or reprogramming method (Figure 19 F). Furthermore, the distribution of differentially expressed genes on the X-chromosome was investigated, revealing a local enrichment in NSCs and pluripotent cells as indicated by pink and cyan lines, respectively (Figure 19 G). Several hotspots of DEGs on the X-chromosome could be identified, including p11.4, p11.23, q13, q22.3, q24, q25/q26.1 and q28.

Together these data show that isogenic human NSC lines of ESC- and iPSC-origin display a remarkable similar gene expression signature and that many of the identified X-chromosomal genes are already altered in the pluripotent state.

4.2.2 Comparison of different reprogramming strategies

There was a significant X-chromosomal enrichment of DEGs for all pair wise comparisons (Figure 19 F) and analysis of the expression data in dependence on the experiment, genetic background and reprogramming method also revealed an overrepresentation of

differentially expressed genes on the X-chromosome (data not shown). However, the effect was not that prominent amongst all differentially expressed genes in the third experiment (H9.2 background). In general, iPSC-NSCs derived from H9.2 cells displayed a higher variance also on autosomal genes when compared to their ESC-derived counterpart than it was seen for the other experiments (Figure 19 A-C).

4.3 Enriched differential methylation on the X-Chromosome coincides with differential expression

To analyze whether the transcriptional changes might be associated with alterations in DNA methylation, DNA from three I3 iPSC-derived NSC lines (I3OK#2, I3OK#3, I3OK#4) and their matched ESC-derived counterpart (I3OK) were bisulfite converted and hybridized (all in biological duplicates) to an Illumina Infinium Human Methylation 450K bead chip. This array enables a genome-wide analysis of DNA methylation levels of more than 450,000 CpGs covering putative regulatory regions of 99% of all RefSeq genes. To further explore whether any aberrant DNA methylation pattern in NSCs is still preserved after differentiation to neurons, corresponding ESC- and iPSC-NSC-derived neurons (single samples) were also included in the methylation analysis.

This revealed a highly similar methylome between ESC- and iPSC-derived NSCs and neurons with a correlation of about 96% and a small interclonal variance of NSCs and neurons derived from the three different iPSC clones of at least 97%. However, hierarchical clustering of global DNA methylation data points to a higher similarity for neurons and NSCs of each clone than for ESC- and iPSC-derived cells (Figure 20 A). The same was true for principal component analysis (Figure 20 B).

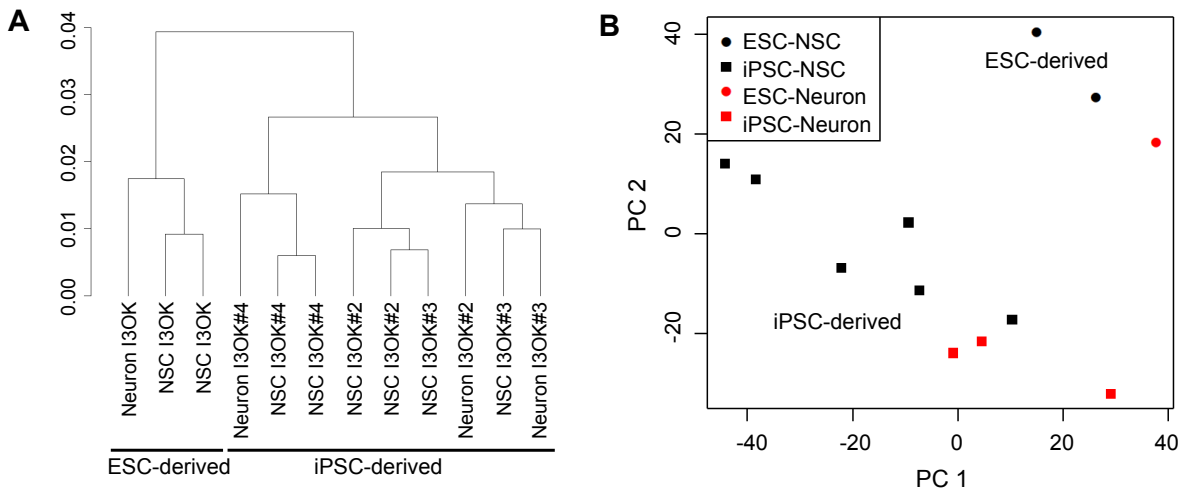


Figure 20: ESC- and iPSC-derived cells show a highly similar methylome. (A) Hierarchical clustering and **(B)** principal component analysis of global DNA methylation profiles of ESC- and iPSC-derived NSCs and neurons. NSC and neurons were analyzed as duplicates and single samples, respectively.

Pairwise comparison of differentially methylated CpGs (DMCGs) between iPSC-NSCs and ESC-NSCs (delta Beta >0.2) revealed an overlap of only 9552 CpGs in all single comparisons (Figure 21 A). Similar results could be seen in iPSC- and ESC-derived neurons (Figure 21 B) and about 67% of DMCGs in NSCs correlated with DMCGs in the neuronal cells (Figure 21 C). In more detail, 9419 of DMCGs in iPSC-derived NSCs were hyper- and 10586 were hypomethylated compared to ESC-NSCs, whereas in iPSC-derived neurons more DMCGs were hypermethylated when compared to their ESC-derived counterpart (Figure 21 D).

There was no overrepresentation of transcription start sites in DMCGs when compared to the overall array annotation (Figure 21 E, upper panels). In iPSC-NSCs, more hypomethylated DMCGs were located in CpG islands and more hypermethylated CpGs in non-island DNA sequences (Figure 21 E, lower panels), which was even more distinct in DMCGs in iPSC-derived neurons.

On gene level, the most prominent alterations in DNA methylation occurred within the genes *VENTX* and *PNPLA4*. iPSC-derived NSCs predominantly displayed hypomethylated CpG sites of the *VENT-like homeobox protein 2* (*VENTX*) gene. The predicted maternally imprinted gene *VENTX* is important for lineage commitment, especially in hematopoiesis (Scerbo et al., 2012; Gao et al., 2012) and plays a role in cell senescence by influencing the WNT-signaling pathway (Gao et al., 2010). However, a correlation between DNA methylation and gene expression could not be detected for the *VENTX* gene.

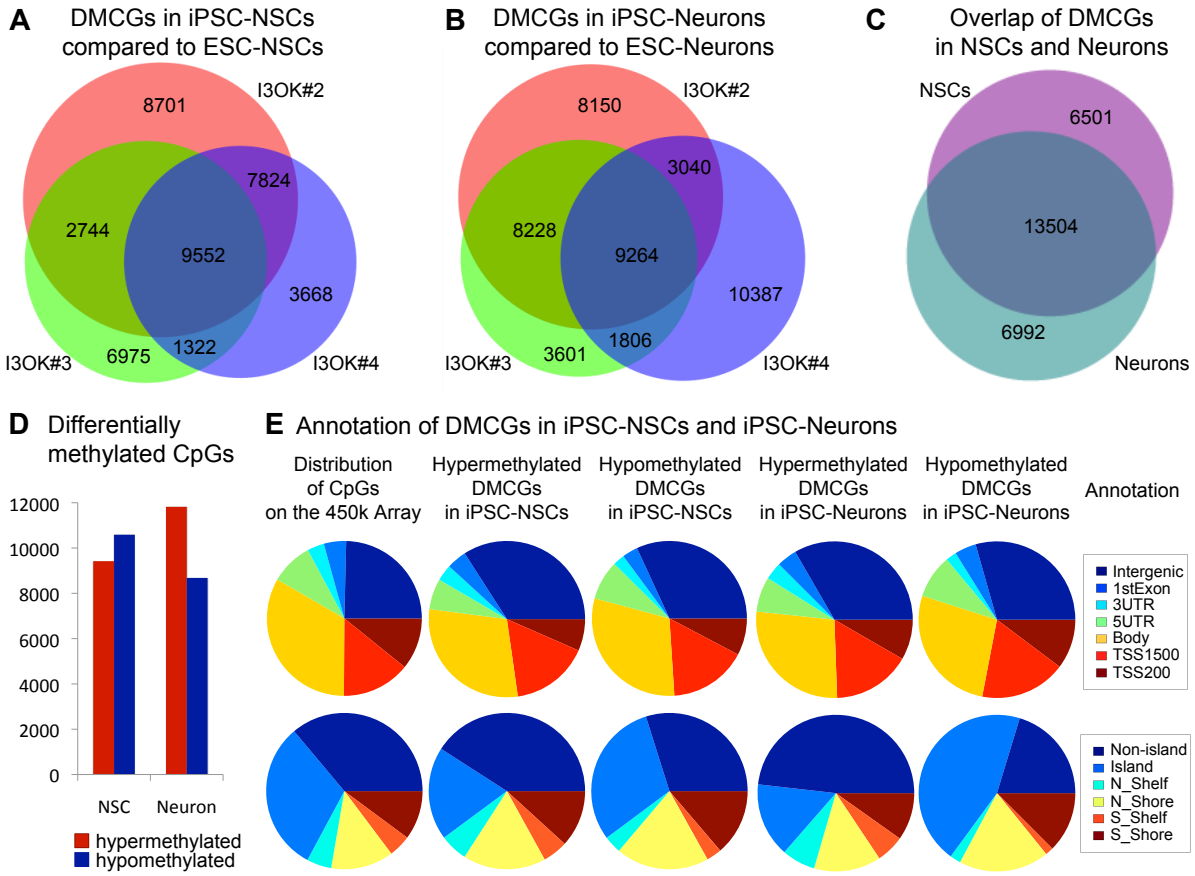


Figure 21: Pairwise comparison of **(A)** iPSC-NSCs with ESC-NSCs and **(B)** iPSC-derived neurons with ESC-derived neurons. **(C)** Overlap of differentially methylated CpGs in NSCs and neurons when comparing the iPSC-derived lines (I3OK#2, I3OK#3, I3OK#4) with their ESC-derived (I3OK) counterpart. **(D)** Hyper- and hypomethylated fractions of DMCGs in NSCs and neurons. **(E)** Annotation of DMCGs in NSCs and neurons in regard of gene regions (upper panels) and CpG islands (lower panels). UTR: untranslated region; TSS: transcription start site; N-shore, S-shore: directly adjacent upstream and downstream CpG island-flanking regions, respectively; N-shelf, S-shelf: directly adjacent upstream and downstream shore-flanking regions, respectively.

Furthermore, iPSC-derived NSCs showed hypermethylation of the X-chromosomal *patatin like phospholipase domain containing 4 gene* (*PNPLA4*). *PNPLA4* encodes for a phospholipases with triacylglycerol lipase and transacylase activity and may be involved in retinol metabolism in keratinocytes (Gao and Simon, 2005) and adipocyte triglyceride homeostasis (Steinberg et al., 2007). Expression of *PNPLA4* was significantly reduced (pairwise comparisons, p-value < 0.02) exclusively in the corresponding iPSC-NSC lines of experiment 2. Moreover, enrichment analysis of all genes with at least one corresponding DMCG (deltaBeta>0.2) between iPSC- and ESC-derived NSCs revealed a slight overrepresentation (p-value = 0.01426) of imprinted genes.

Similar to the observations when analyzing gene expression profiles, hierarchical clustering revealed a much more profound separation of ESC- and iPSC-derived NSCs when investigating X-chromosomal genes separately than in the case of all genes or autosomal

genes with a correlation of 85% (X-chromosomal genes) and 96.5% (autosomal genes), respectively (Figure 22 A-C), indicating a major impact of differentially methylated X-chromosomal CpGs on the overall difference in DNA methylation.

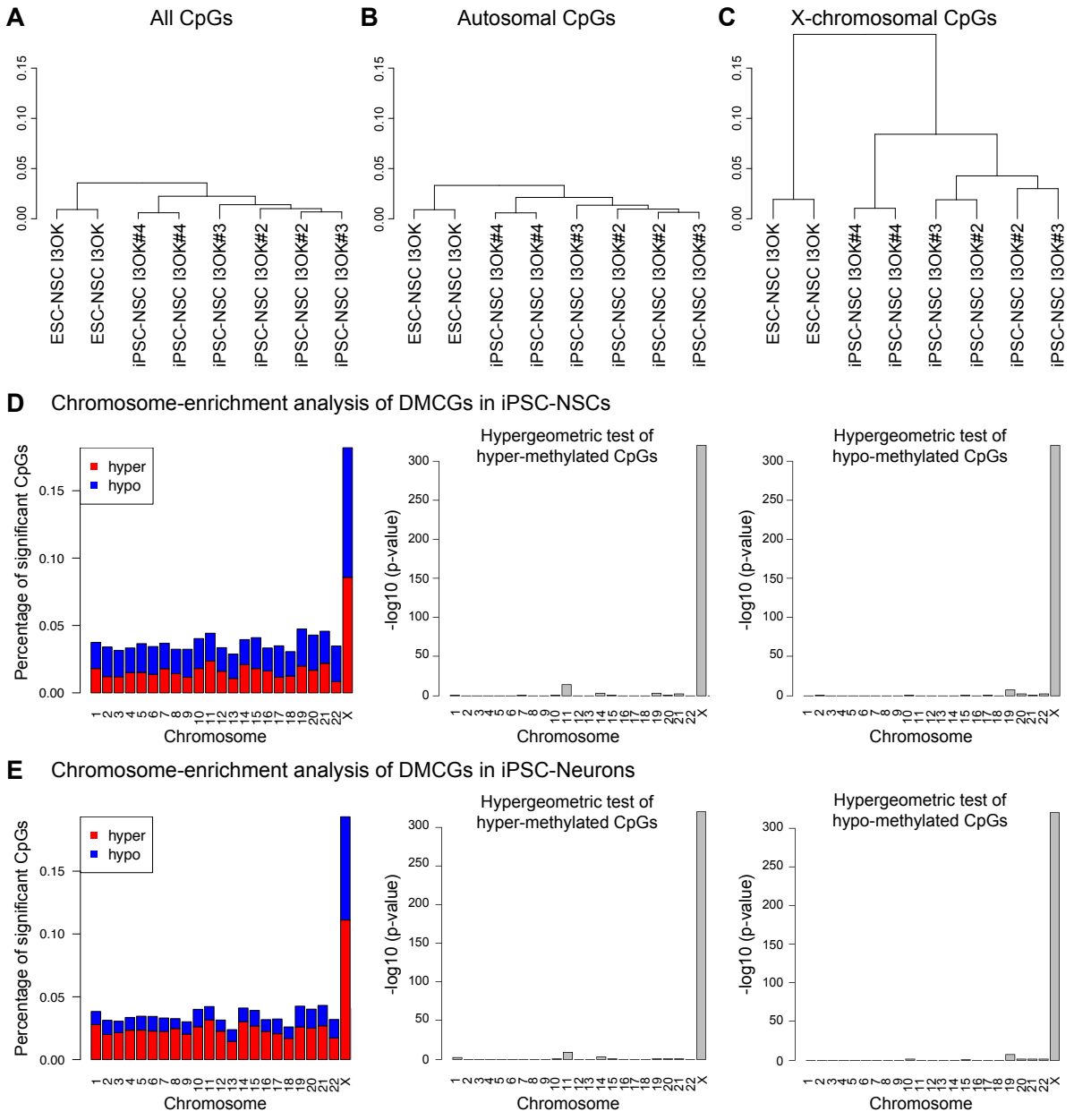


Figure 22: Differential methylation is enriched on the X-chromosome. Global hierarchical clustering of DNA methylation data **(A)** of all analyzed CpGs, **(B)** autosomal CpGs and **(C)** X-chromosomal CpGs. Samples were analyzed in duplicates. **(D)** Chromosome-enrichment analysis of DMCGs in ESC- and iPSC-derived NSCs (delta beta >0.2). Graphical representation of DMCGs in percentages of measured CpGs on a given chromosome. Hyper- and hypomethylated fractions are indicated in red and blue, respectively and hypergeometric tests of hyper- and hypomethylated CpGs. **(E)** Chromosome-enrichment analysis of DMCGs in ESC- and iPSC-derived neurons (delta beta >0.2). Graphical representation of DMCGs in percentages of measured CpGs on a given chromosome. Hyper- and hypomethylated fractions are indicated in red and blue, respectively and hypergeometric tests of hyper- and hypomethylated CpGs.

When plotting the differentially methylated CpGs between iPSC- and ESC-derived NSCs according to the respective chromosomes a clear overrepresentation on the X-chromosome was detectable (delta beta >0.2; Figure 22 D). Interestingly, many X-chromosomal CpGs were either hypermethylated or hypomethylated in iPSC-derived NSCs whereas the ESC-derived NSCs showed a beta-value of approximately 0.5 in those CpGs (Figure 22 D, Figure 23 A). This differential methylation pattern was retained after differentiation to neurons, which resembled the overall chromosomal distribution as seen in NSCs but with a higher percentage of hypermethylated CpG sites on all chromosomes (Figure 22 E).

Moreover, differentially methylated CpGs in iPSC-NSCs (both hyper- and hypomethylated) correspond in part regionally precisely with differentially expressed genes as seen by coherent local enrichment of those differentially methylated CpGs with local enrichment of differentially expressed genes (Figure 23 B).

In summary, the methylation pattern of ESC- and iPSC-derived NSCs correlate to the observed gene expression pattern with only minimal changes observed between both populations and a clear accumulation of alterations on the X-chromosome.

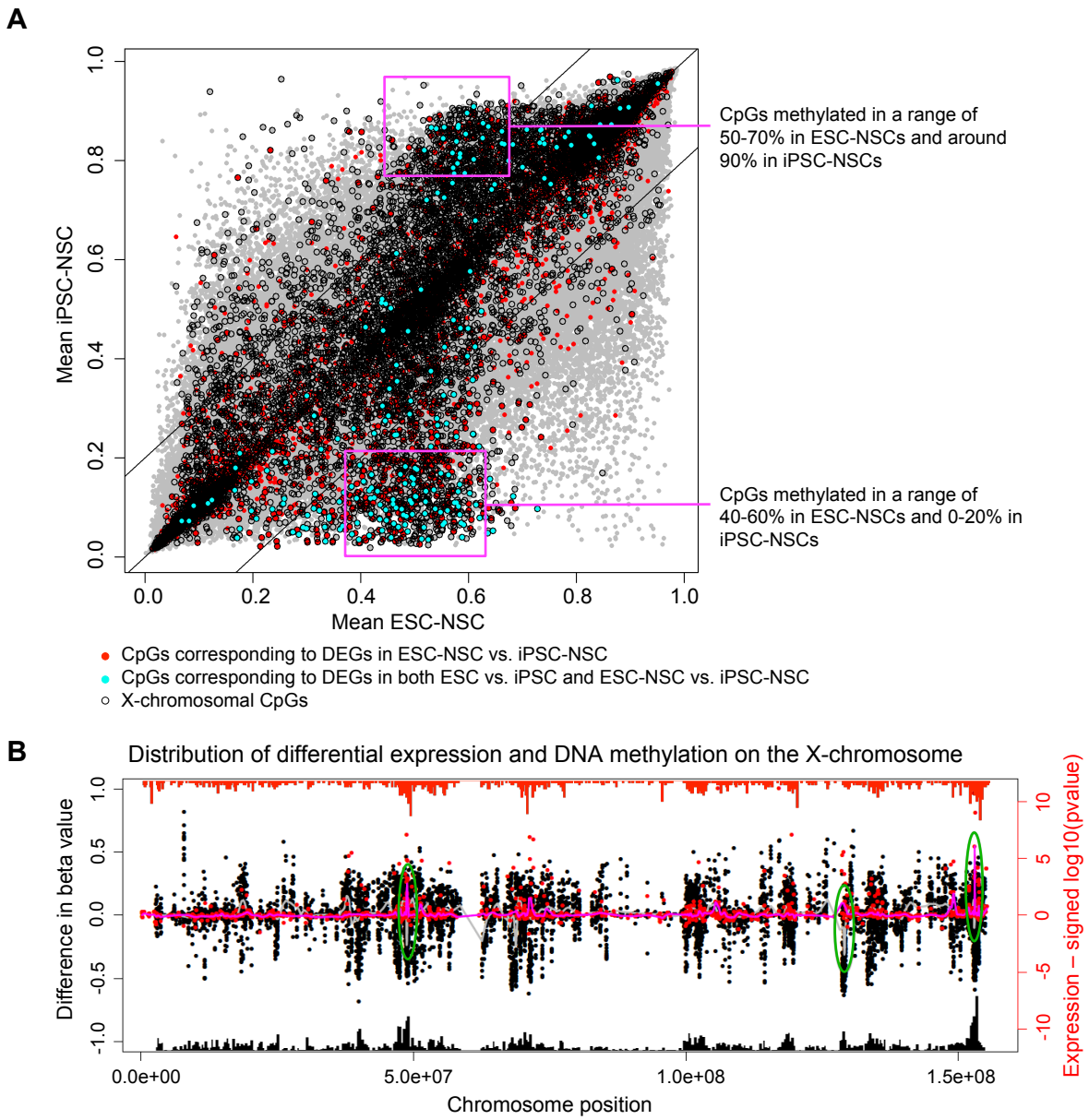


Figure 23: Enriched differential methylation on the X-chromosome coincides with differential expression. **(A)** Scatter plot of global DNA methylation patterns comparing iPSC-NSCs with ESC-NSCs. Lines indicate a difference in beta > 0.2 **(B)** Distribution of DEGs and DMCGs over the X-chromosome are indicated with red and black dots, respectively. All analyzed probes and CpGs are indicated with red and black histogram on the outer axis, respectively. Pink and gray lines indicate local enrichment of DEGs and DMCGs comparing iPSC-NSCs with ESC-NSCs, respectively. Green circles indicate coherent differences in gene expression and DNA methylation.

4.4 Implications of the reprogramming process on the X-chromosome inactivation status

To test whether the differences in expression and methylation between reprogrammed and non-reprogrammed cells are based on alterations in the X-chromosome inactivation status of the cell lines, *XIST* expression in ESC- and iPSC-derived NSCs as well as female primary NSCs (primary hindbrain neuroepithelial stem cells; Sun et al., 2008) was analyzed. Primary NSCs underwent regular *in vivo* differentiation and therefore should display regular X-

chromosome inactivation. Transcriptional analysis of *XIST* via qPCR revealed no significant difference in its expression comparing neural stem cells before and after reprogramming, but lower expression levels in ESC- and iPSC-derived lines than in primary neural stem cells (Figure 24 A). This may suggest a loss of this inactivation mark already in the starting populations.

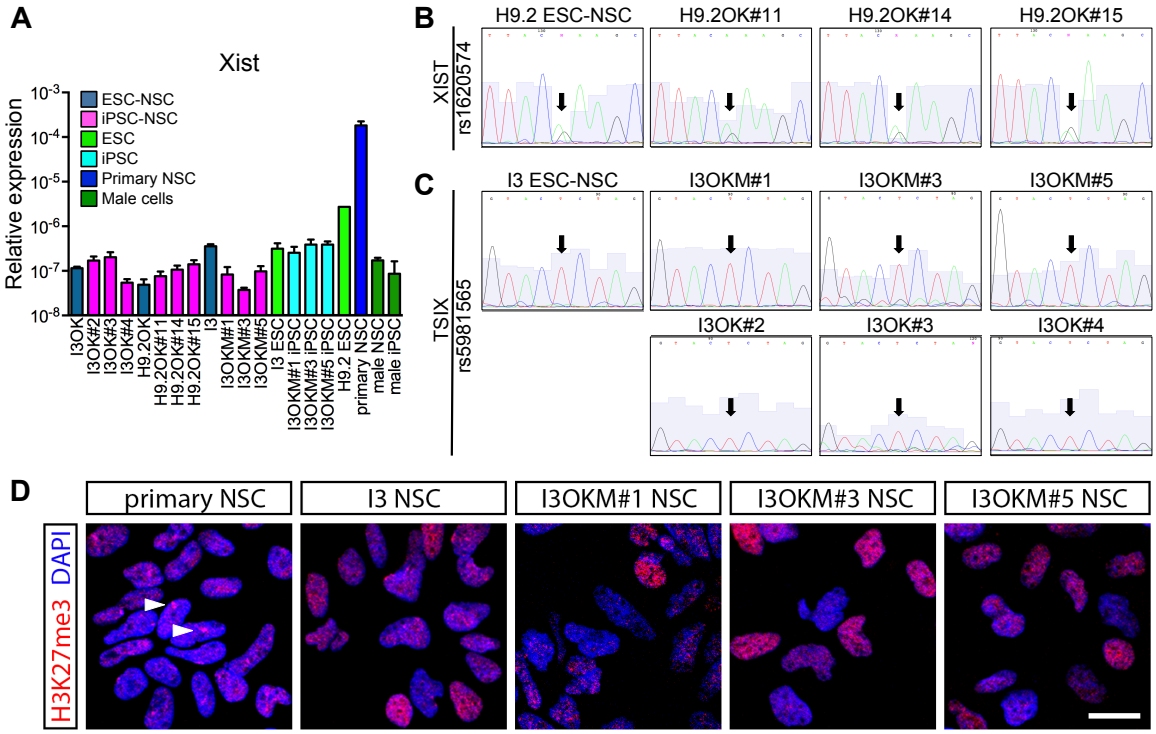


Figure 24: XCI status in iPSC-derived NSCs. **(A)** Relative *XIST* expression of ESC- and iPSC-derived NSCs. Primary NSC: female primary hindbrain neuroepithelial stem cell line SAi5; male NSC: I6 ESC-derived NSCs; male iPSC: I6 NSC-derived iPSCs. **(B, C)** Sequencing data of X-chromosomal genes with heterozygous SNPs. iPSC-derived NSCs show a biallelic expression of *XIST* and a monoallelic expression of *TSIX* with no skewing to the other X-chromosome after reprogramming. Arrows indicate the site of the heterozygous SNP. Blue background indicates Phred score for each base call. **(D)** H3K27me3 heterochromatin staining of ESC- and iPSC-derived NSCs. Arrow heads indicating H3K27me3 accumulations. Scale bar: 20µm.

To assess whether the same X-chromosome was inactivated before and after the reprogramming process, allele-specific expression of the long non-coding RNA *XIST* (X-inactive specific transcript), which is the key player of X-chromosome inactivation, and its antisense RNA *TSIX* were analyzed. To that end, the previously obtained SNP analysis data was screened for heterozygous SNPs (BAF of 0.5) in the genetic loci of *XIST* and *TSIX* and the RT-PCR-products of both RNAs containing corresponding SNPs were sequenced. Since a heterozygous SNP could only be identified for *XIST* in the H9.2 genetic background and for *TSIX* in the I3 genetic background, solely the respective cell lines could be analyzed for their allele-specific expression. This revealed a biallelic expression of *XIST* and a monoallelic

expression of *TSIX* with no skewing to the other X-chromosome in all lines analyzed (Figure 24 B, C).

Furthermore, X-chromosome inactivation is characterized by the formation of heterochromatin. As such, XCI leads to an accumulation of the repressive histone modification H3K27me3 on the inactivated X-chromosome. Likewise to *XIST* expression, this inactivation mark has been described to possibly erode in iPSCs (Mekhoubad et al., 2012). Analysis via immunocytochemical staining revealed a lack of the typical H3K27me3-foci in most of the cells irrespective of their ESC- or iPSC-origin, implying a deterioration of this mark already in the starting population, which seems not to be reestablished by the reprogramming process (Figure 24 D).

To test whether the loss of those XCI marks actually leads to biallelic expression and thereby to altered transcript levels of X-linked genes, allele-specific expression analysis was also performed on X-chromosomal genes containing heterozygous SNPs. However, the presence of heterozygous SNPs in the identified DEGs was limited and therefore only two genes could be analyzed for their allele-specific expression. For the first gene, *HEPH*, contains two heterozygous SNPs in both genetic backgrounds (I3 and H9.2) but was only in a few iPSC-derived NSC lines differentially expressed and was therefore not amongst the overall identified DEGs. The second gene analyzed, *TMEM255A*, was differentially expressed in all analyzed iPSC-derived lines (FC>2) but contains only a heterozygous SNP in the H9.2 background, therefore only H9.2-derived NSCs were analyzed. Both genes are typically affected by XCI (Vallot et al., 2015).

A distinct loss of monoallelic expression of *HEPH* could be identified only in one iPSC-derived NSC line (H9.2OK#15), whereas X-chromosomal gene *TMEM255A* remained monoallelically expressed in all H9.2-derived NSCs (Figure 25 A). Interestingly, the iPSC-derived NSC cell line showing biallelic expression did not display an elevated expression of *HEPH* compared to ESC-derived NSCs and cell lines exhibiting elevated gene expression in iPSC-derived NSCs stayed monoallelically expressed (Figure 25 A, B). The different elevated expression levels of *TMEM255A* could also not be associated with biallelic expression.

These data suggest that, XCI marks are altered already in the starting population of the applied isogenic system yet, at least in the analyzed X-chromosomal genes, the differential transcript levels between iPSC- and ESC-NSCs do not correlate with biallelic expression.

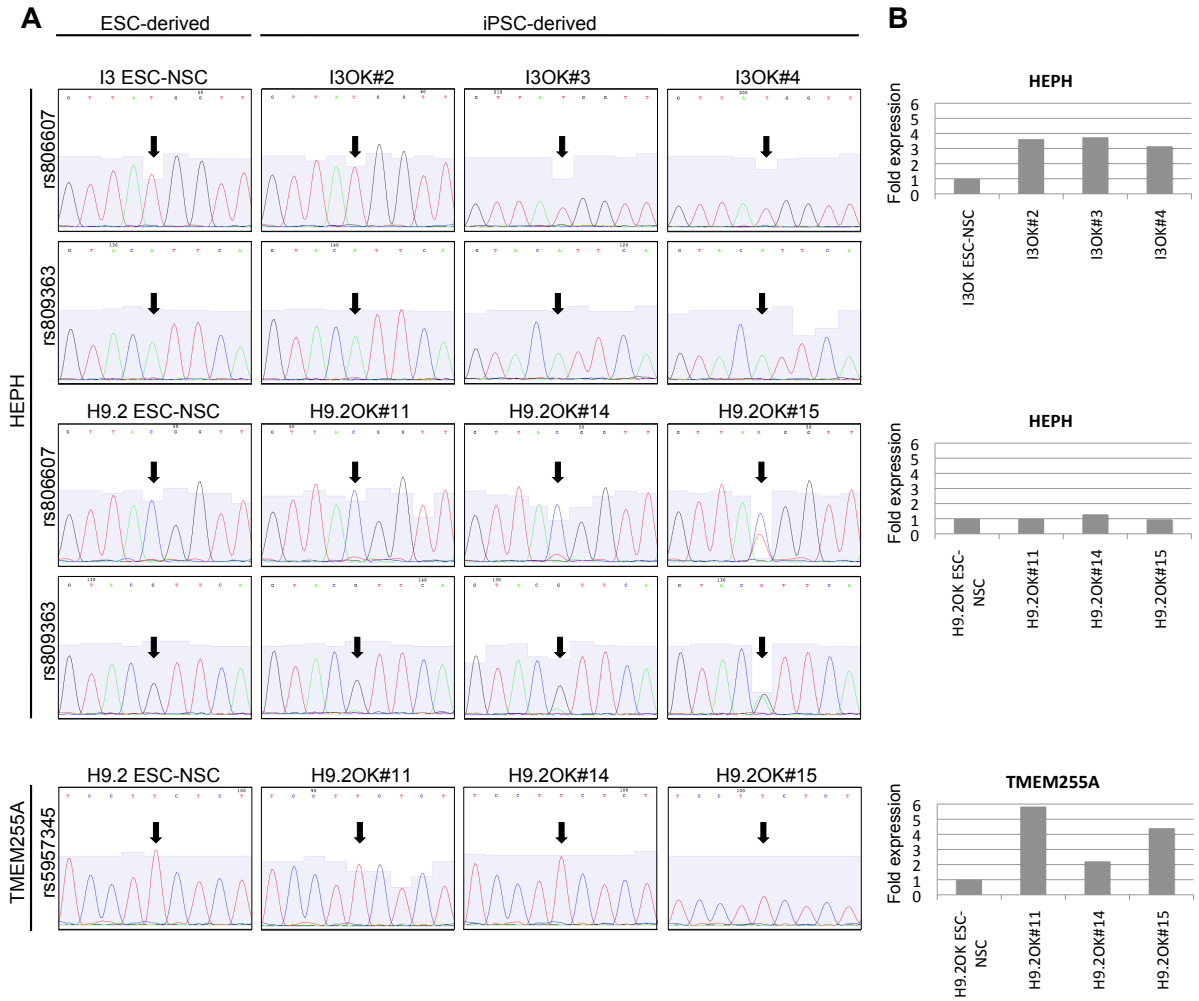


Figure 25: Allele-specific X-chromosomal gene expression in iPSC-derived NSCs. **(A)** Sequencing data of X-chromosomal genes with heterozygous SNPs. iPSC-derived NSCs show partial loss of monoallelic expression of X-chromosomal gene *HEPH* (rs806607, rs809363), whereas X-chromosomal gene *TMEM255A* (rs5957345) remains monoallelically expressed as confirmed by Sanger sequencing. Arrows indicate the site of the heterozygous SNP. Blue background indicates Phred score for each base call. **(B)** Corresponding fold expression of *HEPH* and *TMEM255A* in iPSC-NSCs.

5 Discussion

Reprogramming to pluripotency is a complex process with the requirement of massive epigenetic rearrangements that might affect diverse cellular characteristics. Several studies in the mouse and human system already addressed the question whether iPSCs resemble ESCs in their cellular properties. However, their equivalence remained controversial and in order to specifically evaluate whether cells derived from iPSCs did not acquire any characteristic that corrupt downstream applications such as iPSC-based disease modeling and regenerative approaches, these studies have to be extended to human iPSC-derived somatic cell populations. Thus, this work focused mainly on ESC- and iPSC-derived NSCs in the established human isogenic system, in order to identify reprogramming-associated alterations. First, a comprehensive analysis of transcriptional profiles of ESC- and iPSC-derived NSCs was performed. Second, global DNA methylation analysis widened the insight into reprogramming-associated changes and could be correlated to the expression data. Since, alongside to a remarkably similar transcriptome and methylome, predominantly X-chromosomal genes and CpGs were found to be altered, the third part of this thesis focused on deciphering the coherence of reprogramming and X-chromosome inactivation.

5.1 Assessment of the isogenic stem cell system

In the course of this study, three different approaches were applied to generate human iPSCs. With the retroviral transduction of the reprogramming factors, the first described classical strategy to convert somatic cells into iPSCs (Takahashi and Yamanaka, 2006; Takahashi et al., 2007) was included in this study. Retroviruses can provide a convenient vehicle for the stable insertion of exogenous DNA into the host genome of mitotically active mammalian cells (Barquinero et al., 2004; Aasen et al., 2008). Due to a moderate efficiency and reasonable inadequacy in silencing, especially of the *KLF4* transgene in several iPSC clones, the reprogramming method was extended to a more advanced strategy based on tetracycline-dependent expression of stably incorporated reprogramming factors via lentiviral transduction (Wernig et al., 2008; Maherali et al., 2008; Hockemeyer et al., 2008). However, the usage of retroviral or lentiviral vectors to deliver the reprogramming factors causes, due to their random genome-integrating nature, a variety of insertions into the genome of the target cells, which could potentially cause disruption of endogenous genes or aberrant activation of adjacent genes and also pose a general risk of possible transgene-

reactivation at a later point in time (Wernig et al., 2007; Yu et al., 2007; Okita et al., 2011; Ramos-Mejía et al., 2012). This caused serious concerns regarding the suitability and therapeutic safety of cells generated via genome-integrating reprogramming approaches (Sun et al., 2010). Indeed, an elevated tumor formation propensity in germline-competent mouse iPSC-chimera progeny has been described, which could be attributed to the reactivation of transgenic oncogene c-Myc (Okita et al., 2007), and it has been shown that an insufficient silencing of reprogramming transgenes can impede lineage-specific differentiation (Hochedlinger et al., 2005; Yu et al., 2007; Ramos-Mejía et al., 2012; Toivonen et al., 2013). Thus, though the used retroviral vectors should promote transgene suppression once pluripotency is reached (Yao et al., 2004; Takahashi and Yamanaka, 2006) and tetracycline-dependent expression should be terminated upon withdrawal of tetracycline in the culture media (Wernig et al., 2008; Maherali et al., 2008; Hockemeyer et al., 2008), residual transgene expression or reactivation can pose a serious problem for their safety and reliability in later applications (Okita et al., 2007; Ramos-Mejía et al., 2012; Sommer et al., 2012; Toivonen et al., 2013).

To avoid these possible drawbacks, various alternative induction strategies have been developed in the recent years to either (i) subsequently remove integrated reprogramming factors from the established iPSC lines via transposase-mediated withdrawal of the transgene-containing transposons in *piggyBac* approaches or via *Cre* recombinase treatment if flanking *loxP* sites are present in the transgene cassette (Woltjen et al., 2009; Kaji et al., 2009; Yusa et al., 2009) or (ii) to transiently express the reprogramming factors via integration-free methods including episomal vectors (Okita et al., 2008, 2011; Yu et al., 2009; Jia et al., 2010; Chou et al., 2011), adenoviruses (Stadtfeld et al., 2008b; Zhou and Freed, 2009), Sendai viruses (Fusaki et al., 2009; Nishimura et al., 2011, 2017), synthetic RNAs (Warren et al., 2010; Mandal and Rossi, 2013; Yoshioka et al., 2013), recombinant proteins (Zhou et al., 2009; Kim et al., 2009b) and even chemically, without the aid of any transgene, via small molecules in mouse fibroblasts (Hou et al., 2013). However, despite this variety of alternative approaches, most of them possess a lower efficiency than the first-generation methods and include inconvenient repetitive transfection rounds, laborious subcloning and screening for transgene-free iPSC clones (Robinton and Daley, 2012).

Because of its reasonable efficiency and feasible handling, Sendai-viral transduction of reprogramming factors was chosen as additional integration-free method in this study. This

unsegmented negative-strand RNA-based virus self-replicates exclusively in the cytoplasm of infected cells, thus neither an intermediate DNA stage nor chromosomal integration occurs (Bitzer et al., 2003). Alongside to increased temperature sensitivity caused by mutations, used Sendai-viral vectors are incapable of transmissible-virion production, due to a deletion of the *envelope fusion gene* (*f-gene*). Therefore, vector concentration attenuates upon cell division and gets lost over time in culture, resulting in transgene-free and virus-free iPSC lines (Inoue et al., 2003; Fusaki et al., 2009; Ban et al., 2011).

Furthermore, alternatively to the fist described OSKM reprogramming cocktail, diverse combinations of transcription factors had been used for reprogramming to pluripotency. Studies in the recent years revealed, that single factors of the OKSM cocktail could be replaced with others or even omitted (Nakagawa et al., 2008) whereas the combination of OCT4, SOX2, NANOG and LIN28 is the most prominent among these alternative combinations. Although these alternate procedures do exist and though l-Myc demonstrated a much lower oncogenic transformation activity than c-Myc (Nakagawa et al., 2010), this study focused on the originally described OSKM factors since these still represent the most widely used combination for reprogramming. Nevertheless, SOX2 was omitted in the reprogramming of NSCs, as these already express this factor on a level comparable to pluripotent cells and therefore an additional supplementation was not necessary. Furthermore, c-Myc was excluded in some reprogramming attempts because of its described carcinogenic properties and to decipher its effects on iPSCs and iPSC-derived NSCs.

Moreover, reprogramming to pluripotency is not always a complete process. In fact, the majority of cells during this cellular conversion only gets partially reprogrammed, some of which may reach pluripotency at a later time point, others may become arrested in pre-iPSC stages instead (Mikkelsen et al., 2008; Chan et al., 2009). These intermediate or partially reprogrammed cells display reactivation of certain early onset pluripotency-related genes, but insufficient suppression of lineage-specific transcription factors and often remain transgene dependent (Takahashi and Yamanaka, 2006; Mikkelsen et al., 2008; Buganim et al., 2012). During the course of the presented study, partial reprogramming could also be noticed, as colonies formed during the reprogramming process repeatedly stopped growing, collapsed or showed unspecific cell proliferations. Hence, accurate choice and monitoring of

iPSCs is essential to verify bona fide pluripotency and to exclude incompletely reprogrammed iPSC lines.

5.1.1 Induced pluripotent stem cells display bona fide pluripotency

The combination of multiple specific characteristics reliably demonstrates complete pluripotency of human iPSCs. Morphological identifiers for successful reprogramming are a flat, roundish and compact colony shape with clearly defined borders, consisting of cells with large nuclei, prominent nucleoli and a high nucleus-to-cytoplasm ratio and furthermore, these cells show a fast proliferative behavior (Takahashi et al., 2007; Ghule et al., 2008; Smith et al., 2009; Chan et al., 2009; Robinton and Daley, 2012). Only colonies meeting these morphologic characteristics were selected to generate clonal iPSC lines. Furthermore, pluripotent stem cells are characterized by a set of several molecular features, like a high alkaline phosphatase and telomerase activity, the expression of characteristic cell-surface antigens (TRA-1-60, TRA-1-81, SSEA3), as well as transcription of specific miRNAs and pluripotency-associated transcription factors NANOG, OCT4, REX1 and SOX2 (Nichols et al., 1998; Henderson et al., 2002; Mitsui et al., 2003; Suh et al., 2004; Wernig et al., 2007; Smith et al., 2009). All iPSC lines evaluated in this study constantly possessed characteristic pluripotent morphology, were positive for alkaline phosphatase activity and expressed characteristic pluripotency markers.

However, ultimate functional proof of pluripotency is the capability to differentiate into cells of all three embryonic germ layers. This potential can be assessed by undirected differentiation *in vitro* via embryoid body (EB) formation. These in suspension spontaneously formed free-floating spherical aggregates structurally mimic the pregastrulation stage of embryonic development (Itskovitz-Eldor et al., 2000; Rust et al., 2006). In fact, all evaluated iPSC lines produced meso-, endo- and ectodermal offspring. However, EB-assays cannot imitate the complete complex structural interactions of organogenesis during embryonic development. *In vivo* data therefore provide a more stringent quality indicator (Smith et al., 2009). While mouse iPSCs injected into a host blastocyst contribute to germ line competent chimeras or even show full-term development via tetraploid complementation (Okita et al., 2007; Kang et al., 2009), this was not applicable with human iPSCs, due to ethical reasons. Alternatively, the next stringent method, the teratoma formation assay, was conducted with a subset of the generated iPSC lines via intratesticular injection into immunodeficient

SCID/beige mice, resulting in teratoma-like growth containing derivates of all three germ layers (Smith et al., 2009). To conclude, all applied reprogramming approaches successfully generated iPSC lines with full pluripotent potential.

Potential obstacles to the application of retroviral-mediated reprogramming were revealed in the evaluation of transgene silencing. Several iPSC lines still displayed residual transgene expression, especially for the KLF4 transgene, and where therefore excluded from subsequent experiments. This highlights also the need to generate and analyze multiple iPSC lines to retain the option to omit deficient ones. Furthermore, all tetracycline-inducible lentiviral-reprogramming approaches showed down-regulation of transgenes upon doxycycline withdrawal and no residual viral RNA could be detected in Sendai-viral reprogramming after several passages in culture.

Moreover, the genomic integrity of the established iPSC lines was analyzed via genome-wide high-resolution single nucleotide polymorphism (SNP) to correlate results of consecutive experiments with possible aberrations. Indeed, retroviral and lentiviral reprogrammed iPSCs acquired some genomic alterations when compared to their cell lines of origin, whereas lines generated via non-integrating Sendai virus were karyotypically normal. Higher copy number variations (CNVs) in integrating iPSCs than in non-integrating were described in the literature (Sugiura et al., 2014; Steichen et al., 2014; Kang et al., 2015b). However, other studies revealed neither significant differences in the incidence of chromosomal malformation between different reprogramming methods (Gore et al., 2011; Bhutani et al., 2016) nor between iPSCs and ESCs lines (Taapken et al., 2011; Turinetto et al., 2017). In actuality, pluripotent stem cells acquire, during prolonged culture, genomic aberrations with a bias for major hotspots on chromosomes 12, 17, 20 and X (Draper et al., 2004; Lefort et al., 2008; Laurent et al., 2011; Taapken et al., 2011; Martins-Taylor et al., 2011) and minor defects on chromosomes 1, 2, 7, 14, 21, 22 (Närvä et al., 2010). Since several evaluated lentivirally reprogrammed iPSCs exhibit similar defects, in terms of their size and location, this may indicate that these variations were already present as a mosaic at a low undetectable percentage in the stable inducible NSC line of origin and may had selective advantageous effects on the reprogramming process itself and/or on the proliferation behavior in the pluripotent state (Mayshar et al., 2010; Martins-Taylor et al., 2011; Abyzov et al., 2012). Furthermore, there are indications for human pluripotent stem cell derived neural progenitor cells that suggest that these copy number variations are especially gained during

culture and manipulation of cells (Corrales et al., 2012).

5.1.2 iPSCs possess comparable neural differentiation potential to ESCs

In the past, some studies indicated an impaired neural differentiation propensity in iPSCs (Hu et al., 2010; Boulting et al., 2011). Hu and colleagues performed directed neural differentiation of human iPSCs and ESCs. On the one hand they were able to show, that iPSCs differentiated into neuroepithelial cells, neurons and glia, on the other hand, over the same time course, they obtained reduced efficiencies in comparison to ESCs. While others like Boulting and colleagues found comparable efficacies of motor neuron differentiation for the majority of analyzed iPSC lines (Boulting et al., 2011; Marei et al., 2017) and that variances, between iPSCs and ESCs, in neural differentiation rather reflect impairment of individual iPSC lines than a general issue of induced pluripotency (Boulting et al., 2011; Toivonen et al., 2013).

However, in the presented study, no difference in differentiation propensity between iPSCs and ESCs could be detected, neither in the generation of NSC nor in undirected terminal differentiation of the generated NSCs. During differentiation into NSCs, iPSC lines gave rise to neural tube-like structures with comparable efficacies like ESCs and resulting iPSC-derived NSC populations showed similar characteristic morphology, rosette-like clustering, proliferation behavior and specific marker expression as seen for their ESC-derived counterparts. Furthermore, there was no detectable difference in differentiation propensity when undirected neural differentiation was performed. iPSC- as well as ESC-derived NSCs differentiated mainly into neurons and only in a small extent into GFAP-positive astrocytes and the neural cultures of both NSC populations contained neurons positive for a GABAergic subtype. Furthermore, the applied reprogramming system differs from most other studies in the aspect that the iPSCs were not fibroblast-derived and as the reprogrammed cells are differentiated into the very same cell type of their origin, residual epigenetic memory of other lineages should therefore not hamper neural differentiation (Ohi et al., 2011; Bar-Nur et al., 2011; Kim et al., 2011b). This “memory” phenomenon was first described in isogenic murine iPSCs derived from different cell types that showed variable differentiation propensities, with higher efficacies in re-differentiation along lineages related to their cell of origin and reduced performance in alternative cell fates (Polo et al., 2010; Kim et al., 2010) and was also detectable in human iPSCs (Ohi et al., 2011; Bar-Nur et al., 2011; Roost et al.,

2017). Epigenetic memory can be explained in part by residual specific DNA methylation signatures of the somatic donor cell (Kim et al., 2010; Bar-Nur et al., 2011). It is especially found in low-passage iPSCs and seems to be largely attenuated by continuous passaging (Polo et al., 2010; Chin et al., 2010; Nishino et al., 2011; Robinton and Daley, 2012). These observations could possibly also explain the discrepancies detected in the differentiation propensity of iPSCs found in some studies. Consequently, for iPSCs derived from fibroblasts, or iPSCs with other non-ectodermal origins, lineage-specific epigenetic marks that may not be correctly erased during reprogramming might lead to an impaired neural differentiation. However, in case epigenetic memory had an effect on the used isogenic stem cell system, it could result in a positive bias of the applied iPSCs/iPSC-derived cells for the neural lineage (Bar-Nur et al., 2011; Pfaff et al., 2012). Still, such an effect was neither detected in undirected differentiation via EB-formation nor during generation or differentiation of iPSC-NSCs.

5.2 Suitability of the human isogenic stem cell system for the analysis of reprogramming-associated alterations

5.2.1 An isogenic background is essential for comparative analysis of iPSCs and ESCs

While this study was conducted several research groups tried to address the question, whether reprogramming leads to *bona fide* pluripotency by comparing iPSCs of different origins with ESCs, which represent the gold standard of pluripotent cell lines, and therefore can reliably and safely be used in regenerative medicine, disease modeling and drug discovery. This resulted in a mere controversial picture as some studies reported differences in expression, DNA and histone methylation, posttranslational modifications, genetic stability, differentiation potential, teratoma formation propensity or a higher variance and clonal abnormalities in iPSCs (Chin et al., 2009; Miura et al., 2009; Ghosh et al., 2010; Hu et al., 2010; Lister et al., 2011; Ohi et al., 2011; Nazor et al., 2012; Ruiz et al., 2012), whereas others reported high comparability and difficulties or inability to distinguish between lines of both pluripotent populations (Guenther et al., 2010; Newman and Cooper, 2010; Bock et al., 2011; Mallon et al., 2014; Féraud et al., 2016; Marei et al., 2017).

Trying to dissect those different findings revealed a major obstacle in the diverging study-design with different sample sizes (biological and technical replicates), (number of) genetic backgrounds, lines of opposite sex, different reprogramming methods (integrating/non-

integrating, different factor combinations) and culture conditions and duration that presumably influenced their outcome (Newman and Cooper, 2010; Stadtfeld et al., 2012; Yamanaka, 2012). There are indications that different genetic backgrounds mainly account for transcriptional variance between iPSCs and ESCs and therefore a large number of donors is important in comparative analysis (Rouhani et al., 2014). Moreover, Phanstiel and colleagues were able to highlight, that additionally to the analysis of several cell lines, the analysis of multiple biological replicates is important to overcome biological and technical variance and generate statistically significant results in order to detect reprogramming-associated differences rather than single sample-specific aberrations (Phanstiel et al., 2011). Indeed, in general it seems that studies with low sample sizes rather identified differences between iPSCs and ESCs, whereas studies with larger sample sets found it difficult to distinguish between both cell types and rather identified individual clone-specific differences (Yamanaka, 2012). However, though greater sample sets may reduce the impact of diverse confounders, like different genetic backgrounds, they may also increase the experimental noise, e.g. in transcriptional and DNA methylation data, which can mask subtle but relevant molecular differences.

Therefore, for adequate sensitivity it is important to exclude or account for as many confounding variances as possible. Some studies have already taken this into account and eliminated one of the main interfering factors, the different genetic backgrounds, by using isogenic cell systems to study reprogramming-related differences and their biological implications. Isogenic iPSCs and ESCs were found to obtain similar differentiation potentials into myeloid and erythroid lineages (Féraud et al., 2016) and no differentiation bias could be identified in undirected EB-differentiations of isogenic iPSCs and ESCs via transcription analysis of 77 developmental markers of all three germ layers (Choi et al., 2015). Furthermore, while non-isogenic approaches identified 1267-3947 genes as differentially expressed via transcriptional analyses (Chin et al., 2009; Marchetto et al., 2009), genetically matched comparative analysis of iPSCs and ESCs could narrow down the quantity of altered transcripts to 49-154 or even found no difference in gene expression (Teichroeb et al., 2011; Mallon et al., 2014; Choi et al., 2015). Indicating that diverse genetic origins indeed can profoundly influence the results of comparative analysis of iPSCs and ESCs. Teichroeb and colleagues compared six iPSC lines with isogenic female ESCs (H9) and found a highly similar transcriptome with only minor differences between both populations (102-154 genes with 2-

fold deviation) and about 3.9 times as many when comparing unrelated iPSCs and ESCs (Teichroeb et al., 2011). These results were supported by Choi and colleagues comparing iPSC and isogenic ESCs (three lines each) from two male backgrounds (HUES2 and HUES3) identifying only 49 differentially expressed genes (2-fold) and similar DNA methylation pattern in matching iPSCs and ESCs, whereas greater differences were found to be connected with different genetic backgrounds (Choi et al., 2015). Furthermore, Mallon and colleagues could not find any statistical significant transcriptional difference between three iPSC lines and isogenic ESCs (Welch-modified t-test) and even detected higher variance between individual samples than between the two pluripotent populations (Mallon et al., 2014).

5.2.2 Profiling of somatic cells is essential to approve suitability of iPSCs for research and therapy

Although there are some studies that have compared isogenic cell lines from both pluripotent populations, little is known about the possible influences of the reprogramming process on somatic cells and whether any differences that may exist are functionally relevant for their use in regenerative medicine, disease modeling and drug discovery. Possibly, alterations may not be detectable in the pluripotent state but, nevertheless, could become relevant in somatic cell type used in research and therapy. First studies suggest also high similarities of isogenic iPSC- and ESC-derived somatic cell populations. Analyzing global transcriptional pattern, Mallon and colleagues could not identify any major difference between ESC- and iPSC-derived embryoid bodies differentiated under mesodermal- and ectodermal-specific media conditions (Mallon et al., 2014). Furthermore, Choi and colleagues performed RNA sequencing analysis (RNAseq) of genetically matched iPSC- and ESC-derived fibroblast-like cells identifying only two genes as differentially expressed (Choi et al., 2015). Furthermore, besides to parthenogenetic mesenchymal stem cells (MSCs), Vassena and colleagues also analyzed and compared two biparental isogenic iPSC-derived MSC lines with their ESC-derived counterparts via genome-wide expression analysis, identifying a substantially higher amount of differentially expressed genes in parthenogenetic cells and only 165 altered genes before and after reprogramming in biparental MSCs (Vassena et al., 2012). However, low sample size or rather heterogenic cell populations and relatively unspecific cell types may obscure minor differences. Therefore, a

defined and stable cell type with relevance for research and therapy would be preferable for comparative analysis of iPSC- and ESC-derived somatic progeny. The presented isogenic stem cell system is based on a specific neural stem cell population that qualifies as a robust, homogenous and well-defined population for these criteria (Koch et al., 2009; Falk et al., 2012; Fujimoto et al., 2012; Doerr et al., 2015; Ali et al., 2016; Poppe et al., 2018). Appropriate sensitivity to detect even subtle differences between iPSC- and ESC-derived populations also seems to be achieved as even minor changes, like a different supplier of the same media component, were detectable. In order to prevent impairment of later data interpretation by this confounder, the first generated data set was therefore excluded from further analysis.

5.3 iPSC- and ESC-derived NSCs display a highly similar transcriptome and DNA methylation pattern

Global transcriptional analysis revealed a remarkable similar transcriptome for isogenic ESC- and iPSC-derived NSCs, and unsupervised clustering via principal component analysis indicated that differences between different genetic backgrounds (H9.2- and I3-derived NSCs) were more prominent than those between isogenic ESC- and iPSC-derived NSCs. This is in line with other studies, identifying distinct genetic backgrounds as major contributor of transcriptional and epigenetic variation whereas differences caused by the cellular origin of iPSC and ESCs only account very little (Rouhani et al., 2014; Choi et al., 2015), which was also supported by results on iPSC-derived fibroblast-like derivates (Choi et al., 2015). Despite the high similarity, comparative analysis of iPSC- and ESCs-derived NSCs revealed 36 genes to deviate at a 2-fold difference (adjusted p-value <0.01). Interestingly, 33 of the differentially transcribed genes were upregulated and only three were downregulated in iPSC-NSCs compared to their ESC-derived counterparts. Though obtained in the somatic state, this quantity is congruent to the range of differentially expressed genes (DEGs) in other isogenic studies on pluripotent cells (Teichroeb et al., 2011; Mallon et al., 2014; Choi et al., 2015) and likewise to Choi and coworkers, gene ontology term analysis of the differentially expressed genes revealed no enriched molecular function, biological process or cellular component (Choi et al., 2015). Furthermore, neither in pluriTest-analysis nor in the plotting of neutrality versus pluripotency differences between iPSCs and ESCs or between their NSC-progeny could be detected. All analyzed samples revealed a joint clustering typical for either their

pluripotent state or characteristically for NSCs between terminally differentiated neurons and pluripotent samples. This ascertains similar pluripotency of iPSCs and ESCs and similar differentiation of all NSC lines and, moreover, no memory effect was detectable.

Transcriptional profiling data of NSCs derived from isogenic iPSCs and ESCs was extended by global DNA methylation analysis of a small isogenic sample set of NSCs and neurons. Hierarchical clustering and principal component analysis revealed a clear segregation of reprogrammed and non-reprogrammed samples. Those differences seem to be even more prominent than those between NSCs and neurons. However, iPSC- and ESC-derived NSCs display a highly similar DNA methylation pattern with a correlation of 96%, which is in line with results obtained in studies on isogenic pluripotent cells. Choi and colleagues found a 95-96% accordance comparing two iPSC lines of two male backgrounds with their genetically matched ESC counterparts (Choi et al., 2015) and Mallon and colleagues even detected 98-99% analyzing three male iPSCs lines (Mallon et al., 2014). Differentially methylated CpGs (DMCGs) between single comparisons of iPSC- and ESC-derived cells overlapped, yet, substantial clone-specific alterations could be detected and a large share of differences seen between NSCs derived of both populations seems to be passed on to neurons differentiated thereof. The localization of DMCGs largely reflects the distribution of certain genomic sites on the array and no bias for transcription start sites could be detected, however CpG islands tend to be rather hypomethylated and non-island sites rather hypermethylated especially in the comparisons of iPSC- and ESC-derived neurons. This is in line with Nishino and colleagues finding that methylation differences (especially hypermethylation) between ESCs and iPSCs are mainly random (Nishino et al., 2011; Nishino and Umezawa, 2016).

5.4 Consequences of reprogramming on genomic imprinting

Several studies indicated that the reprogramming process might disturbs imprinted gene regulation (Pick et al., 2009; Stadtfeld et al., 2010; Teichroeb et al., 2011; Hiura et al., 2013), which could have implications for disease modeling and clinical safety. Genomic imprinting is an epigenetically regulated phenomenon that leads to parent-of-origin-specific monoallelic expression of dosage-sensitive genes (Barlow and Bartolomei, 2014). During early embryonic development, several imprinted genes are involved in tissue specification and there are some indications that increased transcript levels caused by biallelic expression of particular imprinted genes may result in biased differentiation of ESCs towards specific tissues (Prelle

et al., 2000). As a consequence of their imprinting status, certain pluripotent cell lines may display therefore an altered differentiation propensity along specific cell lineages, which may also be able to confound certain readout parameters in disease modeling. Furthermore, altered dosage of several imprinted genes, like MEG3, MEST, IGF2 and H19, has been correlated with human diseases, including cancers (Nakagawa et al., 2001; Kohda et al., 2001; Pedersen, 2002; Ulaner, 2003; Astuti et al., 2005; Lu et al., 2009; Chak et al., 2017). In the mouse system disturbed developmental potential of some iPSC in tetraploid (4n) complementation assays could be correlated with impaired expression or methylation of imprinted genes (Stadtfeld et al., 2010; Liu et al., 2010; Chang et al., 2014). Stadtfeld and colleagues could show transcriptional repression of the maternally expressed genes Gtl2, Rian and Mirg of the imprinted Dlk1-Dio3 gene cluster as well as corresponding hypermethylated regions in mouse iPSCs compared to isogenic ESCs, whereas, Chang and colleagues found hypomethylation of the imprinted gene Zrsr1 in non-4n complementation-competent iPSC lines (Chang et al., 2014). Meta-analysis of genome-wide expression data revealed aberrant levels of imprinted PEG3 and MEG3 (also known as GTL2) in several human iPSC lines (Pick et al., 2009) and Teichroeb and colleagues detected an aberrant suppression of imprinted neuronatin (NNAT) expression, on transcript and protein level, in all six analyzed iPSC lines compared to isogenic human ESCs (Phanstiel et al., 2011; Teichroeb et al., 2011). NNAT is associated with neural development (Wijnholds et al., 1995) and was found to be downregulated in pituitary adenomas (Revill et al., 2009). Furthermore, a study of Pólvora-Brandão and colleagues indicated impaired DNA methylation of the imprinted cluster affected in Prader-Willi or Angelman syndromes (chr15q11-q13). Bisulfite-sequencing revealed hypermethylation of MKRN3 differentially methylated region (DMR) in all analyzed iPSC lines and loss of maternal methylation at the imprinting center in one out of five isogenic iPSC clones which correlated with biallelic expression of SNRPN (Pólvora-Brandão et al., 2018).

Dissecting the differences between iPSC- and ESC-derived NSCs in the presented study revealed that though DLK1 expression was elevated in some lines, no consistent difference in gene expression of the DLK1-DIO3 cluster was detectable between iPSC- and ESC-derived NSCs. However, Dlk1 imprinting was reported to be selectively absent in mouse NSCs and astrocytes during postnatal neurogenesis and that this regulation of specific biallelic expression of Dlk1 is required for normal neural development (Ferron et al., 2011; Oh and

Katsanis, 2011). Therefore, DLK1 might neither be imprinted in human NSCs. Furthermore, no suppression of the imprinted gene NNAT was evident in iPSC-NSCs, yet, there was a putative imprinted gene ZNF215 altered in its expression. Yet, in general, no enrichment for imprinted transcripts amongst the identified DEGs could be detected. Whole genome DNA methylation data revealed that one predominantly affected gene in the comparison of iPSC- and ESC-derived neural progeny was the VENT-like homeobox protein 2 (VENTX) gene, which is predicted to be maternally imprinted (http://www.geneimprint.org) and was extensively hypomethylated in all analyzed iPSC-derived NSCs and neurons. While analyzed ESC-derived NSCs and neurons were almost completely methylated, iPSC-derived cells were broadly demethylated, yet, this change in methylation did not correlate with elevated transcriptional levels. Indeed, normal allele-specific imprinted gene expression may persist in the absence of DNA methylation, as other regulatory epigenetic mechanisms could sustain normal allele-specific transcription. In mouse ESCs, allele-specific histone modifications at the differentially methylated region of the imprinted Kcnq1 domain were reported to be sufficient to maintain monoallelic gene expression in cell lines devoid of DNA methylation (Lewis et al., 2004; Umlauf et al., 2004). Moreover, though no enrichment of differentially expressed imprinted genes could be detected between iPSC- and ESC-derived NSCs, DNA methylation data revealed a slight overrepresentation (p-value = 0.01426) of imprinted genes amongst all genes with at least one corresponding DMCG (deltaBeta>0.2). Taken together, this might, nonetheless, suggest a susceptibility of the reprogramming process to impair imprinted gene regulation.

It is interesting to consider what circumstances might influence the stability of epigenetic imprinting in iPSCs and pluripotent cells in general. Some distinct genetic backgrounds seem to be more prone for impaired imprinting of certain genes, like the human ESC line HES-4 for biallelic expression of MEG3 and parthenogenetic cell lines seem to be specifically vulnerable for loss of imprinted marks (Rugg-Gunn et al., 2007; Vassena et al., 2012).

Furthermore, besides to other factors, like the reprogramming cocktail stoichiometry and cell type of origin (Carey et al., 2011; Xu et al., 2015), that can have an effect on normal imprinted gene regulation of iPSCs, specific culture conditions during establishment and maintenance of pluripotency possess a strong influence on the stability of epigenetic imprinting (Stadtfeld et al., 2012). Addition of serum replacement (SR), which is widely used in human ESC culture, to the media prevents hypermethylation of an intergenic differentially

methylated region (IG-DMR) of the Dlk1-Dio3 cluster and silencing of Meg3 expression in mouse iPSCs, which also could be observed by addition of the SR-component ascorbic acid (also known as vitamin C; Stadtfeld et al., 2012). Apart from reprogramming, also some mouse and human ESCs seem to be unsuccessful in maintaining proper imprinting upon prolonged passaging, resulting in skewed DNA methylation and/or abnormal biallelic expression of imprinted transcripts (Feil et al., 1997; Dean et al., 1998; Rugg-Gunn et al., 2005, 2007; Adewumi et al., 2007). Furthermore, Swanzey and colleagues could show in mouse iPSCs via a Dlk1-imprinting-reporter, that allele-specific Dlk1 expression erodes by prolonged culture and differentiation under elevated (atmospheric) oxygen levels (Swanzey and Stadtfeld, 2016). Therefore, alterations in imprinted gene expression and epigenetic regulation in iPSC-derived NSCs may not only be influenced by the reprogramming process, but also by a prolonged period in the pluripotent state.

5.5 Consequences of reprogramming on X-chromosomal gene expression and methylation

Amongst isogenic approaches it was especially noticeable that the sex of the analyzed cell lines seems to be reflected in the results of the studies with 102-154 and even 0-49 identified differentially expressed genes for female (Teichroeb et al., 2011) and male (Mallon et al., 2014; Choi et al., 2015) iPSCs respectively. Likewise to Teichroeb *et al*., the presented work was based on female cell lines and a bias for X-chromosomal genes, amongst the identified DEGs, was detected. This was true, not only in joint analysis, but also for every single comparison of each line. Separate analysis of autosomal and X-linked genes could show that only the specific consideration of X-chromosomal genes led to a clear segregation in unsupervised principal component analysis (PCA). This indicates that the main differences between iPSC- and ESC-NSCs are due to transcriptional alteration on the X-chromosome and that this seems to be a general effect not only a clone-specific one. Furthermore, a joint upregulation of several X-chromosomal DEGs was detected in the comparison of iPSC and ESC as well as between their NSC progeny, pointing out that altered X-linked transcription was already present in the iPSCs and was presumably just passed on to their NSC derivates. Moreover, upregulation of X-chromosomal genes was locally enriched and several hotspots of DEGs could be identified, including p11.4, p11.23, q13, q22.3, q24, q25/q26.1 and q28. This is in line with findings of Teichroeb and colleagues, which could identify the regions

p11.4 and q24 as hotspots of predominately upregulated genes when comparing iPSCs with isogenic ESCs. Furthermore, also on gene level, similar to their findings, the transcripts of *SRPX, RPGR, MID1IP1, CUL4B* (adjusted p-value <0.01) and *TMEM255A* (also known as *FAM70A;* FC>2, adjusted p-value <0.01) were identified to be upregulated in iPSC-derived NSCs (Teichroeb et al., 2011).

Though, DNA methylation profiles were also very similar between iPSC- and ESC-derived NSCs as well as for neurons derived thereof, separate analysis of autosomal and X-chromosomal CpGs revealed a distinct lower correlation of iPSC-NSCs to ESC-NSCs of 85% in X-chromosomal methylation sites compared to 96,5% for autosomal CpGs, indicating a significant difference between iPSC- and ESC-NSCs in X-linked methylation. These results are consistent with previous findings, showing that the majority of methylation differences between female human pluripotent lines are located on the X chromosome, while only minor X-linked variation is found in male lines (Bock et al., 2011; Mekhoubad et al., 2012; Nazor et al., 2012). Overrepresentation of X-chromosomal CpGs could be detected in NSCs as well as in neurons derived from iPSCs and ESCs, indicating, that these differences supposedly are maintained during differentiation. Interestingly, many X-chromosomal CpGs were either hypermethylated or hypomethylated in iPSC-derived NSCs whereas the ESC-derived NSCs displayed a beta-value of approximately 0.5 in those CpGs. This could indicate a loss of X chromosome inactivation in iPSC-derived NSCs. Likewise to Teichroeb and colleagues, the obtained results point to a putative partial loss of XCI as indicated by regionally-restricted transcriptional up-regulation and loss of DNA methylation of X-chromosomal genes in iPSCs-derived cells (Teichroeb et al., 2011; Nazor et al., 2012). However, while whole genome DNA methylation data revealed restricted regions of significant hyper- and hypomethylation of X-chromosomal CpG sites, solely upregulated transcription of X-chromosomal genes was evident. Hypermethylation of those regions, therefore, might just reflect a further stabilization of an already non-transcribed status of the regarding genes. Furthermore, this might also indicate a general vulnerability of these regions for aberrant methylation. There is some evidence in human iPSCs that random aberrant methylation sites are mainly hypermethylated compared to ESCs (Nishino et al., 2011). This suggests that, although these sites might have been hypomethylated in their parental cells, they become abnormally hypermethylated during reprogramming. Such altered hypermethylation has been described to occur most clearly in early passages and to

decrease with prolonged culture to levels typically found in ESCs (Nishino and Umezawa, 2016). However, the apparent specific differences of X-linked gene expression and methylation point to alterations in dosage compensation of the X chromosome in iPSC-derived NSCs.

5.6 Dosage compensation in iPSC-derived neural stem cells

Dosage compensation or X chromosome inactivation (XCI) is a highly complex epigenetic regulatory process that balances gene expression in mammalian cells of two X chromosomes in the female to match the dosage of one X chromosome of males (Lyon, 1961). This transcriptional silencing of one random X chromosome in female cells is initiated upon differentiation in early development and once established, XCI is stably transmitted to daughter cells and maintained throughout subsequent somatic cell divisions leading to a mosaic in the inactivation of the parental or maternal allele in the organism (Barakat and Gribnau, 2010). The long non-coding RNA *XIST* (inactivated X chromosome-specific transcript or Xi specific transcript) is thought to be the major initiator of XCI. Upon expression, *XIST* RNA coats one X chromosome *in cis* and thereby attracting other effectors of the XCI-machinery, which finally leads to transcriptional silencing and heterochromatin formation (Moindrot and Brockdorff, 2016; da Rocha and Heard, 2017; Balaton et al., 2018). Mouse ESCs do not yet express *Xist* and posses two active X chromosomes (XaXa) while primed mouse epiblast cells and human (primed) ESCs, which resemble mouse epiblast cells also in other criteria, are thought to have already executed XCI leading to one inactivated chromosome (XaXi; Lessing et al., 2013). However, the XCI status seems to be highly variable in human pluripotent cells (Wutz, 2012; Nguyen et al., 2013) and *XIST* expression and H3K27me3 accumulation was proposed to classify the XCI status of the cells (Silva et al., 2008).

Disturbed XCI could have implications for iPSC-derived cells regarding their clinical safety and could confound certain readout parameters in disease modeling. Similar to genetic imprinting, the dosage of X-linked transcripts is critical and elevated levels, caused by biallelic expression, are associated with various human pathologies like intellectual disability, hemophagocytic lymphohistiocytosis and cancers (Pageau et al., 2007; Agrelo and Wutz, 2010; Kang et al., 2015a; Chaligné et al., 2015; Holle et al., 2015; Fieremans et al., 2016; Geens and Chuva De Sousa Lopes, 2017).

In the presented study, quantitative PCR analysis of *XIST* expression revealed comparable levels in iPSCs and ESCs, as well as in NSCs derived thereof. All lines displayed lower expression than found in female primary NSCs irrespective of their somatic or pluripotent state and rather at levels of male cells. Yet, these results were consistent with reports describing *XIST* expression levels to be heterogenous both between different human ESC populations and between individual cells within the same culture (Minkovsky et al., 2012), and it was found that expression levels in I3 ESCs and H9 ESCs are comparable to levels in male ESCs (Adewumi et al., 2007; Mekhoubad et al., 2012). Furthermore, during inactivation of XCI, *XIST* is usually monoallelically expressed on the silenced X chromosome (Mutzel et al., 2019). Taking advantage of the clonal XCI pattern found in the I3 and H9.2 backgrounds, it is possible to decipher between monoallelic and biallelic expression via the analysis of heterozygous SNPs. While H9.2 ESCs represent a sub-clonally expanded line of the widely used H9 ESC line, which was furthermore already described to display a clonal XCI status (Shen et al., 2008; Mitjavila-Garcia et al., 2010), the I3 genetic background also seems to be clonal in its XCI pattern, which is presumably also due to a clonal selection during derivation and/or propagation rather than a biased selection of one X chromosome for inactivation (Dvash et al., 2010). Though *TSIX* (*XIST* antisense RNA) showed monoallelic expression, *XIST* expression was biallelically expressed in the analyzed NSC lines. The human non-coding RNA *TSIX* was described in the literature to be co-expressed with *XIST* and only transcribed from the inactive X chromosome (Migeon et al., 2002) and biallelic expression of *XIST* was determined in pre-inactivation states in human preimplantation embryos in the combination with two active X chromosomes (XaXa; Petropoulos et al., 2016). This might indicate, since *TSIX* is still allele-specifically expressed, a maintenance of dosage compensation, yet, low and biallelic expression of *XIST* in somatic iPSC- and ESC-derived NSCs on the contrary might indicate a loss of XCI or at least of this inactivation mark. Though *XIST* expression and coating of the inactivated X chromosome (Xi) is important to initiate XCI (Pontier and Gribnau, 2011), there is, however, some evidence, that it is not necessary to maintain dosage compensation (Brown and Willard, 1994; Csankovszki et al., 1999) and moreover, reactivation of biallelic X-linked gene expression can also take place while *XIST* is expressed and coats the chromosome. Therefore, it might not be sufficient to classify XCI with *XIST* expression alone (Vallot et al., 2015).

Another hallmark of dosage compensation is the formation of heterochromatin on the

inactivated X chromosome, which can be detected as H3K27me3 foci in immunocytochemistry. However, in comparison to female, primary NSCs both iPSC- and ESC-derived NSCs were negative for this marker. This might indicate an additional loss of this XCI mark, which was already present in the starting NSC population. Indeed, several human female ESC lines were described to display a loss of XCI marks (Enver et al., 2005; Silva et al., 2008; Shen et al., 2008; Dvash et al., 2010). Over time in culture, pluripotent cells can undergo an erosion of X chromosomal dosage compensation (Diaz Perez et al., 2012; Mekhoubad et al., 2012; Vallot et al., 2015). This erosion of XCI (XaXe, class III) was characterized by the loss of *XIST* expression and H3K27me3 foci by Mekhoubad and colleagues. Furthermore, they could show that gradual loss of XCI correlates with transcriptional de-repression of X-linked *HPRT* gene (Mekhoubad et al., 2012). Interestingly, while some groups found that erosion of normal monoallelic expression of X-linked genes in the pluripotent state could not be reestablished upon differentiation (Mekhoubad et al., 2012; Nazor et al., 2012; Patel et al., 2017), there are some indications that this is not true for all genes and some transcript revert to normal allele-specific expression (Vallot et al., 2015).

However, other studies could not directly correlate the loss of *XIST* expression and H3K27me3 foci with loss of XCI (Brown and Willard, 1994; Csankovszki et al., 2001; Silva et al., 2008; Shen et al., 2008; Tchieu et al., 2010). Therefore, though those XCI inactivation marks may be lost, normal monoallelic gene activity might still be maintained. Finally, the ultimate proof of loss of dosage compensation on the X-chromosome is the verification of biallelic gene expression. Comparison of allele-specific transcription and change in overall expression levels of X-linked genes revealed that, though *TMEM255A* showed an higher expression in all iPSC-derived NSCs compared to their ESC-derived counterparts and was shown by Teichroeb and colleagues to be also upregulated in iPSCs (Teichroeb et al., 2011), it was monoallelically expressed in all analyzed NSC lines. Therefore, upregulation of X-chromosomal genes in iPSC-NSC seems not necessarily be due to biallelic expression and loss of XCI. Furthermore, monoallelic expression of *HEPH* was confirmed for two iPSC-NSC lines, yet, one line showed biallelic expression, though it was not elevated in comparison to ESC-NSCs. Despite the limited data on allele-specific expression, due to the lack of informative SNPs of the DEGs in the applied genetic backgrounds, the expression of *HEPH* from both

alleles in this NSC line may rather indicate a gene and clone specific random event of biallelic transcription than a general loss of dosage compensation.

However, a recent study revealed that not every usually inactivated X-chromosomal gene becomes biallelically expressed upon erosion of XCI (Vallot et al., 2015). Vallot and colleagues described *HEPH* and *TMEM255A* to be monoallelic expressed in H9 ESCs irrespective of eroded or stable dosage compensation. Furthermore, all X-linked differentially expressed transcripts that were also analyzed by Vallot and colleagues (p-value of <0.01 *AIFM1*, *CDKL5*, *DLG3*, *MID1IP1*, *PIN4*; p-value of <0.01 and fold-change >2 *SLC9A7*, *TMEM255A*) were found to be monoallelically expressed even in eroded H9 ESCs (Vallot et al., 2015). The allele-specific expression of differentially expressed X-chromosomal genes like *TMEM255A* therefore might still be further stabilized via other epigenetic (XCI-) mechanisms.

Heterochromatin formation on the silenced X-chromosome is achieved via at least two inhibiting histone modifications H3K27me3 and H3K9me3 in human cells (Chadwick and Willard, 2004). In human ESCs H3K27me3 is clearly and specifically enriched on the X chromosome and H3K9me3 mark is predominantly found on the X chromosome and chromosome 19, however, in human somatic cells the degree of enrichment differs (Vallot et al., 2015). While H3K27me3 is still mainly enriched on the Xi in somatic cells, H3K9me3 is also highly represented on several autosomes and the active X chromosome (Chadwick and Willard, 2004; Vallot et al., 2015). Moreover, these two heterochromatin types show a non-random, region-specific and non-overlapping distribution throughout the inactivated X chromosome (Chadwick and Willard, 2004; Vallot et al., 2015). However, also the loss of XCI seems to be regionally restricted in iPSC and ESCs (Teichroeb et al., 2011) and reactivation was described to happen predominantly at H3K27me3 domains (Vallot et al., 2015; Cantone and Fisher, 2017). This may also be reflected by the gene set enrichment of the differentially expressed genes in the presented study (13 of the 36 DEGs; including 4 out of 8 identified X-linked DEGs), which revealed an association with H3K27me3 modification gene set in SK-N-SH cells.

Furthermore, though DNA methylation might stabilize dosage compensation in general, in the applied system altered DNA methylation was found in the same regions as elevated transcription. This may indicate that XCI is nevertheless affected, at least in some extent and/or genes. It is therefore tempting to speculate that X-chromosomal DNA methylation

and/or H3K27me3, as the last remaining inactivation marks of some genomic regions, are particularly susceptible to alterations during the reprogramming process due to its enormous epigenetic rearrangements or are in generally susceptible in the pluripotent state. Findings of Cantone and colleagues support this assumption by showing that treatment of human fibroblasts with the DNA demethylating agent 5-deoxy-azacytidine does not alter allele-specific expression of inactivated genes, however, they could ascertain a set of genes to become susceptible to reactivation upon induction of pluripotency (Cantone et al., 2017).

Still, not every X-chromosomal gene might also be affected of XCI. Genes located in the pseudoautosomal regions (PAR) are, with the exception of two genes, typically not subjected to XCI (Ciccodicola et al., 2000; Blaschke and Rappold, 2006). These regions show sequence identity between the sex chromosomes and reside at the ends of the X and Y chromosomes (PAR1, PAR2). In addition to genes located in these regions, other X-linked genes do not undergo XCI. In human cells about 15 to 20 % of X-linked genes escape inactivation (also known as escapees) in a tissue-specific manner (Cotton et al., 2011; Deng et al., 2014; Peeters et al., 2014; Tukiainen et al., 2017; Balaton et al., 2018). Escapees seem furthermore to differ between different human populations or individuals and some are supposedly associated with mental retardation (Zhang et al., 2013). About half of those XCI escaping genes are consistently expressed from both alleles while the other half varies between individuals or tissues in their inactivation status (Balaton et al., 2015). Since research in mouse cells indicated that escapees supposedly retain an accessible chromatin during XCI, indicated by the absence of repressive H3K27me3 and H3K9me3 histone modifiers, they could be in general more susceptible for reactivation (Brinkman et al., 2006; Valley et al., 2006; Cantone et al., 2017). However, comparing the identified DEGs of the presented study with data of escapees identified in female human ESC, human B lymphocyte cell lines, human-mouse-hybrid cell lines, human fibroblasts and various human tissues or primate samples in general, no enrichment for XCI bypassing genes could be detected (Carrel and Willard, 2005; Park et al., 2010; Cotton et al., 2013, 2015; Zhang et al., 2013; Vallot et al., 2015; Balaton et al., 2015).

Another, yet in the field controversially discussed, proposed part of X-chromosomal dosage compensation includes balancing of monoallelic expression of one X chromosome to the dosage of biallelic expression of autosomal chromosomes via upregulation of the active X chromosome (Ohno, 1967; Nguyen and Disteche, 2006; Barakat and Gribnau, 2010; Deng et

al., 2011; Sangrithi et al., 2017). This is thought to be achieved via transcriptional and post-transcriptional regulation, including enhanced transcription initiation, H4K16 acetylation and elevated RNA half-life, and may also be involved in the detected differences. Alterations in this proposed upregulation of the active X-linked genes may therefore explain the observed differences in expression levels of TMEM255A in iPSC-derived NSCs, while its monoallelic expression seems to be maintained.

However, there are some indications that not reprogramming per se is the reason for loss of XCI but rather the pluripotent state itself. Nazor and colleagues found that XCI is maintained upon reprogramming in early passages of iPSCs but gets subsequently lost with time in culture (Nazor et al., 2012; Briggs et al., 2015). Furthermore, there seems to be a general difference in the susceptibility to XCI instability in pluripotent cells compared to somatic cells. Apart from some cancers, loss of sex chromosome dosage compensation was never described in differentiated cell populations (Vallot et al., 2015; Geens et al., 2016) and even forced deletion of *XIST* expression did not lead to massive gene reactivation in somatic cells (Brown and Willard, 1994; Csankovszki et al., 1999). There are multiple factors that might affect the stability of dosage compensation in pluripotent cells. The pluripotent state might be particularly sensitive in this concern due to its specific culture conditions and rapid cell cycling. Selective pro-survival or growth advantage of cells due to over-expression of distinct X-linked genes (Baker et al., 2007; Vallot et al., 2016) may lead to a selection of cells with impaired dosage compensation and may also be reflected in abnormal XCI in several cancers (Ganesan et al., 2005). Additionally, the specific non-overlapping H3K27me3 and H3K9me3 heterochromatin architecture in pluripotent cells, compared to an overlapping distribution of those inactivating histone marks in somatic cells, might make them less stable in maintaining normal dosage compensation (Vallot et al., 2015). Combined with reports, that the reprogramming/pluripotency factors Oct4, Klf4 and Sox2 have as so-called pioneer factors the ability to bind to heterochromatin but not to H3K9me3 regions (Soufi et al., 2012), this may further accounts for a higher susceptibility for variances in X-chromosomal H3K27me3 domains especially in pluripotent cells. Furthermore, the pluripotency-associated transcription factors Oct4, Sox2, Nanog, Rex1 and Klf4 have been implicated in the regulation of *Xist* and *Tsix* and thereby coordinating the initiation of XCI in mouse ESCs (Navarro et al., 2008, 2010; Donohoe et al., 2009; Nesterova et al., 2011; Gontan et al., 2012; Chen et al., 2016). It is conceivable that constant contradicting regulating signals that hinder

XCI might lead in human pluripotent cells, with already initiated XCI, to an intermediate and immature epigenetic control of dosage compensation. Therefore, an initiated but supposedly not yet completely finished XCI in human pluripotent cells might be particularly vulnerable for alterations.

Finally, standard culture conditions, in general, seem suboptimal for human pluripotent cells to reliably and invariably maintain normal dosage compensation. Some studies indicated that human ESCs XCI tends to be more stable under physiological (5%) than under ambient (about 20%) oxygen levels (Lengner et al., 2010; Vallot et al., 2015). However, Briggs and colleagues described *XIST* expression to be differentially affected by the oxygen level over time, resulting in a higher loss of *XIST* transcription in late passages under physiological oxygen concentrations (Briggs et al., 2015). Furthermore, aberrant XCI was reported even in newly established human ESCs under 5% O_2 (de Oliveira Georges et al., 2014). Therefore, it seems that low oxygen conditions cannot prevent changes in XCI but rather only delay them. Moreover, different compositions of the culture media can influence the XCI status. Addition of sodium butyrate to the media was reported to keep ESCs in a pre-inactivated state with two active X chromosomes (XaXa, class I; Ware et al., 2009; Diaz Perez et al., 2012; Anguera et al., 2012). And also the culture on leukemia inhibitor factor (LIF)-expressing feeder cells was proposed to revert the XaXi (class II) of cells to two XaXa (class I) but rather accelerated erosion of XCI (XaXe, class III; Tomoda et al., 2012; Patel et al., 2017). Recently, naïve culture conditions (Takashima et al., 2014; Theunissen et al., 2014) have been proposed to reactivate the inactivated X chromosome (Xi) of primed ESCs resulting in a pre-inactivated state with two active X chromosomes (XaXa; Theunissen et al., 2016; Vallot et al., 2017) but with monoallelic *XIST* expression in the majority of cells (Sahakyan et al., 2017). Recent studies indicate that ground state XaXa iPSCs can directly be generated under naïve culture conditions (Kilens et al., 2018) and that even aberrant XCI of primed pluripotent cells seems to be reset to XaXa under these conditions (Sahakyan et al., 2017). While there are controversial results about the reestablishment of lost XCI upon differentiation (XaXe and XaXa) under primed conditions (Vallot et al., 2015; Patel et al., 2017) and also primed XaXa pluripotent cell-derived populations can be impaired in functional establishment of dosage compensation (Patel et al., 2017), the first results for XaXa in the naïve state indicate that upon differentiation (via an intermediate primed state) these cells reestablish normal *XIST*-mediated XCI (Sahakyan et al., 2017; Vallot et al., 2017). However, monoallelic *XIST*

expression (Sahakyan et al., 2017), non-random XCI upon differentiation and accumulation of repressive H3K27me3 histone modification on the *XIST*-expressing chromosome differ from the status found in blastocysts (Vallot et al., 2017) and slightly different XaXa characteristics found in distinct studies (Sahakyan et al., 2017; Vallot et al., 2017; Kilens et al., 2018) indicating that further validation and optimization of the culture conditions is necessary. However, previous studies indicated that the usage of indirect methods to extrapolate on the dosage compensation status of the cells, such as X:autosome expression ratios, presence of a XIST-coated Barr body, loss of X-linked DNA methylation, or analysis of only single genes for allele-specific expression could easily lead to misinterpretation and may be responsible for some contradicting results (Cantone et al., 2017). Upcoming high throughput technologies, such as allele-specific RNA-seq, and single cell analysis could help to clarify the XCI state in humans and to further optimize culture conditions (Geens and Chuva De Sousa Lopes, 2017).

To conclude, there are indications that XCI was already in part (low *XIST* expression and loss of H3K27me3 foci) eroded in the starting populations. Reprogramming and further culture in the pluripotent state may have continued this erosion though the limited analysis of allele-specific expression could not verify this.

5.7 Limitations of the applied isogenic stem cell system

Several studies indicated that iPSC might retain an epigenetic memory of their cell of origin (Marchetto et al., 2009; Kim et al., 2010). This was reflected by impaired or biased differentiation towards specific lineages as a consequence of their parental cell and partly maintained donor cell DNA methylation pattern (Ohi et al., 2011; Bar-Nur et al., 2011; Kim et al., 2011b). Since iPSC lines applied in this study were reprogrammed from and differentiated into the very same cell type, it is difficult to determine whether those iPSC-derived cells retain an epigenetic memory of their cell type of origin. However, no disparity in differentiation propensity neither in the generation, nor in the terminal differentiation of iPSC- and ESC-derived NSCs could be detected. Moreover, whether and to what extend residual memory of other lineages, like in particular from fibroblasts, which are widely used for iPSC generation, would be forwarded to iPSC-derived cell populations could therefore not be addressed via the presented system. One way to approach this would be to include

other parental somatic cell types for the iPSC generation in this isogenic system and then analyze the resulting iPSC-derived NSCs.

Furthermore, since the focus in this study was the analysis of transcriptional and DNA methylation of somatic iPSC- and ESC-derived NSCs and only a limited number of pluripotent cells was analyzed, it is difficult to distinguish whether alterations are specifically due to the reprogramming process itself or due to additional/prolonged culture especially in the pluripotent state. While analysis of several samples at different time points upon achievement of the pluripotent state (early and late passages) would help to assess this, it is difficult to adjust in general for different duration in culture when comparing iPSC- and ESC-derived NSCs in this isogenic stem cell system.

5.8 Outlook

Though the transcriptional pattern of iPSC- and ESC-NSCs were remarkably similar, the main differences found between those two populations seem to be X-linked. It is hard to decipher, whether and to what extend those altered expression of X-chromosomal genes directly or indirectly impact autosomal gene expression. Therefore, it would be interesting to analyze also different male backgrounds in such a comprehensive and defined isogenic somatic cell approach and, furthermore, ascertain whether the identified X-linked DEGs, which were in part verified as monoallelically expressed, are also altered in their transcription. Furthermore, as low expression levels of *XIST* and loss of H3K27-trimethylation foci in the analyzed ESCs and ESC-derived NSCs indicated an erosion of these marks already before the reprogramming process, it is questionable whether the study of cell lines that still retain these marks would lead to similar results or whether those cells display a distinct level of stability of X-chromosomal transcription levels. Moreover, it would be interesting to investigate differences in histone modifications in iPSC- and ESC-derived populations and thereby determine whether these correlate with the identified DEGs.

Since not only alterations in X-linked gene expression but also imprinted gene transcription is supposedly due to certain suboptimal culture conditions during reprogramming and/or in the pluripotent state, it would be additionally of tremendous importance, for the applicability of both, female and male, ESC and iPSC and their progeny in research and therapy, to find conditions that prevent such aberrations. First steps to resolve this issue have already been attempted with physiological oxygen levels in culture and naïve culture

conditions of the pluripotent cells. However, the understanding of the mechanisms behind the susceptibility for epigenetic alterations in the pluripotent state still stays largely elusive and reports of the absence of some features of the X chromosome in the human embryo (Sahakyan et al., 2017) under those newly developed adjusted culture conditions might indicate that further research in this area and further optimization might still be necessary. Yet, it would be important to analyze the presented isogenic systems under naïve and hypoxic conditions, to clearly decipher the effect of reprogramming on somatic daughter cells and whether no other obstacles are connected with those culture conditions.

Since culture conditions, especially in the pluripotent state may not yet to be optimal, alternative approaches bypassing the pluripotent state like the direct conversion in the designated cell type might also be an option for research and therapy. Like in reprogramming to pluripotency, cells can be directly converted or transdifferentiated via chemical compounds and overexpression of lineage-specific transcription factors. Lineage conversion was described for example from fibroblasts to induced neurons (iN; Vierbuchen et al., 2010; Pang et al., 2011; Li et al., 2015), NSCs (Ring et al., 2012) or cardiomyocytes (Fu et al., 2015; Cao et al., 2016), from B cells and T cells to macrophages (Xie et al., 2004; Laiosa et al., 2006) and from pancreatic exocrine cells to islet beta-cells (Zhou et al., 2008). Furthermore, directly converted cells may be especially advantageous for modeling of late onset-diseases, since they have been described to retain certain age-specific pattern of their cells of origin and are not rejuvenated like reprogrammed iPSCs (Suhr et al., 2010; Nishimura et al., 2013; Frobel et al., 2014; Mertens et al., 2015). However, reports of epigenetic memory from donor tissue (Marro et al., 2011; Noguchi et al., 2018) and different conversion states in dependence of the used genetic background (Kim et al., 2016) indicate that successful lineage conversion has to be tightly evaluated and comprehensive studies are needed to clarify, whether cells generated by direct conversion do not posses other unwanted features that may impact their functionality and safety. Therefore, also when bypassing the pluripotent state, a considerate monitoring and quality control and further studies are needed for a trustworthy application in research and therapy.

Additionally, in case of known single locus disease-mutations, genome editing could not only be combined with the iPSC-technology but also with direct conversion or primary cell systems for applications in autologous therapy or to generate diseased or healthy isogenic controls in disease modeling and therefore overcome variability between genetic

backgrounds and individual clones. Mutations or corrections of specific sites can easily be introduced with recently developed genome engineering techniques, such as zinc finger nucleases (Urnov et al., 2005; Gupta et al., 2012), transcription activator-like effector nucleases (TALENs; Christian et al., 2010; Miller et al., 2011; Mussolino et al., 2011) and, finally, clustered regularly interspaced short palindromic repeats (CRISPRs)/Cas9 (Jinek et al., 2012; Gasiunas et al., 2012; Cong et al., 2013; Mali et al., 2013).

Despite its caveats, direct reprogramming to pluripotency represents a powerful tool for cell replacement therapies. With comprehensive quality control procedures and expanding knowledge on their trustability and safety, iPSC-derived populations make their way to the clinic. After several trials using ESC-derived cell populations like oligodendrocytes for spinal cord injury (Takahashi and Yamanaka, 2016), insulin-producing cells for type 1 diabetes (Takahashi and Yamanaka, 2016; Kondo et al., 2018) and neural precursor cells for Parkinson's disease (Barker et al., 2017) the first iPSC-based trials are planned or have already started (Barker et al., 2017; Guhr et al., 2018; Attwood and Edel, 2019; Bragança et al., 2019; International Clinical Trials Registry Platform, ClinicalTrials.gov; Japanese Clinical Trials Registry UMIN-CTR, umin.ac.jp/ctr/index.htm) like for example autologous and allogenic iPSC-derived retinal epithelial cells (RPE) for patients with wet age-related macular degeneration (UMIN000011929; UMIN000026003; Takahashi and Yamanaka, 2016; Cyranoski, 2014; Mandai et al., 2017; Cyranoski, 2017), allogenic mesenchymal stem cells for graft versus host disease in bone marrow transplant patients (NCT02923375; Rasko et al., 2019), allogenic cardiomyocytes for patients with severe ischemic cardiomyopathy (UMIN000032989; Cyranoski, 2018a; NTC03763136; Mallapaty, 2020) and allogenic dopaminergic precursor cells for Parkinson's disease patients (UMIN000033564; Takahashi, 2017; Cyranoski, 2018b). Those and future trials as well as continuing basic research focused on the molecular mechanisms and processes of cellular reprogramming will further asses the safety and viability of iPSC-derived cell populations.

5.9 General Conclusion

iPSCs and iPSC-derived cell populations hold promising prospects for disease modeling, drug discovery and personalized regenerative cell therapy. Several comparative studies could already demonstrate that human iPSC highly resemble ESCs in their transcriptional and epigenetic signatures, however, less is known whether this equally applies to somatic

progeny of these pluripotent cells. Using a human isogenic stem cell system the presented study shows that also iPSC- and ESC-derived NSCs are remarkably similar in their transcriptome and methylome. The yet identified transcriptional differences between reprogrammed and non-reprogrammed NSCs were minor compared to variations due to distinct genetic backgrounds. This suggests that previously described differences between iPSCs and ESCs were presumably reasoned by distinct genetic backgrounds and seems especially considerable when only small numbers of lines are analyzed which are derived from a single individual as well as analyzing opposite sex of iPSC and ESC lines. This once more underlines the relevance of isogenic systems for disease modeling and drug discovery, where even minor variations could be meaningful but yet often iPSCs or iPSC-derived cells of healthy and diseased donors are analyzed.

Nonetheless, differences in transcriptional and methylation pattern were detected between isogenic iPSC- and ESC-NSCs, which displayed a substantial bias of X-linked genes, indicating an impairment of dosage compensation. Alterations in X chromosome inactivation seem not to be a specific feature of iPSCs but a common aspect of female pluripotent cells under conventional culture conditions. This suggests that presumably the additional sojourn in the pluripotent state itself impacts dosage compensation in iPSC and iPSC-derived NSCs. The unique and complex biology of the X chromosome has often led to its disregard or even exclusion of female backgrounds in genome-wide research studies, yet it constitutes a considerable amount of the genome and is a relevant contributing factor to disease, prevalently in a sex-specific manner (Selmi et al., 2012; Balaton et al., 2018; Schurz et al., 2019). Therefore, female cell lines are of enormous importance not only in drug development and cell replacement therapies for female individuals but also in disease modeling by especially representing a valuable and indispensable tool for a deeper understanding of pathological mechanisms (Skare et al., 2017; Carrette et al., 2018).

Furthermore, also a slight enrichment for altered DNA methylation of CpGs located in imprinted genes was detectable between iPSC- and ESC-derived NSCs. Therefore, when using those cells in research and therapy, a comprehensive and tight epigenetic monitoring of iPSCs and iPSC-derived populations should be executed, in general and in particular of imprinted genes. Moreover, while more detailed knowledge remains to be gained on whether naïve pluripotent culture conditions can prevent transcriptional and epigenetic aberrations during the reprogramming process, a considered experimental design and

choice of readout parameters should be applied and also other disease modeling strategies should be considered, especially in the analysis of X-linked and imprinted diseases.

6 Abbreviations

18S	18S ribosomal RNA
2i	dual inhibition of MEK and GSK3
4n	tetraploid
A	ampere
AA	ascorbic acid
AFP	alpha feto-protein
Amp	ampicillin
AP	alkaline phosphatase
BAF	B allele frequency
BMP	Bone morphogenic protein
bp	base pairs
BSA	bovine serum albumin
C	Celsius
CA	constitutive active
cAMP	cyclic adenosine monophosphate
Cas9	CRISPR-associated protein 9
cDNA	complementary DNA
cm	centimeter
c-Myc	myelocytomatosis oncogene
CNS	central nervous system
CNV	copy number variation
CRE	cyclization recombinase
CRISPR	clustered regularly interspaced short palindromic repeat
C-terminal	carboxyl terminal
DACH1	dachshound homolog 1
DAPI	4´,6-Diamidino-2-phenylindile
DEG	differentially expressed gene
DLK1	delta like non-canonical Notch ligand 1
DMCG	differentially methylated CpG site
DMEM	Dulbecco's modified eagle medium
DMR	differentially methylated region
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTPs	deoxyribonucleotides
DNMT3b	DNA methyltransferase 3 b
dNTPs	deoxynucleoside triphosphates
Dox	doxycycline
EB	embroid body
E. coli	*Escherichia coli*
EDTA	ethylendiamintetraacetate
e.g.	exempli gratia (for example)
EGF	epidermal growth factor
EMT	epithelial-to-mesenchymal transition
EpiSCs	epiblast stem cells
ERK	extracellular signal-regulated kinase
ESC	embryonic stem cell
et al.	et alia (and others)
FACS	fluorescence activated cell sorting
FCS	fetal calf serum
FGF2, (bFGF)	fibroblast growth factor 2, (basic FGF)
x g	times gravity
g	gram
GABA	gamma-aminobutyric acid
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GFAP	glial fibrillary acidic protein
GFP	green fluorescent protein
GSK3	glycogen synthase kinase 3
GT	geltrex
Gy	gray
h	hour
h	human
H2O	water
H3K27me3	histone 3 lysin 27 trimethylation
H3K4me3	histone 3 lysin 4 trimethylation
H3K9me3	histone 3 lysin 9 trimethylation
H4K20	histone 4 lysin 20
HBSS	Hanks balanced salt solution
HCl	hydrochloric acid
HDAC	histone deacetylase
HEK	human embryonic kidney
HEPES	4-(2-Hydroxyethyl)-1-piperazinyl]-ethanesulfonic acid
HEPH	hephaestin
HP1	heterochromatin protein 1
ICM	inner cell mass
IgG	immunoglobulin isotype G
IgM	immunoglobulin isotype M
iPSC	induced pluripotent stem cell
k	Kilo-
Kb	kilo bases
kDa	kilo Dalton
KLF4	krüppel-like factor 4
KOSR	knock out serum replacement
l	liter
LB	lysogeny broth
LIF	leukemia inhibitory factor
Ln	laminin
LPS	lipopolysaccharide
LRR	log R ratio
lt-NES	long-term neuroepithelial stem
µ	micro-
m	milli-
m	mouse
M	molar
MAP2ab	microtubuli-associated protein 2 a and b
MAPK	mitogen-activated protein kinase
Mb	megabase
MEF	mouse embryonic fibroblast
MEK	MAPK/ERK kinase
MET	mesenchymal-to-epithelial transition
MG	matrigel
mg	milligram
min	minute(s)

miRNA	microRNA
ml	milliliter
mRNA	messenger RNA
NaCl	sodium cloride
NaOH	sodium hydroxide
NEAA	non-essential amino acids
NEP	neural epithelial progenitor cell
NES	nestin
ng	nano gramm
nm	nanometer
NP	nucleocapsid protein
NSC	neural stem cells
nt	nucleotide
NTC	no template control
OCT4	octamer-binding transcription factor 4
OKSM	OCT4, KLF4, SOX2 and c-Myc
OSK	OCT4, SOX2 and KLF4
P	Pico-
p	passage
p	plasmid
PAR	pseudoautosomal region
PAX6	paired box gene 6
PBS	phosphate buffered saline
PCA, (PC)	principle component analysis, (principal component)
PCR	polymerase chain reaction
PEG	Polyethylene glycol
Pen	penicillin
PFA	paraformaldehyde
PGC	Primordial germ cell
PLZF	promyelocytic leukemia-associated zinc finger
pmol	pico mol
PNPLA4	patatin like phospholipase domain containing 4
PNS	peripheral nervous system
PO	Poly-L-ornithine
qPCR	quantitative PCR
RA	retinoic acid
rcf	relative centrifugal force
RefSeq	reference sequence
RG	radial glia
RNA	ribonucleic acid
RNase	ribonuclease
R-NSC	rosette neural stem cell
ROCK	Rho-associated protein kinase
rpm	revolutions per minute
rRNA	ribosomal RNA
RT	room temperature
RT-PCR	reverse transcription PCR
rtTA	reverse tetracycline-controlled transactivator
SCID	severe combined immunodeficiency
SCNT	somatic cell nuclear transfer
sec	seconds
SHH	sonic hedgehog
SK-N-SH	Sloan Kettering neuroblastoma cell line
SMA	smooth muscle actin
SMAD	small mother against decapentaplegic homolog
smNPC	small molecule neural precursor cell
SNP	single nucleotide polymorphism
SOX2	SRY-related HMG box gene 2
SRY	sex determining region Y
Strep	streptomycin
TALEN	transcription activator-like effector nucleases
TE	trypsin/EDTA
tetON	tetracycline regulatable gene induction system
TF	transcription factor
TGFβ	transforming growth factor beta
TH	tyrosine hydroxylase
TI	trypsine inhibitor
TMEM255A	transmembrane protein 255A; also known as FAM70A
TRA	tumor-related antigen
TRE	tetracycline responsive element
Tris	tris(hydroxyl methyl)aminomethane
TSIX	XIST antisense RNA
TUBB3	β-III-tubulin; also known as tubulin beta 3 class III
U	unit
UMIN	university medical information network
UV	ultra violet
V	volt
VENTX	VENT homeobox
VPA	valproic acid
v/v	volume per volume
WNT	Wingless/Integrated
WPRE	woodchuck posttranscriptional regulatory element
w/v	weight per volume
Xa	active X chromosome
XCI	X chromosome inactivation
Xe	eroded Xi state
Xi	inactive X chromosome
XIST	X inactive specific transcript
ZNF215	zinc finger protein 215
ZO1	zona occludens 1

7 Abstract

Human induced pluripotent stem cells (iPSCs) could provide a valuable tool for the production of specific somatic cell types for disease modeling or cell replacement. Considering the molecular complexity of the reprogramming process and the requirement of major epigenetic rearrangements for the generation of induced pluripotent stem cells, it is critical whether and to what extent reprogramming-associated alterations can confound readout parameters in disease modeling or influence clinical safety of human iPSC-derived cell populations. Several comparative analyses of iPSCs and human embryonic stem cells (ESCs) already tried to unravel this question. In the majority of studies, however, variances due to distinct genetic backgrounds challenged assertion of their equivalence, indicating the need for genetically matched cells in order to discover even minor differences. Moreover, most approaches focused on pluripotent populations and only little is known about the effects on iPSC-derived somatic cell types. Since some reprogramming-associated alterations may not be detectable in the pluripotent state but, nevertheless, could become relevant in somatic cell populations used in research and therapy, the present study aimed at addressing to what extent somatic stem cells, passed through reprogramming and subsequent differentiation into their original fate, maintain their native transcriptional and methylation signatures. To that end, highly standardized and well-characterized human ESC-derived neural stem cells (ESC-NSCs) were reprogrammed into iPSCs, which were subsequently re-differentiated into NSCs (iPSC-NSCs). Global transcription and DNA methylation profiling of this isogenic system revealed a remarkably similar transcriptome and methylome of both NSC populations with only minor differences. Among these yet identified alterations, there was a disproportionately large fraction of X-chromosomal genes and methylation sites, which regionally coincided with each other. While this data point to a extensive overall reinstallation of transcriptomic and methylation signatures upon sojourn through pluripotency, they also indicate that X-chromosomal genes may escape this reinstallation process and thus corrupt downstream applications such as iPSC-based disease modeling and regenerative approaches. Therefore, these results strongly recommend comprehensive and tight epigenetic monitoring of iPSCs and iPSC-derived populations in their biomedical application as well as considerate experimental design and choice of readout parameters in disease modeling and drug discovery approaches.

8 Zusammenfassung

Humane induzierte pluripotente Stammzellen (iPS) können ein wertvolles Instrument zur Gewinnung spezifischer somatischer Zelltypen für Krankheitsmodelle und Zellersatztherapien darstellen. In Anbetracht der molekularen Komplexität des Reprogrammierungsprozesses und der Notwendigkeit umfassender epigenetischer Neuanordnungen zur Generierung von induzierten pluripotenten Stammzellen, ist es fraglich, ob und in welchem Umfang reprogrammierungsassoziierte Veränderungen Bewertungsparameter in Krankheitsmodellen verfälschen beziehungsweise die klinische Sicherheit von humanen iPS-abgeleiteten Zellpopulationen beeinträchtigen. Einige vergleichende Analysen von iPS-Zellen und humanen embryonalen Stammzellen (ES) gingen bereits dieser Frage nach. In den meisten Studien erschwerten jedoch Abweichungen durch unterschiedliche genetische Hintergründe eine Aussage über deren Gleichwertigkeit, was wiederum auf die Notwendigkeit genetisch abgeglichener Zellen verweist, um auch geringfügige Veränderungen feststellen zu können. Darüber hinaus fokussierten sich die meisten Untersuchungen auf pluripotente Zellpopulationen und nur wenig ist über die Effekte auf iPS-abgeleitete, somatische Zelltypen bekannt. Da manche reprogrammierungs-assoziierte Veränderungen möglicherweise nicht im pluripotenten Status zu ermitteln sind, dennoch aber in für in Forschung und Therapie angewandten, somatischen Zellpopulationen relevant werden könnten, ist es Ziel der vorgelegten Studie herauszustellen, in welchem Umfang somatische Zellen, die eine Reprogrammierung und anschließende Differenzierung in ihren Ausgangszustand durchlaufen haben, ihre ursprünglichen Transkriptions- und Methylierungssignaturen beibehalten. Hierzu wurden hoch standardisierte und charakterisierte humane, ES-abgeleitete, neurale Stammzellen (ES-NSCs) in iPS-Zellen reprogrammiert, welche anschließend wieder zurück in neurale Stammzellen (iPS-NSCs) differenziert wurden. Globale Transkriptions- und DNA-Methylierungsanalysen dieses isogenen Systems zeigten ein bemerkenswert ähnliches Transkriptom und Methylom beider neuraler Stammzellpopulationen mit lediglich geringfügigen Unterschieden. Unter den dennoch identifizierten Veränderungen befand sich ein überproportionaler Anteil X-chromosomaler Gene und Methylierungsstellen, welche regional miteinander übereinstimmten. Während diese Daten auf eine umfassende Reinstallation von Transkriptions- und Methylierungssignaturen nach vorübergehendem Pluripotenzaufenthalt hinweisen, deuten sie auch darauf hin, dass X-chromosomale Gene diesem Prozess entgehen

könnten und somit Anwendungen, wie nachgelagerte iPSC-basierte Krankheitsmodelle und regenerative Methoden, beeinträchtigen. Die gewonnenen Daten legen daher sowohl eine umfangreiche und engmaschige epigenetische Überwachung von iPS-Zellen und hiervon abgeleiteten Populationen bei deren biomedizinischer Anwendung nahe als auch eine umsichtige Experimentgestaltung und Wahl der Bewertungskriterien in Bezug auf Krankheitsmodellierung und Arzneimittelentwicklung.

9 References

Aasen, T., Raya, A., Barrero, M.J., Garreta, E., Consiglio, A., Gonzalez, F., et al. (2008) Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nature Biotechnology*, 26(11), 1276–1284.

Abyzov, A., Mariani, J., Palejev, D., Zhang, Y., Haney, M.S., Tomasini, L., et al. (2012) Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature*, 492(7429), 438–442.

Adewumi, O., Aflatoonian, B., Ahrlund-Richter, L., Amit, M., Andrews, P.W., Beighton, G., et al. (2007) Characterization of human embryonic stem cell lines by the International Stem Cell Initiative. *Nature Biotechnology*, 25(7), 803–816.

Adhikary, S. and Eilers, M. (2005) Transcriptional regulation and transformation by Myc proteins. *Nature Reviews. Molecular Cell Biology*, 6(8), 635–645.

Agrelo, R. and Wutz, A. (2010) ConteXt of change - X inactivation and disease. *EMBO Molecular Medicine*, 2(1), 6–15.

Ali, R.Q., Blomberg, E., Falk, A., Ährlund-Richter, L. and Ulfendahl, M. (2016) Induction of sensory neurons from neuroepithelial stem cells by the ISX9 small molecule. *American Journal of Stem Cells*, 5(1), 19–28.

Amit, M., Carpenter, M.K., Inokuma, M.S., Chiu, C.P., Harris, C.P., Waknitz, M.A., et al. (2000) Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. *Developmental Biology*, 227(2), 271–278.

Anguera, M.C., Sadreyev, R., Zhang, Z., Szanto, A., Payer, B., Sheridan, S.D., et al. (2012) Molecular signatures of human induced pluripotent stem cells highlight sex differences and cancer genes. *Cell Stem Cell*, 11(1), 75–90.

Anokye-Danso, F., Trivedi, C.M., Juhr, D., Gupta, M., Cui, Z., Tian, Y., et al. (2011) Highly efficient miRNA-mediated reprogramming of mouse and human somatic cells to pluripotency. *Cell Stem Cell*, 8(4), 376–388.

Aoi, T., Yae, K., Nakagawa, M., Ichisaka, T., Okita, K., Takahashi, K., et al. (2008) Generation of pluripotent stem cells from adult mouse liver and stomach cells. *Science*, 321(5889), 699–702.

Astuti, D., Latif, F., Wagner, K., Gentle, D., Cooper, W.N., Catchpoole, D., et al. (2005) Epigenetic alteration at the DLK1-GTL2 imprinted domain in human neoplasia: analysis of neuroblastoma, phaeochromocytoma and Wilms' tumour. *British Journal of Cancer*, 92(8), 1574–1580.

Attwood, S. and Edel, M. (2019) iPS-cell technology and the problem of genetic instability—can it ever be safe for clinical use? *Journal of Clinical Medicine*, 8(3), 288.

Badcock, G., Pigott, C., Goepel, J. and Andrews, P.W. (1999) The human embryonal carcinoma marker antigen TRA-1-60 is a sialylated keratan sulfate proteoglycan. *Cancer Research*, 59(18), 4715–4719.

Baker, D.E.C., Harrison, N.J., Maltby, E., Smith, K., Moore, H.D., Shaw, P.J., et al. (2007) Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. *Nature Biotechnology*, 25(2), 207–215.

Balaton, B.P., Cotton, A.M. and Brown, C.J. (2015) Derivation of consensus inactivation status for X-linked genes from genome-wide studies. *Biology of Sex Differences*, 6(1), 35.

Balaton, B.P., Dixon-McDougall, T., Peeters, S.B. and Brown, C.J. (2018) The eXceptional nature of the X chromosome. *Human Molecular Genetics*, 27(R2), R242–R249.

Ban, H., Nishishita, N., Fusaki, N., Tabata, T., Saeki, K., Shikamura, M., et al. (2011) Efficient generation of transgene-free human induced pluripotent stem cells (iPSCs) by temperature-sensitive Sendai virus vectors. *Proceedings of the National Academy of Sciences of the United States of America*, 108(34), 14234–14239.

Barakat, T.S. and Gribnau, J. (2010) X chromosome inactivation and embryonic stem cells. *Advances in Experimental Medicine and Biology*, 695, 132–154.

Barker, R.A., Parmar, M., Studer, L. and Takahashi, J. (2017) Human trials of stem cell-derived dopamine neurons for Parkinson's disease: dawn of a new era. *Cell Stem Cell*, 21(5), 569–573.

Barlow, D.P. and Bartolomei, M.S. (2014) Genomic imprinting in mammals. *Cold Spring Harbor Perspectives in Biology*, 6(2), a018382.

Bar-Nur, O., Russ, H.A., Efrat, S. and Benvenisty, N. (2011) Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell*, 9(1), 17–23.

Barquinero, J., Eixarch, H. and Pérez-Melgosa, M. (2004) Retroviral vectors: new applications for an old tool. *Gene Therapy*, 11(S1), S3–S9.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell*, 125(2), 315–326.

Bhutani, K., Nazor, K.L., Williams, R., Tran, H., Dai, H., Džakula, Ž., et al. (2016) Whole-genome mutational burden analysis of three pluripotency induction methods. *Nature Communications*, 7(1), 10536.

Bitzer, M., Armeanu, S., Lauer, U.M. and Neubert, W.J. (2003) Sendai virus vectors as an emerging negative-strand RNA viral vector system. *The Journal of Gene Medicine*, 5(7), 543–553.

Blaschke, R.J. and Rappold, G. (2006) The pseudoautosomal regions, SHOX and disease. *Current Opinion in Genetics & Development*, 16(3), 233–239.

Bock, C., Kiskinis, E., Verstappen, G., Gu, H., Boulting, G., Smith, Z.D., et al. (2011) Reference maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell*, 144(3), 439–452.

Boermans, H.J., Percy, D.H., Stirtzinger, T. and Croy, B.A. (1992) Engraftment of severe combined immune deficient/beige mice with bovine foetal lymphoid tissues. *Veterinary Immunology and Immunopathology*, 34(3–4), 273–289.

Boulting, G.L., Kiskinis, E., Croft, G.F., Amoroso, M.W., Oakley, D.H., Wainger, B.J., et al. (2011) A functionally characterized test set of human induced pluripotent stem cells. *Nature Biotechnology*, 29(3), 279–286.

Bragança, J., Lopes, J.A., Mendes-Silva, L. and Santos, J.M.A. (2019) Induced pluripotent stem cells, a giant leap for mankind therapeutic applications. *World Journal of Stem Cells*, 11(7), 421–430.

Breton, A., Sharma, R., Diaz, A.C., Parham, A.G., Graham, A., Neil, C., et al. (2013) Derivation and characterization of induced pluripotent stem cells from equine fibroblasts. *Stem Cells and Development*, 22(4), 611–621.

Briggs, S.F., Dominguez, A.A., Chavez, S.L. and 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.

Brinkman, A.B., Roelofsen, T., Pennings, S.W.C., Martens, J.H.A., Jenuwein, T. and Stunnenberg, H.G. (2006) Histone modification patterns associated with the human X chromosome. *EMBO Reports*, 7(6), 628–634.

Brown, C.J. and Willard, H.F. (1994) The human X-inactivation centre is not required for maintenance of X-chromosome inactivation. *Nature*, 368(6467), 154–156.

Buganim, Y., Faddah, D.A., Cheng, A.W., Itskovich, E., Markoulaki, S., Ganz, K., et al. (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. *Cell*, 150(6), 1209–1222.

Campbell, K.H., McWhir, J., Ritchie, W.A. and Wilmut, I. (1996) Sheep cloned by nuclear transfer from a cultured cell line. *Nature*, 380(6569), 64–66.

Cantone, I., Dharmalingam, G., Chan, Y.-W., Kohler, A.-C., Lenhard, B., Merkenschlager, M., et al. (2017) Allele-specific analysis of cell fusion-mediated pluripotent reprograming reveals distinct and predictive susceptibilities of human X-linked genes to reactivation. *Genome Biology*, 18(1), 2.

Cantone, I. and Fisher, A.G. (2017) Human X chromosome inactivation and reactivation: implications for cell reprogramming and disease. *Philosophical Transactions of the Royal Society of London. Series B*, 372(1733), 20160358.

Cao, N., Huang, Y., Zheng, J., Spencer, C.I., Zhang, Y., Fu, J.-D., et al. (2016) Conversion of human fibroblasts into functional cardiomyocytes by small molecules. *Science*, 352(6290), 1216–1220.

Carey, B.W., Markoulaki, S., Hanna, J.H., Faddah, D.A., Buganim, Y., Kim, J., et al. (2011) Reprogramming factor stoichiometry influences the epigenetic state and biological properties of induced pluripotent stem cells. *Cell Stem Cell*, 9(6), 588–598.

Carrel, L. and Willard, H.F. (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*, 434(7031), 400–404.

Carrette, L.L.G., Wang, C.-Y., Wei, C., Press, W., Ma, W., Kelleher, R.J., et al. (2018) A mixed modality approach towards Xi reactivation for Rett syndrome and other X-linked disorders. *Proceedings of the National Academy of Sciences of the United States of America*, 115(4), E668–E675.

Cerrato, F., Sparago, A., Verde, G., De Crescenzo, A., Citro, V., Cubellis, M.V., et al. (2008) Different mechanisms cause imprinting defects at the IGF2/H19 locus in Beckwith-Wiedemann syndrome and Wilms' tumour. *Human Molecular Genetics*, 17(10), 1427–1435.

Chadwick, B.P. and Willard, H.F. (2004) Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. *Proceedings of the National Academy of Sciences*, 101(50), 17450–17455.

Chak, W.-P., Lung, R.W.-M., Tong, J.H.-M., Chan, S.Y.-Y., Lun, S.W.-M., Tsao, S.-W., et al. (2017) Downregulation of long non-coding RNA MEG3 in nasopharyngeal carcinoma. *Molecular Carcinogenesis*, 56(3), 1041–1054.

Chaligné, R., Popova, T., Mendoza-Parra, M.-A., Saleem, M.-A.M., Gentien, D., Ban, K., et al. (2015) The inactive X chromosome is epigenetically unstable and transcriptionally labile in breast cancer. *Genome Research*, 25(4), 488–503.

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M. and 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.

Chan, E.M., Ratanasirintrawoot, S., Park, I.-H., Manos, P.D., Loh, Y.-H., Huo, H., et al. (2009) Live cell imaging distinguishes bona fide human iPS cells from partially reprogrammed cells. *Nature Biotechnology*, 27(11), 1033–1037.

Chang, G., Gao, S., Hou, X., Xu, Z., Liu, Y., Kang, L., et al. (2014) High-throughput sequencing reveals the disruption of methylation of imprinted gene in induced pluripotent stem cells. *Cell Research*, 24(3), 293–306.

Chen, J., Liu, H., Liu, J., Qi, J., Wei, B., Yang, J., et al. (2013a) H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs. *Nature Genetics*, 45(1), 34–42.

Chen, G., Schell, J.P., Benitez, J.A., Petropoulos, S., Yilmaz, M., Reinius, B., et al. (2016) Single-cell analyses of X Chromosome inactivation dynamics and pluripotency during differentiation. *Genome Research*, 26(10), 1342–1354.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., et al. (2013b) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*, 14, 128.

Chin, M.H., Mason, M.J., Xie, W., Volinia, S., Singer, M., Peterson, C., et al. (2009) Induced pluripotent stem cells and embryonic stem cells are distinguished by gene expression signatures. *Cell Stem Cell*, 5(1), 111–123.

Chin, M.H., Pellegrini, M., Plath, K. and Lowry, W.E. (2010) Molecular analyses of human induced pluripotent stem cells and embryonic stem cells. *Cell Stem Cell*, 7(2), 263–269.

Choi, J., Lee, S., Mallard, W., Clement, K., Tagliazucchi, G.M., Lim, H., et al. (2015) A comparison of genetically matched cell lines reveals the equivalence of human iPSCs and ESCs. *Nature Biotechnology*, 33(11), 1173–1181.

Chou, B.-K., Mali, P., Huang, X., Ye, Z., Dowey, S.N., Resar, L.M., et al. (2011) Efficient human iPS cell derivation by a non-integrating plasmid from blood cells with unique epigenetic and gene expression signatures. *Cell Research*, 21(3), 518–529.

Christian, M., Cermak, T., Doyle, E.L., Schmidt, C., Zhang, F., Hummel, A., et al. (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, 186(2), 757–761.

Cibelli, J.B., Grant, K.A., Chapman, K.B., Cunniff, K., Worst, T., Green, H.L., et al. (2002) Parthenogenetic stem cells in nonhuman primates. *Science*, 295(5556), 819.

Ciccodicola, A., D'Esposito, M., Esposito, T., Gianfrancesco, F., Migliaccio, C., Miano, M.G., et al. (2000) Differentially regulated and evolved genes in the fully sequenced Xq/Yq pseudoautosomal region. *Human Molecular Genetics*, 9(3), 395–401.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., et al. (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science*, 339(6121), 819–823.

Conti, L. and Cattaneo, E. (2010) Neural stem cell systems: physiological players or in vitro entities? *Nature Reviews Neuroscience*, 11(3), 176–187.

Conti, L., Pollard, S.M., Gorba, T., Reitano, E., Toselli, M., Biella, G., et al. (2005) Niche-independent symmetrical self-renewal of a mammalian tissue stem cell. *PLoS Biology*, 3(9), e283.

Corrales, N.L.L., Mrasek, K., Voigt, M., Liehr, T. and Kosyakova, N. (2012) Copy number variations (CNVs) in human pluripotent cell-derived neuroprogenitors. *Gene*, 506(2), 377–379.

Cotton, A.M., Ge, B., Light, N., Adoue, V., Pastinen, T. and Brown, C.J. (2013) Analysis of expressed SNPs identifies variable extents of expression from the human inactive X chromosome. *Genome Biology*, 14(11), R122.

Cotton, A.M., Lam, L., Affleck, J.G., Wilson, I.M., Peñaherrera, M.S., McFadden, D.E., et al. (2011) Chromosome-wide DNA methylation analysis predicts human tissue-specific X inactivation. *Human Genetics*, 130(2), 187–201.

Cotton, A.M., Price, E.M., Jones, M.J., Balaton, B.P., Kobor, M.S. and Brown, C.J. (2015) Landscape of DNA methylation on the X chromosome reflects CpG density, functional chromatin state and X-chromosome inactivation. *Human Molecular Genetics*, 24(6), 1528–1539.

Cowan, C.A., Atienza, J., Melton, D.A. and Eggan, K. (2005) Nuclear reprogramming of somatic cells after fusion with human embryonic stem cells. *Science*, 309(5739), 1369–1373.

Csankovszki, G., Nagy, A. and Jaenisch, R. (2001) Synergism of Xist RNA, DNA methylation, and histone hypoacetylation in maintaining X chromosome inactivation. *The Journal of Cell Biology*, 153(4), 773–784.

Csankovszki, G., Panning, B., Bates, B., Pehrson, J.R. and Jaenisch, R. (1999) Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. *Nature Genetics*, 22(4), 323–324.

Cyranoski, D. (2014) Japanese woman is first recipient of next-generation stem cells. *Nature*, nature.2014.15915.

Cyranoski, D. (2017) Japanese man is first to receive "reprogrammed" stem cells from another person. *Nature*, nature.2017.21730.

Cyranoski, D. (2018a) 'Reprogrammed' stem cells approved to mend human hearts for the first time. *Nature*, 557(7707), 619–620.

Cyranoski, D. (2018b) 'Reprogrammed' stem cells implanted into patient with Parkinson's disease. *Nature*, d41586-018-07407–9.

Dean, W., Bowden, L., Aitchison, A., Klose, J., Moore, T., Meneses, J.J., et al. (1998) Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: association with aberrant phenotypes. *Development*, 125(12), 2273–2282.

Deng, X., Berletch, J.B., Nguyen, D.K. and Disteche, C.M. (2014) X chromosome regulation: diverse patterns in development, tissues and disease. *Nature Reviews. Genetics*, 15(6), 367–378.

Deng, X., Hiatt, J.B., Nguyen, D.K., Ercan, S., Sturgill, D., Hillier, L.W., et al. (2011) Evidence for compensatory upregulation of expressed X-linked genes in mammals, Caenorhabditis elegans and Drosophila melanogaster. *Nature Genetics*, 43(12), 1179–1185.

Diaz Perez, S.V., Kim, R., Li, Z., Marquez, V.E., Patel, S., Plath, K., et al. (2012) Derivation of new human embryonic stem cell lines reveals rapid epigenetic progression in vitro that can be prevented by chemical modification of chromatin. *Human Molecular Genetics*, 21(4), 751–764.

Doerr, J., Böckenhoff, A., Ewald, B., Ladewig, J., Eckhardt, M., Gieselmann, V., et al. (2015) Arylsulfatase a overexpressing human iPSC-derived neural cells reduce CNS sulfatide storage in a mouse model of metachromatic leukodystrophy. *Molecular Therapy*, 23(9), 1519–1531.

Donohoe, M.E., Silva, S.S., Pinter, S.F., Xu, N. and Lee, J.T. (2009) The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. *Nature*, 460(7251), 128–132.

Draper, J.S., Smith, K., Gokhale, P., Moore, H.D., Maltby, E., Johnson, J., et al. (2004) Recurrent gain of chromosomes 17q and 12 in cultured human embryonic stem cells. *Nature Biotechnology*, 22(1), 53–54.

Duggal, G., Warrier, S., Ghimire, S., Broekaert, D., Van der Jeught, M., Lierman, S., et al. (2015) Alternative routes to induce naïve pluripotency in human embryonic stem cells. *Stem Cells*, 33(9), 2686–2698.

Dvash, T., Lavon, N. and Fan, G. (2010) Variations of X chromosome inactivation occur in early passages of female human embryonic stem cells. *PloS One*, 5(6), e11330.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., et al. (2008) Global transcription in pluripotent embryonic stem cells. *Cell Stem Cell*, 2(5), 437–447.

Elkabetz, Y., Panagiotakos, G., Al Shamy, G., Socci, N.D., Tabar, V. and Studer, L. (2008) Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes & Development*, 22(2), 152–165.

Enver, T., Soneji, S., Joshi, C., Brown, J., Iborra, F., Orntoft, T., et al. (2005) Cellular differentiation hierarchies in normal and culture-adapted human embryonic stem cells. *Human Molecular Genetics*, 14(21), 3129–3140.

Evans, M.J. and Kaufman, M.H. (1981) Establishment in culture of pluripotential cells from mouse embryos. *Nature*, 292(5819), 154–156.

Falk, A., Koch, P., Kesavan, J., Takashima, Y., Ladewig, J., Alexander, M., et al. (2012) Capture of neuroepithelial-like stem cells from pluripotent stem cells provides a versatile system for in vitro production of human neurons. *PloS One*, 7(1), e29597.

Feil, R., Boyano, M.D., Allen, N.D. and Kelsey, G. (1997) Parental chromosome-specific chromatin conformation in the imprinted *U2af1-rs1* gene in the mouse. *Journal of Biological Chemistry*, 272(33), 20893–20900.

Feng, B., Jiang, J., Kraus, P., Ng, J.-H., Heng, J.-C.D., Chan, Y.-S., et al. (2009) Reprogramming of fibroblasts into induced pluripotent stem cells with orphan nuclear receptor Esrrb. *Nature Cell Biology*, 11(2), 197–203.

Feng, Q., Lu, S.-J., Klimanskaya, I., Gomes, I., Kim, D., Chung, Y., et al. (2010) Hemangioblastic derivatives from human induced pluripotent stem cells exhibit limited expansion and early senescence. *Stem Cells*, 28(4), 704–712.

Féraud, O., Valogne, Y., Melkus, M.W., Zhang, Y., Oudrhiri, N., Haddad, R., et al. (2016) Donor dependent variations in hematopoietic differentiation among embryonic and induced pluripotent stem cell lines. *PloS One*, 11(3), e0149291.

Ferron, S.R., Charalambous, M., Radford, E., McEwen, K., Wildner, H., Hind, E., et al. (2011) Postnatal loss of Dlk1 imprinting in stem cells and niche astrocytes regulates neurogenesis. *Nature*, 475(7356), 381–385.

Fieremans, N., Van Esch, H., Holvoet, M., Van Goethem, G., Devriendt, K., Rosello, M., et al. (2016) Identification of intellectual disability genes in female patients with a skewed X-inactivation pattern. *Human Mutation*, 37(8), 804–811.

Frobel, J., Hemeda, H., Lenz, M., Abagnale, G., Joussen, S., Denecke, B., et al. (2014) Epigenetic rejuvenation of mesenchymal stromal cells derived from induced pluripotent stem cells. *Stem Cell Reports*, 3(3), 414–422.

Fu, Y., Huang, C., Xu, X., Gu, H., Ye, Y., Jiang, C., et al. (2015) Direct reprogramming of mouse fibroblasts into cardiomyocytes with chemical cocktails. *Cell Research*, 25(9), 1013–1024.

Fujimoto, Y., Abematsu, M., Falk, A., Tsujimura, K., Sanosaka, T., Juliandi, B., et al. (2012) Treatment of a mouse model of spinal cord injury by transplantation of human induced pluripotent stem cell-derived long-term self-renewing neuroepithelial-like stem cells. *Stem Cells*, 30(6), 1163–1173.

Fusaki, N., Ban, H., NISHIYAMA, A., Saeki, K. and Hasegawa, M. (2009) Efficient induction of transgene-free human pluripotent stem cells using a vector based on Sendai virus, an RNA virus that does not integrate into the host genome. *Proceedings of the Japan Academy, Series B*, 85(8), 348–362.

Gafni, O., Weinberger, L., Mansour, A.A., Manor, Y.S., Chomsky, E., Ben-Yosef, D., et al. (2013) Derivation of novel human ground state naive pluripotent stem cells. *Nature*, 504(7479), 282–286.

Ganesan, S., Richardson, A.L., Wang, Z.C., Iglehart, J.D., Miron, A., Feunteun, J., et al. (2005) Abnormalities of the inactive X chromosome are a common feature of BRCA1 mutant and sporadic basal-like breast cancer. *Cold Spring Harbor Symposia on Quantitative Biology*, 70, 93–97.

Gao, H., Le, Y., Wu, X., Silberstein, L.E., Giese, R.W. and Zhu, Z. (2010) VentX, a novel lymphoid-enhancing factor/T-cell factor-associated transcription repressor, is a putative tumor suppressor. *Cancer Research*, 70(1), 202–211.

Gao, J. and Simon, M. (2005) Identification of a novel keratinocyte retinyl ester hydrolase as a transacylase and lipase. *The Journal of Investigative Dermatology*, 124(6), 1259–1266.

Gao, H., Wu, X., Sun, Y., Zhou, S., Silberstein, L.E. and Zhu, Z. (2012) Suppression of homeobox transcription factor VentX promotes expansion of human hematopoietic stem/multipotent progenitor cells. *The Journal of Biological Chemistry*, 287(35), 29979–29987.

García-Cao, M., O'Sullivan, R., Peters, A.H.F.M., Jenuwein, T. and Blasco, M.A. (2004) Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. *Nature Genetics*, 36(1), 94–99.

Gasiunas, G., Barrangou, R., Horvath, P. and Siksnys, V. (2012) Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proceedings of the National Academy of Sciences*, 109(39), E2579–E2586.

Gaspar-Maia, A., Alajem, A., Meshorer, E. and Ramalho-Santos, M. (2011) Open chromatin in pluripotency and reprogramming. *Nature Reviews. Molecular Cell Biology*, 12(1), 36–47.

Geens, M. and Chuva De Sousa Lopes, S.M. (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.

Geens, M., Seriola, A., Barbé, L., Santalo, J., Veiga, A., Dée, K., et al. (2016) Female human pluripotent stem cells rapidly lose X chromosome inactivation marks and progress to a skewed methylation pattern during culture. *Molecular Human Reproduction*, 22(4), 285–298.

Ghosh, Z., Wilson, K.D., Wu, Y., Hu, S., Quertermous, T. and Wu, J.C. (2010) Persistent donor cell gene expression among human induced pluripotent stem cells contributes to differences with human embryonic stem cells. *PloS One*, 5(2), e8975.

Ghule, P.N., Dominski, Z., Yang, X. -c., Marzluff, W.F., Becker, K.A., Harper, J.W., et al. (2008) Staged assembly of histone gene expression machinery at subnuclear foci in the abbreviated cell cycle of human embryonic stem cells. *Proceedings of the National Academy of Sciences*, 105(44), 16964–16969.

Glaser, T., Brose, C., Franceschini, I., Hamann, K., Smorodchenko, A., Zipp, F., et al. (2007) Neural cell adhesion molecule polysialylation enhances the sensitivity of embryonic stem cell-derived neural precursors to migration guidance cues. *Stem Cells*, 25(12), 3016–3025.

Golipour, A., David, L., Liu, Y., Jayakumaran, G., Hirsch, C.L., Trcka, D., et al. (2012) A late transition in somatic cell reprogramming requires regulators distinct from the pluripotency network. *Cell Stem Cell*, 11(6), 769–782.

Gontan, C., Achame, E.M., Demmers, J., Barakat, T.S., Rentmeester, E., van IJcken, W., et al. (2012) RNF12 initiates X-chromosome inactivation by targeting REX1 for degradation. *Nature*, 485(7398), 386–390.

Gonzalo, S., Jaco, I., Fraga, M.F., Chen, T., Li, E., Esteller, M., et al. (2006) DNA methyltransferases control telomere length and telomere recombination in mammalian cells. *Nature Cell Biology*, 8(4), 416–424.

Gore, A., Li, Z., Fung, H.-L., Young, J.E., Agarwal, S., Antosiewicz-Bourget, J., et al. (2011) Somatic coding mutations in human induced pluripotent stem cells. *Nature*, 470(7336), 63–67.

Gorris, R., Fischer, J., Erwes, K.L., Kesavan, J., Peterson, D.A., Alexander, M., et al. (2015) Pluripotent stem cell-derived radial glia-like cells as stable intermediate for efficient generation of human oligodendrocytes. *Glia*, 63(12), 2152–2167.

Guenther, M.G., Frampton, G.M., Soldner, F., Hockemeyer, D., Mitalipova, M., Jaenisch, R., et al. (2010) Chromatin structure and gene expression programs of human embryonic and induced pluripotent stem cells. *Cell Stem Cell*, 7(2), 249–257.

Guhr, A., Kobold, S., Seltmann, S., Seiler Wulczyn, A.E.M., Kurtz, A. and Löser, P. (2018) Recent Trends in Research with Human Pluripotent Stem Cells: Impact of Research and Use of Cell Lines in Experimental Research and Clinical Trials. *Stem Cell Reports*, 11(2), 485–496.

Guo, G., von Meyenn, F., Santos, F., Chen, Y., Reik, W., Bertone, P., et al. (2016) Naive Pluripotent Stem Cells Derived Directly from Isolated Cells of the Human Inner Cell Mass. *Stem Cell Reports*, 6(4), 437–446.

Guo, G., Yang, J., Nichols, J., Hall, J.S., Eyres, I., Mansfield, W., et al. (2009) Klf4 reverts developmentally programmed restriction of ground state pluripotency. *Development*, 136(7), 1063–1069.

Gupta, A., Christensen, R.G., Rayla, A.L., Lakshmanan, A., Stormo, G.D. and Wolfe, S.A. (2012) An optimized two-finger archive for ZFN-mediated gene targeting. *Nature Methods*, 9(6), 588–590.

Gurdon, J.B. (1962) The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles. *Journal of Embryology and Experimental Morphology*, 10, 622–640.

Haase, A., Olmer, R., Schwanke, K., Wunderlich, S., Merkert, S., Hess, C., et al. (2009) Generation of induced pluripotent stem cells from human cord blood. *Cell Stem Cell*, 5(4), 434–441.

Hall, P.A. and Watt, F.M. (1989) Stem cells: the generation and maintenance of cellular diversity. *Development*, 106(4), 619–633.

Hanna, J., Cheng, A.W., Saha, K., Kim, J., Lengner, C.J., Soldner, F., et al. (2010) Human embryonic stem cells with biological and epigenetic characteristics similar to those of mouse ESCs. *Proceedings of the National Academy of Sciences of the United States of America*, 107(20), 9222–9227.

Hanna, J., Markoulaki, S., Schorderet, P., Carey, B.W., Beard, C., Wernig, M., et al. (2008) Direct reprogramming of terminally differentiated mature B lymphocytes to pluripotency. *Cell*, 133(2), 250–264.

Hanna, J., Saha, K., Pando, B., van Zon, J., Lengner, C.J., Creyghton, M.P., et al. (2009) Direct cell reprogramming is a stochastic process amenable to acceleration. *Nature*, 462(7273), 595–601.

Hansson, J., Rafiee, M.R., Reiland, S., Polo, J.M., Gehring, J., Okawa, S., et al. (2012) Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Reports*, 2(6), 1579–1592.

Hayashi, K. and Surani, M.A. (2009) Self-renewing epiblast stem cells exhibit continual delineation of germ cells with epigenetic reprogramming in vitro. *Development*, 136(21), 3549–3556.

Helling, R.B., Goodman, H.M. and Boyer, H.W. (1974) Analysis of endonuclease R-EcoRI fragments of DNA from lambdoid bacteriophages and other viruses by agarose-gel electrophoresis. *Journal of Virology*, 14(5), 1235–1244.

Henderson, J.K., Draper, J.S., Baillie, H.S., Fishel, S., Thomson, J.A., Moore, H., et al. (2002) Preimplantation human embryos and embryonic stem cells show comparable expression of stage-specific embryonic antigens. *Stem Cells*, 20(4), 329–337.

Hiura, H., Toyoda, M., Okae, H., Sakurai, M., Miyauchi, N., Sato, A., et al. (2013) Stability of genomic imprinting in human induced pluripotent stem cells. *BMC Genetics*, 14(1), 32.

Hochedlinger, K., Yamada, Y., Beard, C. and Jaenisch, R. (2005) Ectopic expression of Oct-4 blocks progenitor-cell differentiation and causes dysplasia in epithelial tissues. *Cell*, 121(3), 465–477.

Hockemeyer, D., Soldner, F., Cook, E.G., Gao, Q., Mitalipova, M. and Jaenisch, R. (2008) A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell*, 3(3), 346–353.

Holle, J.R., Marsh, R.A., Holdcroft, A.M., Davies, S.M., Wang, L., Zhang, K., et al. (2015) Hemophagocytic lymphohistiocytosis in a female patient due to a heterozygous *XIAP* mutation and skewed X chromosome inactivation. *Pediatric Blood & Cancer*, 62(7), 1288–1290.

Hou, P., Li, Y., Zhang, X., Liu, C., Guan, J., Li, H., et al. (2013) Pluripotent stem cells induced from mouse somatic cells by small-molecule compounds. *Science*, 341(6146), 651–654.

Hu, B.-Y., Weick, J.P., Yu, J., Ma, L.-X., Zhang, X.-Q., Thomson, J.A., et al. (2010) Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proceedings of the National Academy of Sciences of the United States of America*, 107(9), 4335–4340.

Huang, K., Shen, Y., Xue, Z., Bibikova, M., April, C., Liu, Z., et al. (2013) A panel of CpG methylation sites distinguishes human embryonic stem cells and induced pluripotent stem cells. *Stem Cell Reports*, 2(1), 36–43.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009a) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, 4(1), 44–57.

Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009b) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, 37(1), 1–13.

Huang, X., Tian, C., Liu, M., Wang, Y., Tolmachev, A.V., Sharma, S., et al. (2012) Quantitative proteomic analysis of mouse embryonic fibroblasts and induced pluripotent stem cells using O16/O18 labeling. *Journal of Proteome Research*, 11(4), 2091–2102.

Huangfu, D., Maehr, R., Guo, W., Eijkelenboom, A., Snitow, M., Chen, A.E., et al. (2008a) Induction of pluripotent stem cells by defined factors is greatly improved by small-molecule compounds. *Nature Biotechnology*, 26(7), 795–797.

Huangfu, D., Osafune, K., Maehr, R., Guo, W., Eijkelenboom, A., Chen, S., et al. (2008b) Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. *Nature Biotechnology*, 26(11), 1269–1275.

Hussein, S.M., Batada, N.N., Vuoristo, S., Ching, R.W., Autio, R., Närvä, E., et al. (2011) Copy number variation and selection during reprogramming to pluripotency. *Nature*, 470(7336), 58–62.

Inoue, M., Tokusumi, Y., Ban, H., Kanaya, T., Tokusumi, T., Nagai, Y., et al. (2003) Nontransmissible virus-like particle formation by F-deficient sendai virus is temperature sensitive and reduced by mutations in M and HN proteins. *Journal of Virology*, 77(5), 3238–3246.

Itskovitz-Eldor, J., Schuldiner, M., Karsenti, D., Eden, A., Yanuka, O., Amit, M., et al. (2000) Differentiation of human embryonic stem cells into embryoid bodies compromising the three embryonic germ layers. *Molecular Medicine*, 6(2), 88–95.

Ji, J., Ng, S.H., Sharma, V., Neculai, D., Hussein, S., Sam, M., et al. (2012) Elevated coding mutation rate during the reprogramming of human somatic cells into induced pluripotent stem cells. *Stem Cells*, 30(3), 435–440.

Jia, F., Wilson, K.D., Sun, N., Gupta, D.M., Huang, M., Li, Z., et al. (2010) A nonviral minicircle vector for deriving human iPS cells. *Nature Methods*, 7(3), 197–199.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A. and Charpentier, E. (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 337(6096), 816–821.

Kaji, K., Norrby, K., Paca, A., Mileikovsky, M., Mohseni, P. and Woltjen, K. (2009) Virus-free induction of pluripotency and subsequent excision of reprogramming factors. *Nature*, 458(7239), 771–775.

Kang, J., Lee, H.J., Kim, J., Lee, J.J. and Maeng, L. (2015a) Dysregulation of X chromosome inactivation in high grade ovarian serous adenocarcinoma. *PloS One*, 10(3), e0118927.

Kang, L., Wang, J., Zhang, Y., Kou, Z. and Gao, S. (2009) iPS cells can support full-term development of tetraploid blastocyst-complemented embryos. *Cell Stem Cell*, 5(2), 135–138.

Kang, X., Yu, Q., Huang, Y., Song, B., Chen, Y., Gao, X., et al. (2015b) Effects of integrating and non-integrating reprogramming methods on copy number variation and genomic stability of human induced pluripotent stem cells. *PloS One*, 10(7), e0131128.

Kilens, S., Meistermann, D., Moreno, D., Chariau, C., Gaignerie, A., Reignier, A., et al. (2018) Parallel derivation of isogenic human primed and naive induced pluripotent stem cells. *Nature Communications*, 9(1), 360.

Kim, K., Doi, A., Wen, B., Ng, K., Zhao, R., Cahan, P., et al. (2010) Epigenetic memory in induced pluripotent stem cells. *Nature*, 467(7313), 285–290.

Kim, J.B., Greber, B., Araúzo-Bravo, M.J., Meyer, J., Park, K.I., Zaehres, H., et al. (2009a) Direct reprogramming of human neural stem cells by OCT4. *Nature*, 461(7264), 649–653.

Kim, D., Kim, C.-H., Moon, J.-I., Chung, Y.-G., Chang, M.-Y., Han, B.-S., et al. (2009b) Generation of human induced pluripotent stem cells by direct delivery of reprogramming proteins. *Cell Stem Cell*, 4(6), 472–476.

Kim, J., Lengner, C.J., Kirak, O., Hanna, J., Cassady, J.P., Lodato, M.A., et al. (2011a) Reprogramming of postnatal neurons into induced pluripotent stem cells by defined factors. *Stem Cells*, 29(6), 992–1000.

Kim, S.M., Lim, K.T., Kwak, T.H., Lee, S.C., Im, J.H., Hali, S., et al. (2016) Induced neural stem cells from distinct genetic backgrounds exhibit different reprogramming status. *Stem Cell Research*, 16(2), 460–468.

Kim, J.B., Sebastiano, V., Wu, G., Araúzo-Bravo, M.J., Sasse, P., Gentile, L., et al. (2009c) Oct4-induced pluripotency in adult neural stem cells. *Cell*, 136(3), 411–419.

Kim, J.B., Zaehres, H., Wu, G., Gentile, L., Ko, K., Sebastiano, V., et al. (2008) Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors. *Nature*, 454(7204), 646–650.

Kim, K., Zhao, R., Doi, A., Ng, K., Unternaehrer, J., Cahan, P., et al. (2011b) Donor cell type can influence the epigenome and differentiation potential of human induced pluripotent stem cells. *Nature Biotechnology*, 29(12), 1117–1119.

Knoepfler, P.S. (2008) Why myc? An unexpected ingredient in the stem cell cocktail. *Cell Stem Cell*, 2(1), 18–21.

Kobayashi, H. and Kikyo, N. (2015) Epigenetic regulation of open chromatin in pluripotent stem cells. *Translational Research: The Journal of Laboratory and Clinical Medicine*, 165(1), 18–27.

Koch, P., Opitz, T., Steinbeck, J.A., Ladewig, J. and Brüstle, O. (2009) A rosette-type, self-renewing human ES cell-derived neural stem cell with potential for in vitro instruction and synaptic integration. *Proceedings of the National Academy of Sciences of the United States of America*, 106(9), 3225–3230.

Koche, R.P., Smith, Z.D., Adli, M., Gu, H., Ku, M., Gnirke, A., et al. (2011) Reprogramming factor expression initiates widespread targeted chromatin remodeling. *Cell Stem Cell*, 8(1), 96–105.

Kohda, M., Hoshiya, H., Katoh, M., Tanaka, I., Masuda, R., Takemura, T., et al. (2001) Frequent loss of imprinting of IGF2 and MEST in lung adenocarcinoma. *Molecular Carcinogenesis*, 31(4), 184–191.

Kondo, Y., Toyoda, T., Inagaki, N. and Osafune, K. (2018) iPSC technology-based regenerative therapy for diabetes. *Journal of Diabetes Investigation*, 9(2), 234–243.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., et al. (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, 44(W1), W90-97.

Laiosa, C.V., Stadtfeld, M., Xie, H., de Andres-Aguayo, L. and Graf, T. (2006) Reprogramming of committed T cell progenitors to macrophages and dendritic cells by C/EBPα and PU.1 transcription factors. *Immunity*, 25(5), 731–744.

Laurent, L.C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., et al. (2011) Dynamic changes in the copy number of pluripotency and cell proliferation genes in human ESCs and iPSCs during reprogramming and time in culture. *Cell Stem Cell*, 8(1), 106–118.

Lefort, N., Feyeux, M., Bas, C., Féraud, O., Bennaceur-Griscelli, A., Tachdjian, G., et al. (2008) Human embryonic stem cells reveal recurrent genomic instability at 20q11.21. *Nature Biotechnology*, 26(12), 1364–1366.

Lengner, C.J., Gimelbrant, A.A., Erwin, J.A., Cheng, A.W., Guenther, M.G., Welstead, G.G., et al. (2010) Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell*, 141(5), 872–883.

Lenz, M., Schuldt, B.M., Müller, F.-J. and Schuppert, A. (2013) PhysioSpace: relating gene expression experiments from heterogeneous sources using shared physiological processes. *PloS One*, 8(10), e77627.

Lessing, D., Anguera, M.C. and Lee, J.T. (2013) X chromosome inactivation and epigenetic responses to cellular reprogramming. *Annual Review of Genomics and Human Genetics*, 14(1), 85–110.

Lewis, A., Mitsuya, K., Umlauf, D., Smith, P., Dean, W., Walter, J., et al. (2004) Imprinting on distal chromosome 7 in the placenta involves repressive histone methylation independent of DNA methylation. *Nature Genetics*, 36(12), 1291–1295.

Li, R., Liang, J., Ni, S., Zhou, T., Qing, X., Li, H., et al. (2010) A mesenchymal-to-epithelial transition initiates and is required for the nuclear reprogramming of mouse fibroblasts. *Cell Stem Cell*, 7(1), 51–63.

Li, W., Sun, W., Zhang, Y., Wei, W., Ambasudhan, R., Xia, P., et al. (2011) Rapid induction and long-term self-renewal of primitive neural precursors from human embryonic stem cells by small molecule inhibitors. *Proceedings of the National Academy of Sciences*, 108(20), 8299–8304.

Li, X., Zuo, X., Jing, J., Ma, Y., Wang, J., Liu, D., et al. (2015) Small-molecule-driven direct reprogramming of mouse fibroblasts into functional neurons. *Cell Stem Cell*, 17(2), 195–203.

Lin, C.Y., Lovén, J., Rahl, P.B., Paranal, R.M., Burge, C.B., Bradner, J.E., et al. (2012) Transcriptional amplification in tumor cells with elevated c-Myc. *Cell*, 151(1), 56–67.

Lin, G., OuYang, Q., Zhou, X., Gu, Y., Yuan, D., Li, W., et al. (2007) A highly homozygous and parthenogenetic human embryonic stem cell line derived from a one-pronuclear oocyte following in vitro fertilization procedure. *Cell Research*, 17(12), 999–1007.

Lister, R., Pelizzola, M., Kida, Y.S., Hawkins, R.D., Nery, J.R., Hon, G., et al. (2011) Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature*, 471(7336), 68–73.

Liu, L., Luo, G.-Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., et al. (2010) Activation of the imprinted *Dlk1-Dio3* region correlates with pluripotency levels of mouse stem cells. *Journal of Biological Chemistry*, 285(25), 19483–19490.

Livak, K.J. and Schmittgen, T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, 25(4), 402–408.

Lo Sardo, V., Ferguson, W., Erikson, G.A., Topol, E.J., Baldwin, K.K. and Torkamani, A. (2017) Influence of donor age on induced pluripotent stem cells. *Nature Biotechnology*, 35(1), 69–74.

Lu, Y., Lu, P., Zhu, Z., Xu, H. and Zhu, X. (2009) Loss of imprinting of insulin-like growth factor 2 is associated with increased risk of lymph node metastasis and gastric corpus cancer. *Journal of Experimental & Clinical Cancer Research*, 28(1), 125.

Lyon, M.F. (1961) Gene action in the X-chromosome of the mouse (mus musculus l.). *Nature*, 190(4773), 372–373.

Maekawa, M., Yamaguchi, K., Nakamura, T., Shibukawa, R., Kodanaka, I., Ichisaka, T., et al. (2011) Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature*, 474(7350), 225–229.

Maherali, N., Ahfeldt, T., Rigamonti, A., Utikal, J., Cowan, C. and Hochedlinger, K. (2008) A high-efficiency system for the generation and study of human induced pluripotent stem cells. *Cell Stem Cell*, 3(3), 340–345.

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., et al. (2007) Directly reprogrammed fibroblasts show global epigenetic remodeling and widespread tissue contribution. *Cell Stem Cell*, 1(1), 55–70.

Mali, P., Chou, B.-K., Yen, J., Ye, Z., Zou, J., Dowey, S., et al. (2010) Butyrate greatly enhances derivation of human induced pluripotent stem cells by promoting epigenetic remodeling and the expression of pluripotency-associated genes. *Stem Cells*, 28(4), 713–720.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., et al. (2013) RNA-guided human genome engineering via Cas9. *Science*, 339(6121), 823–826.

Mallapaty, S. (2020) Revealed: two men in China were first to receive pioneering stem-cell treatment for heart disease. *Nature*, 581(7808), 249–250.

Mallon, B.S., Hamilton, R.S., Kozhich, O.A., Johnson, K.R., Fann, Y.C., Rao, M.S., et al. (2014) Comparison of the molecular profiles of human embryonic and induced pluripotent stem cells of isogenic origin. *Stem Cell Research*, 12(2), 376–386.

Mandai, M., Watanabe, A., Kurimoto, Y., Hirami, Y., Morinaga, C., Daimon, T., et al. (2017) Autologous Induced Stem-Cell–Derived Retinal Cells for Macular Degeneration. *New England Journal of Medicine*, 376(11), 1038–1046.

Mandal, P.K. and Rossi, D.J. (2013) Reprogramming human fibroblasts to pluripotency using modified mRNA. *Nature Protocols*, 8(3), 568–582.

Mandel, M. and Higa, A. (1970) Calcium-dependent bacteriophage DNA infection. *Journal of Molecular Biology*, 53(1), 159–162.

Manoli, D.S., Subramanyam, D., Carey, C., Sudin, E., Van Westerhuyzen, J.A., Bales, K.L., et al. (2012) Generation of induced pluripotent stem cells from the prairie vole. *PloS One*, 7(5), e38119.

Marchetto, M.C.N., Yeo, G.W., Kainohana, O., Marsala, M., Gage, F.H. and Muotri, A.R. (2009) Transcriptional signature and memory retention of human-induced pluripotent stem cells. *PloS One*, 4(9), e7076.

Marei, H.E., Althani, A., Lashen, S., Cenciarelli, C. and Hasan, A. (2017) Genetically unmatched human iPSC and ESC exhibit equivalent gene expression and neuronal differentiation potential. *Scientific Reports*, 7(1), 17504.

Marion, R.M., Strati, K., Li, H., Tejera, A., Schoeftner, S., Ortega, S., et al. (2009) Telomeres acquire embryonic stem cell characteristics in induced pluripotent stem cells. *Cell Stem Cell*, 4(2), 141–154.

Marks, H., Kalkan, T., Menafra, R., Denissov, S., Jones, K., Hofemeister, H., et al. (2012) The transcriptional and epigenomic foundations of ground state pluripotency. *Cell*, 149(3), 590–604.

Marro, S., Pang, Z.P., Yang, N., Tsai, M.-C., Qu, K., Chang, H.Y., et al. (2011) Direct lineage conversion of terminally differentiated hepatocytes to functional neurons. *Cell Stem Cell*, 9(4), 374–382.

Martin, G.R. (1981) Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, 78(12), 7634–7638.

Martins-Taylor, K., Nisler, B.S., Taapken, S.M., Compton, T., Crandall, L., Montgomery, K.D., et al. (2011) Recurrent copy number variations in human induced pluripotent stem cells. *Nature Biotechnology*, 29(6), 488–491.

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.-C., Yakir, B., Clark, A.T., et al. (2010) Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. *Cell Stem Cell*, 7(4), 521–531.

Mekhoubad, S., Bock, C., de Boer, A.S., Kiskinis, E., Meissner, A. and Eggan, K. (2012) Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell*, 10(5), 595–609.

Mertens, J., Paquola, A.C.M., Ku, M., Hatch, E., Böhnke, L., Ladjevardi, S., et al. (2015) Directly reprogrammed human neurons retain aging-associated transcriptomic signatures and reveal age-related nucleocytoplasmic defects. *Cell Stem Cell*, 17(6), 705–718.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T. and Misteli, T. (2006) Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. *Developmental Cell*, 10(1), 105–116.

Migeon, B.R., Lee, C.H., Chowdhury, A.K. and Carpenter, H. (2002) Species differences in TSIX/Tsix reveal the roles of these genes in X-chromosome inactivation. *The American Journal of Human Genetics*, 71(2), 286–293.

Mikkelsen, T.S., Hanna, J., Zhang, X., Ku, M., Wernig, M., Schorderet, P., et al. (2008) Dissecting direct reprogramming through integrative genomic analysis. *Nature*, 454(7200), 49–55.

Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., et al. (2011) A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology*, 29(2), 143–148.

Minkovsky, A., Patel, S. and Plath, K. (2012) Concise review: pluripotency and the transcriptional inactivation of the female mammalian X chromosome. *Stem Cells*, 30(1), 48–54.

Mitjavila-Garcia, M.-T., Bonnet, M.L., Yates, F., Haddad, R., Oudrhiri, N., Feraud, O., et al. (2010) Partial reversal of the methylation pattern of the X-linked gene HUMARA during hematopoietic differentiation of human embryonic stem cells. *Journal of Molecular Cell Biology*, 2(5), 291–298.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., et al. (2003) The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*, 113(5), 631–642.

Miura, K., Okada, Y., Aoi, T., Okada, A., Takahashi, K., Okita, K., et al. (2009) Variation in the safety of induced pluripotent stem cell lines. *Nature Biotechnology*, 27(8), 743–745.

Miyoshi, N., Ishii, H., Nagano, H., Haraguchi, N., Dewi, D.L., Kano, Y., et al. (2011) Reprogramming of mouse and human cells to pluripotency using mature microRNAs. *Cell Stem Cell*, 8(6), 633–638.

Moindrot, B. and Brockdorff, N. (2016) RNA binding proteins implicated in Xist-mediated chromosome silencing. *Seminars in Cell & Developmental Biology*, 56, 58–70.

Montserrat, N., Nivet, E., Sancho-Martinez, I., Hishida, T., Kumar, S., Miquel, L., et al. (2013) Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell*, 13(3), 341–350.

Müller, F.-J., Schuldt, B.M., Williams, R., Mason, D., Altun, G., Papapetrou, E.P., et al. (2011) A bioinformatic assay for pluripotency in human cells. *Nature Methods*, 8(4), 315–317.

Munoz, J., Low, T.Y., Kok, Y.J., Chin, A., Frese, C.K., Ding, V., et al. (2011) The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. *Molecular Systems Biology*, 7(1), 550.

Mussolino, C., Morbitzer, R., Lütge, F., Dannemann, N., Lahaye, T. and Cathomen, T. (2011) A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Research*, 39(21), 9283–9293.

Mutzel, V., Okamoto, I., Dunkel, I., Saitou, M., Giorgetti, L., Heard, E., et al. (2019) A symmetric toggle switch explains the onset of random X inactivation in different mammals. *Nature Structural & Molecular Biology*, 26(5), 350–360.

Nakagawa, H., Chadwick, R.B., Peltomaki, P., Plass, C., Nakamura, Y. and de la Chapelle, A. (2001) Loss of imprinting of the insulin-like growth factor II gene occurs by biallelic methylation in a core region of H19-associated CTCF-binding sites in colorectal cancer. *Proceedings of the National Academy of Sciences*, 98(2), 591–596.

Nakagawa, M., Koyanagi, M., Tanabe, K., Takahashi, K., Ichisaka, T., Aoi, T., et al. (2008) Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts. *Nature Biotechnology*, 26(1), 101–106.

Nakagawa, M., Takizawa, N., Narita, M., Ichisaka, T. and Yamanaka, S. (2010) Promotion of direct reprogramming by transformation-deficient Myc. *Proceedings of the National Academy of Sciences*, 107(32), 14152–14157.

Närvä, E., Autio, R., Rahkonen, N., Kong, L., Harrison, N., Kitsberg, D., et al. (2010) High-resolution DNA analysis of human embryonic stem cell lines reveals culture-induced copy number changes and loss of heterozygosity. *Nature Biotechnology*, 28(4), 371–377.

Nat, R., Nilbratt, M., Narkilahti, S., Winblad, B., Hovatta, O. and Nordberg, A. (2007) Neurogenic neuroepithelial and radial glial cells generated from six human embryonic stem cell lines in serum-free suspension and adherent cultures. *Glia*, 55(4), 385–399.

Navarro, P., Chambers, I., Karwacki-Neisius, V., Chureau, C., Morey, C., Rougeulle, C., et al. (2008) Molecular coupling of Xist regulation and pluripotency. *Science*, 321(5896), 1693–1695.

Navarro, P., Oldfield, A., Legoupi, J., Festuccia, N., Dubois, A., Attia, M., et al. (2010) Molecular coupling of Tsix regulation and pluripotency. *Nature*, 468(7322), 457–460.

Nazor, K.L., Altun, G., Lynch, C., Tran, H., Harness, J.V., Slavin, I., et al. (2012) Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. *Cell Stem Cell*, 10(5), 620–634.

Nesterova, T.B., Senner, C.E., Schneider, J., Alcayna-Stevens, T., Tattermusch, A., Hemberger, M., et al. (2011) Pluripotency factor binding and Tsix expression act synergistically to repress Xist in undifferentiated embryonic stem cells. *Epigenetics & Chromatin*, 4(1), 17.

Newman, A.M. and Cooper, J.B. (2010) Lab-specific gene expression signatures in pluripotent stem cells. *Cell Stem Cell*, 7(2), 258–262.

Ng, H.-H. and Surani, M.A. (2011) The transcriptional and signalling networks of pluripotency. *Nature Cell Biology*, 13(5), 490–496.

Nguyen, D.K. and Disteche, C.M. (2006) Dosage compensation of the active X chromosome in mammals. *Nature Genetics*, 38(1), 47–53.

Nguyen, H.T., Geens, M. and Spits, C. (2013) Genetic and epigenetic instability in human pluripotent stem cells. *Human Reproduction Update*, 19(2), 187–205.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., et al. (1998) Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*, 95(3), 379–391.

Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., et al. (2012) c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell*, 151(1), 68–79.

Nishimura, T., Kaneko, S., Kawana-Tachikawa, A., Tajima, Y., Goto, H., Zhu, D., et al. (2013) Generation of rejuvenated antigen-specific T cells by reprogramming to pluripotency and redifferentiation. *Cell Stem Cell*, 12(1), 114–126.

Nishimura, K., Ohtaka, M., Takada, H., Kurisaki, A., Tran, N.V.K., Tran, Y.T.H., et al. (2017) Simple and effective generation of transgene-free induced pluripotent stem cells using an auto-erasable Sendai virus vector responding to microRNA-302. *Stem Cell Research*, 23, 13–19.

Nishimura, K., Sano, M., Ohtaka, M., Furuta, B., Umemura, Y., Nakajima, Y., et al. (2011) Development of defective and persistent Sendai virus vector: a unique gene delivery/expression system ideal for cell reprogramming. *Journal of Biological Chemistry*, 286(6), 4760–4771.

Nishino, K., Toyoda, M., Yamazaki-Inoue, M., Fukawatase, Y., Chikazawa, E., Sakaguchi, H., et al. (2011) DNA methylation dynamics in human induced pluripotent stem cells over time. *PLoS Genetics*, 7(5), e1002085.

Nishino, K. and Umezawa, A. (2016) DNA methylation dynamics in human induced pluripotent stem cells. *Human Cell*, 29(3), 97–100.

Nishizawa, M., Chonabayashi, K., Nomura, M., Tanaka, A., Nakamura, M., Inagaki, A., et al. (2016) Epigenetic variation between human induced pluripotent stem cell lines is an indicator of differentiation capacity. *Cell Stem Cell*, 19(3), 341–354.

Noguchi, H., Miyagi-Shiohira, C. and Nakashima, Y. (2018) Induced tissue-specific stem cells and epigenetic memory in induced pluripotent stem cells. *International Journal of Molecular Sciences*, 19(4).

Oh, E.C. and Katsanis, N. (2011) Imprinting in the brain. *Nature*, 475(7356), 299–300.

Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J.M., Guo, T., et al. (2011) Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature Cell Biology*, 13(5), 1–11.

Ohno, S. (1967) *Sex chromosomes and sex-linked genes*.

Okita, K., Ichisaka, T. and Yamanaka, S. (2007) Generation of germline-competent induced pluripotent stem cells. *Nature*, 448(7151), 313–317.

Okita, K., Matsumura, Y., Sato, Y., Okada, A., Morizane, A., Okamoto, S., et al. (2011) A more efficient method to generate integration-free human iPS cells. *Nature Methods*, 8(5), 409–412.

Okita, K., Nakagawa, M., Hyenjong, H., Ichisaka, T. and Yamanaka, S. (2008) Generation of mouse induced pluripotent stem cells without viral vectors. *Science*, 322(5903), 949–953.

de Oliveira Georges, J.A., Vergani, N., Fonseca, S.A.S., Fraga, A.M., de Mello, J.C.M., Albuquerque, M.C.R.M., et al. (2014) Aberrant patterns of X chromosome inactivation in a new line of human embryonic stem cells established in physiological oxygen concentrations. *Stem Cell Reviews and Reports*, 10(4), 472–479.

O'Malley, J., Skylaki, S., Iwabuchi, K.A., Chantzoura, E., Ruetz, T., Johnsson, A., et al. (2013) High-resolution analysis with novel cell-surface markers identifies routes to iPS cells. *Nature*, 499(7456), 88–91.

Onder, T.T., Kara, N., Cherry, A., Sinha, A.U., Zhu, N., Bernt, K.M., et al. (2012) Chromatin-modifying enzymes as modulators of reprogramming. *Nature*, 483(7391), 598–602.

Ostermann, L., Ladewig, J., Müller, F.-J., Kesavan, J., Tailor, J., Smith, A., et al. (2019) In vitro recapitulation of developmental transitions in human neural stem cells. *Stem Cells (Dayton, Ohio)*, 37(11), 1429–1440.

Ozair, M.Z., Noggle, S., Warmflash, A., Krzyspiak, J.E. and Brivanlou, A.H. (2013) SMAD7 directly converts human embryonic stem cells to telencephalic fate by a default mechanism. *Stem Cells*, 31(1), 35–47.

Pageau, G.J., Hall, L.L., Ganesan, S., Livingston, D.M. and Lawrence, J.B. (2007) The disappearing Barr body in breast and ovarian cancers. *Nature Reviews. Cancer*, 7(8), 628–633.

Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., et al. (2011) Induction of human neuronal cells by defined transcription factors. *Nature*, 476(7359), 220–223.

Pankratz, M.T., Li, X.-J., LaVaute, T.M., Lyons, E.A., Chen, X. and Zhang, S.-C. (2007) Directed neural differentiation of human embryonic stem cells via an obligated primitive anterior stage. *Stem Cells*, 25(6), 1511–1520.

Park, C., Carrel, L. and Makova, K.D. (2010) Strong purifying selection at genes escaping X chromosome inactivation. *Molecular Biology and Evolution*, 27(11), 2446–2450.

Patel, S., Bonora, G., Sahakyan, A., Kim, R., Chronis, C., Langerman, J., et al. (2017) Human embryonic stem cells do not change their X inactivation status during differentiation. *Cell Reports*, 18(1), 54–67.

Pedersen, I.S. (2002) Promoter switch: a novel mechanism causing biallelic PEG1/MEST expression in invasive breast cancer. *Human Molecular Genetics*, 11(12), 1449–1453.

Peeters, S.B., Cotton, A.M. and Brown, C.J. (2014) Variable escape from X-chromosome inactivation: identifying factors that tip the scales towards expression. *BioEssays*, 36(8), 746–756.

Petropoulos, S., Edsgärd, D., Reinius, B., Deng, Q., Panula, S.P., Codeluppi, S., et al. (2016) Single-cell RNA-Seq reveals lineage and X chromosome dynamics in human preimplantation embryos. *Cell*, 165(4), 1012–1026.

Pfaff, N., Lachmann, N., Kohlscheen, S., Sgodda, M., AraUzo-Bravo, M.J., Greber, B., et al. (2012) Efficient hematopoietic redifferentiation of induced pluripotent stem cells derived from primitive murine bone marrow cells. *Stem Cells and Development*, 21(5), 689–701.

Phanstiel, D.H., Brumbaugh, J., Wenger, C.D., Tian, S., Probasco, M.D., Bailey, D.J., et al. (2011) Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nature Methods*, 8(10), 821–827.

Pick, M., Stelzer, Y., Bar-Nur, O., Mayshar, Y., Eden, A. and Benvenisty, N. (2009) Clone- and gene-specific aberrations of parental imprinting in human induced pluripotent stem cells. *Stem Cells*, 27(11), 2686–2690.

Polo, J.M., Anderssen, E., Walsh, R.M., Schwarz, B.A., Nefzger, C.M., Lim, S.M., et al. (2012) A molecular roadmap of reprogramming somatic cells into iPS cells. *Cell*, 151(7), 1617–1632.

Polo, J.M., Liu, S., Figueroa, M.E., Kulalert, W., Eminli, S., Tan, K.Y., et al. (2010) Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nature Biotechnology*, 28(8), 848–855.

Pólvora-Brandão, D., Joaquim, M., Godinho, I., Aprile, D., Álvaro, A.R., Onofre, I., et al. (2018) Loss of hierarchical imprinting regulation at the Prader–Willi/Angelman syndrome locus in human iPSCs. *Human Molecular Genetics*, 27(23), 3999–4011.

Pomp, O., Dreesen, O., Leong, D.F.M., Meller-Pomp, O., Tan, T.T., Zhou, F., et al. (2011) Unexpected X chromosome skewing during culture and reprogramming of human somatic cells can be alleviated by exogenous telomerase. *Cell Stem Cell*, 9(2), 156–165.

Pontier, D.B. and Gribnau, J. (2011) Xist regulation and function eXplored. *Human Genetics*, 130(2), 223–236.

Poppe, D., Doerr, J., Schneider, M., Wilkens, R., Steinbeck, J.A., Ladewig, J., et al. (2018) Genome editing in neuroepithelial stem cells to generate human neurons with high adenosine-releasing capacity. *Stem Cells Translational Medicine*, 7(6), 477–486.

Potten, C.S. and Loeffler, M. (1990) Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. *Development*, 110(4), 1001–1020.

Pour, M., Pilzer, I., Rosner, R., Smith, Z.D., Meissner, A. and Nachman, I. (2015) Epigenetic predisposition to reprogramming fates in somatic cells. *EMBO reports*, 16(3), 370–378.

Prelle, K., Wobus, A.M., Krebs, O., Blum, W.F. and Wolf, E. (2000) Overexpression of insulin-like growth factor-II in mouse embryonic stem cells promotes myogenic differentiation. *Biochemical and Biophysical Research Communications*, 277(3), 631–638.

Rahl, P.B., Lin, C.Y., Seila, A.C., Flynn, R.A., McCuine, S., Burge, C.B., et al. (2010) c-Myc regulates transcriptional pause release. *Cell*, 141(3), 432–445.

Ramos-Mejía, V., Montes, R., Bueno, C., Ayllón, V., Real, P.J., Rodríguez, R., et al. (2012) Residual expression of the reprogramming factors prevents differentiation of iPSC generated from human fibroblasts and cord blood CD34+ progenitors. *PloS One*, 7(4), e35824.

Rasko, J.E.J., Patel, A., Griffin, J.E., Gilleece, M.H., Radia, R., Yeung, D.T., et al. (2019) Results of the First Completed Clinical Trial of an iPSC-Derived Product: CYP-001 in Steroid-Resistant Acute GvHD. *Biology of Blood and Marrow Transplantation*, 25(3), S255–S256.

Redmer, T., Diecke, S., Grigoryan, T., Quiroga-Negreira, A., Birchmeier, W. and Besser, D. (2011) E-cadherin is crucial for embryonic stem cell pluripotency and can replace OCT4 during somatic cell reprogramming. *Nature Publishing Group*, 12(7), 720–726.

Reinhardt, P., Glatza, M., Hemmer, K., Tsytsyura, Y., Thiel, C.S., Höing, S., et al. (2013) Derivation and expansion using only small molecules of human neural progenitors for neurodegenerative disease modeling. *PloS One*, 8(3), e59252.

Revill, K., Dudley, K.J., Clayton, R.N., McNicol, A.M. and Farrell, W.E. (2009) Loss of neuronatin expression is associated with promoter hypermethylation in pituitary adenoma. *Endocrine-Related Cancer*, 16(2), 537–548.

Ring, K.L., Tong, L.M., Balestra, M.E., Javier, R., Andrews-Zwilling, Y., Li, G., et al. (2012) Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. *Cell Stem Cell*, 11(1), 100–109.

Robinton, D.A. and Daley, G.Q. (2012) The promise of induced pluripotent stem cells in research and therapy. *Nature*, 481(7381), 295–305.

da Rocha, S.T. and Heard, E. (2017) Novel players in X inactivation: insights into Xist-mediated gene silencing and chromosome conformation. *Nature Structural & Molecular Biology*, 24(3), 197–204.

Roost, M.S., Slieker, R.C., Bialecka, M., van Iperen, L., Gomes Fernandes, M.M., He, N., et al. (2017) DNA methylation and transcriptional trajectories during human development and reprogramming of isogenic pluripotent stem cells. *Nature Communications*, 8(1), 908.

Rossant, J. (2008) Stem cells and early lineage development. *Cell*, 132(4), 527–531.

Rossant, J., Gardner, R.L. and Alexandre, H.L. (1978) Investigation of the potency of cells from the postimplantation mouse embryo by blastocyst injection: a preliminary report. *Journal of Embryology and Experimental Morphology*, 48, 239–247.

Rouhani, F., Kumasaka, N., de Brito, M.C., Bradley, A., Vallier, L. and Gaffney, D. (2014) Genetic background drives transcriptional variation in human induced pluripotent stem cells. *PLoS Genetics*, 10(6), e1004432.

Rouhani, F.J., Nik-Zainal, S., Wuster, A., Li, Y., Conte, N., Koike-Yusa, H., et al. (2016) Mutational history of a human cell lineage from somatic to induced pluripotent stem cells. *PLoS Genetics*, 12(4), e1005932.

Rugg-Gunn, P.J., Ferguson-Smith, A.C. and Pedersen, R.A. (2005) Epigenetic status of human embryonic stem cells. *Nature Genetics*, 37(6), 585–587.

Rugg-Gunn, P.J., Ferguson-Smith, A.C. and Pedersen, R.A. (2007) Status of genomic imprinting in human embryonic stem cells as revealed by a large cohort of independently derived and maintained lines. *Human Molecular Genetics*, 16(R2), R243–R251.

Ruiz, S., Diep, D., Gore, A., Panopoulos, A.D., Montserrat, N., Plongthongkum, N., et al. (2012) Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells. *Proceedings of the National Academy of Sciences*, 109(40), 16196–16201.

Rust, W.L., Sadasivam, A. and Dunn, N.R. (2006) Three-dimensional extracellular matrix stimulates gastrulation-like events in human embryoid bodies. *Stem Cells and Development*, 15(6), 889–904.

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T.W., et al. (2017) Human naive pluripotent stem cells model X chromosome dampening and X inactivation. *Cell Stem Cell*, 20(1), 87–101.

Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., et al. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science*, 230(4732), 1350–1354.

Samavarchi-Tehrani, P., Golipour, A., David, L., Sung, H.-K., Beyer, T.A., Datti, A., et al. (2010) Functional genomics reveals a BMP-driven mesenchymal-to-epithelial transition in the initiation of somatic cell reprogramming. *Cell Stem Cell*, 7(1), 64–77.

Sangrithi, M.N., Royo, H., Mahadevaiah, S.K., Ojarikre, O., Bhaw, L., Sesay, A., et al. (2017) Non-Canonical and Sexually Dimorphic X Dosage Compensation States in the Mouse and Human Germline. *Developmental Cell*, 40(3), 289-301.e3.

Scerbo, P., Girardot, F., Vivien, C., Markov, G.V., Luxardi, G., Demeneix, B., et al. (2012) Ventx factors function as Nanog-like guardians of developmental potential in Xenopus. *PloS One*, 7(5), e36855.

Schopperle, W.M. and DeWolf, W.C. (2007) The TRA-1-60 and TRA-1-81 human pluripotent stem cell markers are expressed on podocalyxin in embryonal carcinoma. *Stem Cells*, 25(3), 723–730.

Schurz, H., Salie, M., Tromp, G., Hoal, E.G., Kinnear, C.J. and Möller, M. (2019) The X chromosome and sex-specific effects in infectious disease susceptibility. *Human Genomics*, 13(1), 2.

Selmi, C., Brunetta, E., Raimondo, M.G. and Meroni, P.L. (2012) The X chromosome and the sex ratio of autoimmunity. *Autoimmunity Reviews*, 11(6–7), A531–A537.

Shao, K., Koch, C., Gupta, M.K., Lin, Q., Lenz, M., Laufs, S., et al. (2013) Induced pluripotent mesenchymal stromal cell clones retain donor-derived differences in DNA methylation profiles. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, 21(1), 240–250.

Shen, Y., Matsuno, Y., Fouse, S.D., Rao, N., Root, S., Xu, R., et al. (2008) X-inactivation in female human embryonic stem cells is in a nonrandom pattern and prone to epigenetic alterations. *Proceedings of the National Academy of Sciences*, 105(12), 4709–4714.

Shi, Y., Do, J.T., Desponts, C., Hahm, H.S., Schöler, H.R. and Ding, S. (2008) A combined chemical and genetic approach for the generation of induced pluripotent stem cells. *Cell Stem Cell*, 2(6), 525–528.

Shipony, Z., Mukamel, Z., Cohen, N.M., Landan, G., Chomsky, E., Zeliger, S.R., et al. (2014) Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells. *Nature*, 513(7516), 115–119.

Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., et al. (2013) Induction of pluripotency in mouse somatic cells with lineage specifiers. *Cell*, 153(5), 963–975.

Silva, S.S., Rowntree, R.K., Mekhoubad, S. and Lee, J.T. (2008) X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. *Proceedings of the National Academy of Sciences*, 105(12), 4820–4825.

Skare, Ø., Gjessing, H.K., Gjerdevik, M., Haaland, Ø.A., Romanowska, J., Lie, R.T., et al. (2017) A new approach to chromosome-wide analysis of X-linked markers identifies new associations in Asian and European case-parent triads of orofacial clefts. *PloS One*, 12(9), e0183772.

Smith, K.P., Luong, M.X. and Stein, G.S. (2009) Pluripotency: Toward a gold standard for human ES and iPS cells. *Journal of Cellular Physiology*, 220(1), 21–29.

Smith, Z.D., Nachman, I., Regev, A. and Meissner, A. (2010) Dynamic single-cell imaging of direct reprogramming reveals an early specifying event. *Nature Biotechnology*, 28(5), 521–526.

Sommer, C.A., Christodoulou, C., Gianotti-Sommer, A., Shen, S.S., Sailaja, B.S., Hezroni, H., et al. (2012) Residual Expression of Reprogramming Factors Affects the Transcriptional Program and Epigenetic Signatures of Induced Pluripotent Stem Cells. *PloS One*, 7(12), e51711.

Soufi, A. (2014) Mechanisms for enhancing cellular reprogramming. *Current Opinion in Genetics & Development*, 25, 101–109.

Soufi, A., Donahue, G. and Zaret, K.S. (2012) Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*, 151(5), 994–1004.

Sridharan, R., Gonzales-Cope, M., Chronis, C., Bonora, G., McKee, R., Huang, C., et al. (2013) Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1γ in reprogramming to pluripotency. *Nature Cell Biology*, 15(7), 872–882.

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., et al. (2010) Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature*, 465(7295), 175–181.

Stadtfeld, M., Apostolou, E., Ferrari, F., Choi, J., Walsh, R.M., Chen, T., et al. (2012) Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all–iPS cell mice from terminally differentiated B cells. *Nature Genetics*, 44(4), 398–405.

Stadtfeld, M., Brennand, K. and Hochedlinger, K. (2008a) Reprogramming of pancreatic beta cells into induced pluripotent stem cells. *Current Biology*, 18(12), 890–894.

Stadtfeld, M., Nagaya, M., Utikal, J., Weir, G. and Hochedlinger, K. (2008b) Induced pluripotent stem cells generated without viral integration. *Science*, 322(5903), 945–949.

Steichen, C., Luce, E., Maluenda, J., Tosca, L., Moreno-Gimeno, I., Desterke, C., et al. (2014) Messenger RNA-versus retrovirus-based induced pluripotent stem cell reprogramming strategies: analysis of genomic integrity. *Stem Cells Translational Medicine*, 3(6), 686–691.

Steinberg, G.R., Kemp, B.E. and Watt, M.J. (2007) Adipocyte triglyceride lipase expression in human obesity. *American Journal of Physiology. Endocrinology and Metabolism*, 293(4), E958-964.

Subramanyam, D., Lamouille, S., Judson, R.L., Liu, J.Y., Bucay, N., Derynck, R., et al. (2011) Multiple targets of miR-302 and miR-372 promote reprogramming of human fibroblasts to induced pluripotent stem cells. *Nature Biotechnology*, 29(5), 443–448.

Sugiura, M., Kasama, Y., Araki, R., Hoki, Y., Sunayama, M., Uda, M., et al. (2014) Induced pluripotent stem cell generation-associated point mutations arise during the initial stages of the conversion of these cells. *Stem Cell Reports*, 2(1), 52–63.

Suh, M.-R., Lee, Y., Kim, J.Y., Kim, S.-K., Moon, S.-H., Lee, J.Y., et al. (2004) Human embryonic stem cells express a unique set of microRNAs. *Developmental Biology*, 270(2), 488–498.

Suhr, S.T., Chang, E.A., Tjong, J., Alcasid, N., Perkins, G.A., Goissis, M.D., et al. (2010) Mitochondrial rejuvenation after induced pluripotency. *PloS One*, 5(11), e14095.

Sun, N., Longaker, M.T. and Wu, J.C. (2010) Human iPS cell-based therapy: considerations before clinical applications. *Cell Cycle*, 9(5), 880–885.

Sun, Y., Pollard, S., Conti, L., Toselli, M., Biella, G., Parkin, G., et al. (2008) Long-term tripotent differentiation capacity of human neural stem (NS) cells in adherent culture. *Molecular and Cellular Neurosciences*, 38(2), 245–258.

Swanzey, E. and Stadtfeld, M. (2016) A reporter model to visualize imprinting stability at the *Dlk1* locus during mouse development and in pluripotent cells. *Development*, 143(22), 4161–4166.

Taapken, S.M., Nisler, B.S., Newton, M.A., Sampsell-Barron, T.L., Leonhard, K.A., McIntire, E.M., et al. (2011) Karyotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. *Nature Biotechnology*, 29(4), 313–314.

Tada, M., Takahama, Y., Abe, K., Nakatsuji, N. and Tada, T. (2001) Nuclear reprogramming of somatic cells by in vitro hybridization with ES cells. *Current Biology*, 11(19), 1553–1558.

Takahashi, J. (2017) Strategies for bringing stem cell-derived dopamine neurons to the clinic: The Kyoto trial. *Progress in Brain Research*, 230, 213–226.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., et al. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5), 861–872.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Sasaki, A., Yamamoto, M., et al. (2014) Induction of pluripotency in human somatic cells via a transient state resembling primitive streak-like mesendoderm. *Nature Communications*, 5, 3678.

Takahashi, K. and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, 126(4), 663–676.

Takahashi, K. and Yamanaka, S. (2016) A decade of transcription factor-mediated reprogramming to pluripotency. *Nature Reviews Molecular Cell Biology*, 17(3), 183–193.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., et al. (2014) Resetting transcription factor control circuitry toward ground-state pluripotency in human. *Cell*, 158(6), 1254–1269.

Tchieu, J., Kuoy, E., Chin, M.H., Trinh, H., Patterson, M., Sherman, S.P., et al. (2010) Female human iPSCs retain an inactive X chromosome. *Cell Stem Cell*, 7(3), 329–342.

Teichroeb, J.H., Betts, D.H. and Vaziri, H. (2011) Suppression of the imprinted gene NNAT and X-chromosome gene activation in isogenic human iPS cells. *PloS One*, 6(10), e23436.

Tesar, P.J., Chenoweth, J.G., Brook, F.A., Davies, T.J., Evans, E.P., Mack, D.L., et al. (2007) New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature*, 448(7150), 196–199.

Theunissen, T.W., Friedli, M., He, Y., Planet, E., O'Neil, R.C., Markoulaki, S., et al. (2016) Molecular criteria for defining the naive human pluripotent state. *Cell Stem Cell*, 19(4), 502–515.

Theunissen, T.W., Powell, B.E., Wang, H., Mitalipova, M., Faddah, D.A., Reddy, J., et al. (2014) Systematic identification of culture conditions for induction and maintenance of naive human pluripotency. *Cell Stem Cell*, 15(4), 471–487.

Thiery, J.P., Acloque, H., Huang, R.Y.J. and Nieto, M.A. (2009) Epithelial-mesenchymal transitions in development and disease. *Cell*, 139(5), 871–890.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., et al. (1998) Embryonic stem cell lines derived from human blastocysts. *Science*, 282(5391), 1145–1147.

Toivonen, S., Ojala, M., Hyysalo, A., Ilmarinen, T., Rajala, K., Pekkanen-Mattila, M., et al. (2013) Comparative analysis of targeted differentiation of human induced pluripotent stem cells (hiPSCs) and human embryonic stem cells reveals variability associated with incomplete transgene silencing in retrovirally derived hiPSC lines. *Stem Cells Translational Medicine*, 2(2), 83–93.

Tomoda, K., Takahashi, K., Leung, K., Okada, A., Narita, M., Yamada, N.A., et al. (2012) Derivation conditions impact X-inactivation status in female human induced pluripotent stem cells. *Cell Stem Cell*, 11(1), 91–99.

Tropepe, V., Hitoshi, S., Sirard, C., Mak, T.W., Rossant, J. and van der Kooy, D. (2001) Direct neural fate specification from embryonic stem cells: a primitive mammalian neural stem cell stage acquired through a default mechanism. *Neuron*, 30(1), 65–78.

Tukiainen, T., Villani, A.-C., Yen, A., Rivas, M.A., Marshall, J.L., Satija, R., et al. (2017) Landscape of X chromosome inactivation across human tissues. *Nature*, 550(7675), 244–248.

Turinetto, V., Orlando, L. and Giachino, C. (2017) Induced Pluripotent Stem Cells: Advances in the Quest for Genetic Stability during Reprogramming Process. *International Journal of Molecular Sciences*, 18(9), 1952.

Ulaner, G.A. (2003) Loss of imprinting of IGF2 and H19 in osteosarcoma is accompanied by reciprocal methylation changes of a CTCF-binding site. *Human Molecular Genetics*, 12(5), 535–549.

Umlauf, D., Goto, Y., Cao, R., Cerqueira, F., Wagschal, A., Zhang, Y., et al. (2004) Imprinting along the Kcnq1 domain on mouse chromosome 7 involves repressive histone methylation and recruitment of polycomb group complexes. *Nature Genetics*, 36(12), 1296–1300.

Urnov, F.D., Miller, J.C., Lee, Y.-L., Beausejour, C.M., Rock, J.M., Augustus, S., et al. (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 435(7042), 646–651.

Valley, C.M., Pertz, L.M., Balakumaran, B.S. and Willard, H.F. (2006) Chromosome-wide, allele-specific analysis of the histone code on the human X chromosome. *Human Molecular Genetics*, 15(15), 2335–2347.

Vallot, C., Ouimette, J.-F., Makhlouf, M., Féraud, O., Pontis, J., Côme, J., et al. (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.

Vallot, C., Ouimette, J.-F. and Rougeulle, C. (2016) Establishment of X chromosome inactivation and epigenomic features of the inactive X depend on cellular contexts. *BioEssays*, 38(9), 869–880.

Vallot, C., Patrat, C., Collier, A.J., Huret, C., Casanova, M., Liyakat Ali, T.M., et al. (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.

Vassena, R., Montserrat, N., Carrasco Canal, B., Aran, B., de Oñate, L., Veiga, A., et al. (2012) Accumulation of instability in serial differentiation and reprogramming of parthenogenetic human cells. *Human molecular genetics*, 21(15), 3366–3373.

Vierbuchen, T., Ostermeier, A., Pang, Z.P., Kokubu, Y., Südhof, T.C. and Wernig, M. (2010) Direct conversion of fibroblasts to functional neurons by defined factors. *Nature*, 463(7284), 1035–1041.

Ware, C.B., Wang, L., Mecham, B.H., Shen, L., Nelson, A.M., Bar, M., et al. (2009) Histone deacetylase inhibition elicits an evolutionarily conserved self-renewal program in embryonic stem cells. *Cell Stem Cell*, 4(4), 359–369.

Warren, L., Manos, P.D., Ahfeldt, T., Loh, Y.-H., Li, H., Lau, F., et al. (2010) Highly efficient reprogramming to pluripotency and directed differentiation of human cells with synthetic modified mRNA. *Cell Stem Cell*, 7(5), 618–630.

Weissman, I.L. (2000) Stem Cells: units of development, units of regeneration, and units in evolution. *Cell*, 100(1), 157–168.

Wernig, M., Lengner, C.J., Hanna, J., Lodato, M.A., Steine, E., Foreman, R., et al. (2008) A drug-inducible transgenic system for direct reprogramming of multiple somatic cell types. *Nature Biotechnology*, 26(8), 916–924.

Wernig, M., Meissner, A., Foreman, R., Brambrink, T., Ku, M., Hochedlinger, K., et al. (2007) In vitro reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature*, 448(7151), 318–324.

Wesselschmidt, R.L. (2011) The teratoma assay: an in vivo assessment of pluripotency. *Methods in Molecular Biology*, 767, 231–241.

West, F.D., Terlouw, S.L., Kwon, D.J., Mumaw, J.L., Dhara, S.K., Hasneen, K., et al. (2010) Porcine induced pluripotent stem cells produce chimeric offspring. *Stem Cells and Development*, 19(8), 1211–1220.

Wijnholds, J., Chowdhury, K., Wehr, R. and Gruss, P. (1995) Segment-specific expression of the neuronatin gene during early hindbrain development. *Developmental Biology*, 171(1), 73–84.

Wilmut, I., Schnieke, A.E., McWhir, J., Kind, A.J. and Campbell, K.H.S. (1997) Viable offspring derived from fetal and adult mammalian cells. *Nature*, 385(6619), 810–813.

Wilson, K.D., Venkatasubrahmanyam, S., Jia, F., Sun, N., Butte, A.J. and Wu, J.C. (2009) MicroRNA profiling of human-induced pluripotent stem cells. *Stem Cells and Development*, 18(5), 749–758.

Woltjen, K., Michael, I.P., Mohseni, P., Desai, R., Mileikovsky, M., Hämäläinen, R., et al. (2009) piggyBac transposition reprograms fibroblasts to induced pluripotent stem cells. *Nature*, 458(7239), 766–770.

Wray, J., Kalkan, T. and Smith, A.G. (2010) The ground state of pluripotency. *Biochemical Society Transactions*, 38(4), 1027–1032.

Wu, Y., Zhang, Y., Mishra, A., Tardif, S.D. and Hornsby, P.J. (2010) Generation of induced pluripotent stem cells from newborn marmoset skin fibroblasts. *Stem Cell Research*, 4(3), 180–188.

Wutz, A. (2012) Epigenetic alterations in human pluripotent stem cells: a tale of two cultures. *Cell Stem Cell*, 11(1), 9–15.

Xie, H., Ye, M., Feng, R. and Graf, T. (2004) Stepwise reprogramming of B cells into macrophages. *Cell*, 117(5), 663–676.

Xu, X., Smorag, L., Nakamura, T., Kimura, T., Dressel, R., Fitzner, A., et al. (2015) Dppa3 expression is critical for generation of fully reprogrammed iPS cells and maintenance of Dlk1-Dio3 imprinting. *Nature Communications*, 6(1), 6008.

Yamada, M., Johannesson, B., Sagi, I., Burnett, L.C., Kort, D.H., Prosser, R.W., et al. (2014) Human oocytes reprogram adult somatic nuclei of a type 1 diabetic to diploid pluripotent stem cells. *Nature*, 510(7506), 533–536.

Yamanaka, S. (2012) Induced pluripotent stem cells: past, present, and future. *Cell Stem Cell*, 10(6), 678–684.

Yao, S., Sukonnik, T., Kean, T., Bharadwaj, R.R., Pasceri, P. and Ellis, J. (2004) Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, 10(1), 27–36.

Yoshihara, M., Araki, R., Kasama, Y., Sunayama, M., Abe, M., Nishida, K., et al. (2017) Hotspots of de novo point mutations in induced pluripotent stem cells. *Cell Reports*, 21(2), 308–315.

Yoshioka, N., Gros, E., Li, H.-R., Kumar, S., Deacon, D.C., Maron, C., et al. (2013) Efficient generation of human iPSCs by a synthetic self-replicative RNA. *Cell Stem Cell*, 13(2), 246–254.

Yu, J., Hu, K., Smuga-Otto, K., Tian, S., Stewart, R., Slukvin, I.I., et al. (2009) Human induced pluripotent stem cells free of vector and transgene sequences. *Science*, 324(5928), 797–801.

Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., et al. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science*, 318(5858), 1917–1920.

Yusa, K., Rad, R., Takeda, J. and Bradley, A. (2009) Generation of transgene-free induced pluripotent mouse stem cells by the piggyBac transposon. *Nature Methods*, 6(5), 363–369.

Zhang, Y., Castillo-Morales, A., Jiang, M., Zhu, Y., Hu, L., Urrutia, A.O., et al. (2013) Genes that escape X-inactivation in humans have high intraspecific variability in expression, are associated with mental impairment but are not slow evolving. *Molecular Biology and Evolution*, 30(12), 2588–2601.

Zhao, T., Zhang, Z.-N., Rong, Z. and Xu, Y. (2011) Immunogenicity of induced pluripotent stem cells. *Nature*, 474(7350), 212–215.

Zhou, T., Benda, C., Dunzinger, S., Huang, Y., Ho, J.C., Yang, J., et al. (2012) Generation of human induced pluripotent stem cells from urine samples. *Nature Protocols*, 7(12), 2080–2089.

Zhou, Q., Brown, J., Kanarek, A., Rajagopal, J. and Melton, D.A. (2008) In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature*, 455(7213), 627–632.

Zhou, W. and Freed, C.R. (2009) Adenoviral gene delivery can reprogram human fibroblasts to induced pluripotent stem cells. *Stem Cells*, 27(11), 2667–2674.

Zhou, H., Wu, S., Joo, J.Y., Zhu, S., Han, D.W., Lin, T., et al. (2009) Generation of induced pluripotent stem cells using recombinant proteins. *Cell Stem Cell*, 4(5), 381–384.

10 Danksagung

Herrn Prof. Dr. Oliver Brüstle möchte ich sowohl für die Chance danken, dass ich zu einem solch spannenden Thema unter ausgezeichneten Forschungsbedingungen an seinem Institut promovieren durfte als auch für die Möglichkeit, dass ich während meiner Zeit als Doktorandin an renommierten internationalen Konferenzen teilnehmen konnte.

Mein Dank gilt außerdem Herrn Prof. Dr. Walter Witke für seine Bereitschaft zur Übernahme der Zweitgutachterschaft sowie Frau PD Dr. Gerhild van Echten-Deckert und Herrn Prof. Dr. Martin Sander für ihre Mitwirkung an meiner Prüfungskommission.

Herrn Prof. Dr. Philipp Koch danke ich für seine hilfreichen wissenschaftlichen Anmerkungen.

Herrn Dr. Michael Lenz und Herrn Prof. Dr. Martin Zenke bin ich im besonderen Maße für die überaus angenehme Zusammenarbeit sowie den anregenden wissenschaftlichen Austausch im Rahmen unserer Kollaboration dankbar.

Kathrin Rubel-Stüber, Dr. Daniel Poppe und Dr. Jonas Doerr danke ich weiterhin sehr für ihre wertvolle Freundschaft und die vielen gemeinsamen Stunden in der Zellkultur, vor allem auch an den Wochenenden, die durch sie äußerst bereichert wurden.

Meiner ehemaligen Arbeitsgruppe, in der sich über die Jahre viele gute Freundschaften entwickelt haben, möchte ich für die lustige und ausgelassene Arbeitsatmosphäre danken; mein besonderer Dank gilt dabei Jérôme, Ksenia, Svenja, Steffi, Laura, Karolina und Jasmin.

Svenja Brück und David Kühne danke ich zudem für die hervorragende technische Unterstützung in der Arbeitsgruppe sowie Anke Leinhaas für die verlässliche Durchführung der Transplantationen im Zuge der Teratomassays.

Allen weiteren netten und hilfsbereiten Kolleg*innen am Institut für Rekonstruktive Neurobiologie möchte ich für ihre Unterstützung im Institutsalltag und das positive Arbeitsklima herzlichst danken.

Für das fleißige Korrekturlesen bedanke ich mich vielmals bei Katja, Daniel, Jonas und Felix.

Katja, Felix, Sofia und Matthias danke ich darüber hinaus für den erfrischenden Tapetenwechsel und Patrick, dass er mich ein Stück des Weges begleitet hat.

Außerdem danke ich all meinen Freund*innen für Aufmunterung, Zuspruch und Zerstreuung, motivierende Gesten und Worte, Beistand und Rat und noch so vieles mehr. Insbesondere möchte ich dabei Katja, Sofia, Kathrin, Fabian, Erick, Felix, Daniel und Jonas von Herzen danken.

Ganz besonderer Dank gilt schließlich meinen Eltern und meiner Familie für die tatkräftige Unterstützung, den immerwährenden Rückhalt sowie den sicheren Hafen, den ich immer ansteuern kann.

11 Publications

Parts of this thesis have been published at the following locations:

Haubenreich C., Lenz M., Walter J., Schuppert A., Koch P., Zenke M. and Brüstle O.
Epigenetic and transcriptional alterations in human neural stem cells subjected to reprogramming and subsequent redifferentiation (in preparation).

de Boni L.*, Gasparoni G.*, **Haubenreich C.***, Tierling S., Schmitt I., Peitz M., Koch P., Walter J., Wüllner U. and Brüstle O.
DNA methylation alteration in iPSC- and hESC-derived neurons: potential implications for neurological disease modeling
Clinical Epigenetics 2018 Jan 29; 10:13. https://doi.org/10.1186/s13148-018-0440-0

**equal contribution*