General discussion

This thesis set out to test the functionality of human iPS technology for human brain disease modeling. In the preceding chapters I reported on several studies where we successfully used iPS technology to answer questions on human molecular and cellular neurobiological functioning. We established a simplified protocol for obtaining mature neuronal networks and revealed on transcriptional regulation of human *BDNF* and sublocalization of human UBE3A. We reported that the reprogramming procedure leads to silencing of the *FMR1* gene even in a healthy individual without concomitant methylation of the full mutation. Lastly, we identified a functional variant associated with lower risk for AD.

Nonetheless, since the emergence of iPS technology several features of its use have come to light that require proper attention. The largest and most disturbing discovery is that not all pluripotent stem cells are equal in their capacity to differentiate into desired cell types *in vitro*. Numerous studies now point towards variation at the genetic and epigenetic level between clones that result in functional variability between cell lines and heterogeneity between clones. Below I discuss the sources of this variability and how we have combatted these in our studies.

DONOR CELL-INDUCED GENETIC VARIABILITY

The first introduction of genetic variability arises with the choice of donor cell from which an iPS line is generated. Nowadays, many different cell types have proven suitable as donor cell. In the initial publication on reprogramming by Takahashi et al. dermal fibroblasts and fibroblast-like synoviocytes were used[1]. Ever since other groups confirmed that also blood erythroblasts, hair keratinonocytes[2,3], cells from tubular networks from the ureters, bladder and urethra disposed in urine[4,5], and dental pulp cells[6] are converted to iPS by the Yamanaka factors Oct4, Sox2, Klf4, c-Myc. Also cells derived from lesser accessible tissues proved sufficient such as neural stem cells, hematopoietic stem cells and liver cells[7]. Although all of these cells are originally formed from different germ layers and their conversion towards a pluripotent state is possible, increasing reports document that the efficiency differs as a function of the donor cell source[7,8]. This may depend on endogenous expression of the Yamanaka factors themselves[9]. Regardless, there seem to be no limitations depending on sex, ethnic group, disease condition, or interestingly age. This latter point however may require extra attention. As individuals age, their DNA accumulates mutations either induced by the environment or because of mistakes in the DNA proofreading process during cell division[10,11]. These somatic mutations not necessarily turn into harmful tissue for the individual, yet this phenomenon in iPS-based studies may pose a problem: the starting donor cell culture may be genetically heterogeneous. Several groups indeed confirm this[12–14]. Albeit a small population of cells, there are unique mutations not present in the culture as whole. Next to inherent heterogeneity of the used tissue, a mutational load for cell divisions (approximately 0.02 per cell division[12]) also applies. While the contribution of variability to the culture is small, the subsequent step

in iPS line generation requires reprogramming and colony picking. Here individual cells form individual colonies and initial neglectable variety runs the risks of being established within a cell line. Interestingly also, several studies suggest that somatic mosaicism, the presence of multiple cell clones with different genotypes in the same individual, is common in normal development[15,16]. This poses a dilemma on modeling. What is the reference genome or are the reference genomes? What is the contribution of each? Are somatic mutation facilitating the phenotype in an individual or are they non-functional?

In our studies we have tried to deal with donor cell variability is several ways. Firstly, our iPS lines were derived from skin fibroblasts, where our oldest donor was 57 years old and our youngest donor 3 years old. We made use of skin fibroblasts because of their large source and ease for culturing. This would keep the culture-induced mutation rate as low as possible.

Ideally, we would use younger cells, such as hematopoetic stem cells which are rare in peripheral blood, but rich in bone marrow, umbilical cord blood and placenta[17]. Moreover, these last two have multi-lineage differentiation potential and a low mutational load. However in practice this may pose a problem. Such cells are not commonly stored. Since a large group of psychiatric disorders and degenerative disorders present themselves only decades after birth, a large source of donor cells may be the next best option in line for modeling them with iPS. Next to fibroblasts, another convenient source of cells are urine-derive donor cells[4,5]. Also no medical assistance is necessary to obtain them. However little is known about this derived source. Peripheral blood also represents itself as a rich source, yet it contains erasable immunogenic marks, and may contain infections[7,8]. Overall, conscious decisions should be made with respect to donor cell type, and quality control checks for spotting heterogeneity in donor cell population may be of help.

REPROGRAMMING-INDUCED GENETIC VARIABILITY

Apart from variability induced by the donor cell population, several groups have reported on additional mutations and genomic alterations after reprogramming. Gore et al. indicated that in 22 tested iPS line an average of 6 exomic mutations per line was gained. It is unclear though at which passage number the lines were tested. Interestingly, Ji et al. indicate an average of 12 mutations per iPS cell line at passage 6[13]. Their study focused on the derivation of 5 individual iPS lines from one fibroblast source. Additionally, large chromosomal aberrations were also found in derived iPS lines. Several groups report on abnormal chromosomal aneuploidy (multiple copies of the same chromosome), chromosomal trisomies[14,18], copy-number variants (CNV)[18,19], and deletions and duplications[18,20]. Taapken et al.[21] reported that of 552 cultures of 219 iPS lines, 12,5% of the cultures have an abnormal karyotype. This indicates that significant genomic aberrations emerge during reprogramming, colony picking, and expansion.

The genes affected by these mutations were not random. Many of the mutations were found in genes related to cancer[12], and culturing specifically selected for them[12]. These mutations may give the cells a growth advantage. CNVs were also found in genes with established roles in cancer[18]. Most chromosomal aberration were detected on chromosome 12 and 17 which carry genes benefitting embryonic tumors and stem cell adaptation[14]. Similarly, Hussein et al. ascertained that compared to 6596 common CNVs found in 270 healthy individuals, 37% of the found CNVs were novel but enriched in maintaining an undifferentiated state, or associated with human ES differentiation and maintenance[19]. They also indicated that deletions were commonly found in common fragile sites in the genome and subtelomeric regions. Although others could not confirm that[18]. On a karyotype level trisomy 12 was the predominant abnormality in 31,9% of the hundreds of iPS lines tested. However 42% of the located chromosomal abnormalities were nonrecurrent between lines.

Testing for mutations in the gene however only represents the genomic status in that moment, as mutations seem to be acquired and lost with passaging. In a small study Ji et al. indicated that at passage 12, 2 of the 5 tested iPS lines had lost 2, and 1 point mutations, and two iPS lines had gained 1, and 3 point mutations[13]. In another study an increase of 4 mutations from passage 9 to passage 40 was found[12]. A rough estimation therefore is approximately 1 mutation per 10 passages. However both studies examined mutational burden in the exome. Additional mutations may have been incorporated in the non-coding genome as well. Apart from the exome, the DNA also holds regulatory sequences, the proper functioning of which ensures adequate transcriptional regulation of the cell[22,23]. Therefore, the amount and effect of acquired mutations may in fact be higher. Long-term culture also increases genomic abnormalities, where aneuploidy is rare in low passage iPS, but increase at later passages[24]. For example in one iPS line (hiPSC 18)[25,26] Marshay et al. measured a normal karyotype at passage 45, passage 58 presented a mosaic cell line with normal cells and trisomic cells containing three copies of chromosome 12. However at passage 63 the line had acquired a full trisomy of chromosome 12. Deletions were mostly found in early passages (passage 5-8), and duplications in later passages (passage 25-34)[20]. Some early deletions actually receded, indicating that they are positively selected for during reprogramming, but negatively selected for during passaging. With regard to CNVs Hussein et al. found that they were negatively correlated with passage numbers[19]. This indicated that with passaging CNVs were selected against, and their number and length decreased over passaging time. Over time, therefore, cultures were mosaic. Others however did not find an association between CNVs and passage number[18].

To ensure that our iPS lines did not carry genetic abnormalities we checked their karyotype between p5-p10 after colony picking and every 10 passages. We kept our lines in culture for the least amount of time necessary. Lines with aberrant karyotypes were not used for subsequent studies. Yet we did not perform exome or whole-genome sequencing at any of the passages. A major challenge we encountered was that individual clones are selected not only in the reprogramming procedure, but also in iPS maintenance. Culture of any given line in

routine-practice therefore is highly branched. A way to combat this disadvantage is to work with highly efficient reprogramming strategies, and iPS maintenance protocols that are robust and standardized such that colony picking is prevented as much as possible. Next to this, an administrative system to keep close track of genetically surveyed lines, and their pedigree relationship between cryopreserved stocks, live cultures and cells from which data is derived may greatly benefit detecting any genetic abnormalities that may obscure experimental data.

REPROGRAMING-INDUCED EPIGENETIC VARIABILITY

Next to genetic variability, also epigenetic variability occurs in cell culture. In essence cellular reprogramming as is done by the Yamanaka factors, results in the repression of genes responsible for differentiation and activation of genes responsible for reprogramming. Here epigenetic marks are responsible for the gene-specific expression.

Different types of epigenetic marks exist. They are divided in two major classes[27]: DNA methylation and histone modifications. DNA methylation is a biochemical process where a methyl group (CH3) is covalently bound to the cytosine in the DNA. Through this modification access to the DNA is hampered. Also methyl-CG-binding domain proteins can be recruited. They remodel histones and form compact, inactive chromatin so-called heterochromatin. Regularly high repeats of CG's are found near gene promoters and transcriptional start sites. These are called CG-islands. These islands are targets for methylation. Methylation of CG-islands generally leads to inhibition of transcriptional activity of genes in their vicinity, whereas unmethylated CG-islands allow activation.

For most genetic locations DNA methylation is identical on both alleles. However, at imprinted genes and X-chromosomes though, only a single allele is methylated normally. This results in silencing and parental-specific expression of this gene. At this point about 60 human genes are known to be imprinted[28]. There are imprints that are established in the germline, whereas others are derived in somatic cells during early embryonic development. Imprinting defects are amongst others associated with neurodevelopmental diseases such as Silver-Russell, Beckwith-Wiedermann, Prader-Willi syndromes and Angelman Syndrome[29].

The second class of epigenetic marks is histone modification. Histones are proteins around which the DNA winds itself. Wound up DNA together with the histone is called a nucleosome. Histones can also undergo covalent modifications such as acetylation, phosphorylation, methylation, SUMOylation, and ubiquitination[30].

A another class of epigenetic-related processes is covered by regulation through noncoding RNA expression[27]. It has become evident that noncoding RNAs are involved in controlling several epigenomic phenomena. One example is the dosage compensation mechanism of the X-chromosome through the long non-coding RNA, *XIST* (X-inactive specific transcript). This mechanism ensures X-chromosome inactivation (XCI). However noncoding RNAs are

also involved with silencing genes and repetitive DNA sequences by post-transcriptional and transcriptional RNA interference-related pathways through microRNAs and siRNAs.

Cellular reprogramming requires the substitution of the donor cell epigenetic marks, which normally are stably inherited through subsequent divisions, with that of the epigenetic marks specific to iPS cells[29]. Subsequent modeling of human brain cells then obliges remodeling of the iPS epigenetic landscape to that of the desired brain cell. This however turns out not to be straightforward. In the original study by Takahashi et al. iPS were promoted for their comparison to ES cells with respect to morphology, proliferation, gene expression and differentiation potential[1]. However, at the epigenetic level iPS and ES cells share some differences. For example, when DNA methylation patterns in iPS cells are compared to those in ES cells, differentially methylated regions (DMR) in genes are detected[31–34]. Similarly, at several genes methylation patterns are found in iPS that are specific to the donor cell, but are not found in ES cells. This epigenetic memory phenomenon can either be labeled as aberrant or incomplete reprogramming, or as an iPS-specific epigenetic signature. Lister et al. indicated that 51-56% of 3507 DMRs in CG islands found between iPS on the one hand and donor cells or ES cells on the other hand, were specific to the iPS cells only. Sixty-nine percent of these DMRs were present in at least two iPS lines, and 16% of the DMRs were found in 5 iPS lines. These may represent iPS-specific epigenetic signatures. These iPS-specific signature marks were spread over the genome so they did not specifically disrupt certain processes. In these 5 lines 92% of the DMRs turned out to be hypomethylated compared to the donor cell, indicating that it mostly was methylation that was not properly reset.

Ohi et al. found a similar trend. They differentiated hepatocytes, newborn foreskin fibroblasts, and adult melanocytes to iPSs. In low passage iPS cells (below 20) they found that genes that were expressed at high levels in donor cells, were repressed in iPS, yet their expression remained higher than in ES cells. The same applied for poorly expressed genes in donor cells: they were more highly expressed in iPS, but not as high as in ES cells. Next to this they found that DMRs were not dependent on any of the donor cell type. However, they did find a non-random pattern of incompletely silenced genes. These genes tended to be physically isolated from other genes that did undergo silencing. This could indicate that the silencing machinery or DNA methyltransferases may be inefficient or delayed at certain donor genes.

Also, when iPS were differentiated to trophoblast lineage cells, hundreds of DMRs were found between ES cells and ES-derived tryphoblasts[31]. The differences were attributed to donor cell DMRs, and iPS-specific DMRs. This indicated that aberrant methylation is maintained in differentiated lineages. Bar-Nur et al. reprogrammed pancreatic islet beta cells towards iPS, and found that pluripotency genes indeed were active, however donor cell genes were more methylated[34]. Also, hypomethylated genes in the islet cells were still hypomethylated in the iPS line, while normally methylated in lines derived from fibroblasts, or in ES cells. Next to this 29 mega-regions of dissimilar methylation were found in genomes[31]. Half of them were greater

than 1 MB, the largest was 4.8 MB. Many of these regions were found in close proximity to centromeres and telomeres.

Nazor et al. also found aberrant methylation in differentiated cell types[35]. They studied methylation in several female iPS lines, and discovered that numerous had partial or low methylation of X-chromosomes. This coincided with *XIST* expression, where a higher expression of the non-coding RNA *XIST* that mediates silencing, was related to higher methylation levels on the X-chromosome. This difference was found even though all clones were passaged and managed in the same way. Also, where the majority of lines in early passages showed XCI and *XIST* expression, at late passages they showed loss of XCI and *XIST* expression. Similar patterns were observed by Mekhoubad et al.[36] This loss of imprinting resulted in biallelic expression of the X-chromosomes. When these cells were differentiated to the NPC and OPC lineages, these partial methylation patterns persisted. Apart from epigenetic changes that are established during reprogramming and passed through to differentiated lineages, epigenetic changes thus also occur during passaging over time.

We did not perform assays on DNA methylation patterns in our derived iPS or differentiated neural cells. However we did experience the epigenetic altering effect of reprogramming in our studies. In chapter 5 we worked with fibroblasts from a healthy individual who carries a full mutation of the *FMR1* gene. Where a full mutation normally induces silencing of the gene by DNA methylation of the *FMR1* promoter and additional histone modifications, this individual carried unmethylated *FMR1* alleles in fibroblasts and showed *FMR1* expression. To study the effects of epigenetic silencing of *FMR1* in fragile X syndrome we reprogrammed these fibroblasts into iPS lines. However in the iPS state, the *FMR1* promoter of this healthy individual was methylated. This illustrates an example of the effect of reprogramming on the epigenome that render iPS unusable for modeling. Since in this case the epigenetic silencing process was our area of focus, this discrepancy in methylation status came to our attention. However, certain epigenetic marks may play subtle roles in disease modeling, and where the involved marks are even unknown, these as of yet unpredictable differential epigenetic marks may cause variability and faulty results.

One of the limitations of our study is that we did not evaluate the methylation pattern of the PWS-IC of the cells used for our UBE3A localization experiments (Chapter 4). A methylated PWS-IC inhibits expression of UBE3A-ATS. This long non-coding RNA silences expression of the *UBE3A* gene. The UBE3A-ATS is exclusively expressed in neurons. As such, in neurons derived from iPS with unmethylated PWS-ICs no UBE3A expression would be observed. Nonetheless, we observed UBE3A expression in NPCs derived from iPS generated from fibroblasts from an AS patient and in NPCs and neurons of a healthy control. Therefore we assume that the PWS-IC must have been methylated. However, we are not able to ascertain that UBE3A expression in neurons from the healthy control was not due to two active UBE3A alleles. Nonetheless, several studies reported on fibroblast-derived iPS lines where the PWS-IC centers in different iPS lines was methylated[37–39], confirming the use of iPS for studying

UBE3A expression. However a recent study found differential methylation of PWS-IC and subsequent aberrant expression of the closely located *SNRPN* gene, pointing towards the necessity to thoroughly check methylation status in genes under investigation[40].

INHERENT GENETIC VARIABILITY

Several researchers have established acquired genetic and epigenetic variability not to be the biggest source of *in vitro* variation[41–43]. Inherent genetic variation between individuals seems to play a much larger part. Burrows et al. collected blood and fibroblasts for iPS reprogramming from two males, and two female individuals. Deriving multiple lines from each donor cell source allowed them to compare cell-type of origin, epigenetic memory, and their intra- and inter-individual components to variability. Their gene expression and DNA methylation data showed that the contribution of cell type of origin to variation in gene expression and methylation data was very small. There was an epigenetic memory of the donor cells in the iPS lines, but this contributed only marginally to variation. This was also confirmed by others[43]. Burrows et al. concluded that only a handful of differentially methylated sites influenced regulatory variation but that genetic background captured a much greater proportion of the variation seen in gene expression and methylation assays.

This fact was also supported by Kyttälä et al.[42] who showed that only 7-25% of the DMRs resemble those from the donor cell. On average 70% of these DMRs are equal to those found in ES cells. They find that the majority of variance found in gene expression and methylation assays is dependent on genetic background. The genes differentially expressed between donors were mostly those encoding transcriptional factors related to maintenance and differentiation of iPS. Also, when iPS were differentiated they found that the differences in gene expression in iPS were reflected in the gene expression in differentiated cells types as well.

Subsequent studies were able to more concretely define genetic variability to the outcomes of measured *in vitro* variance[18]. In a comparative study using hundreds of lines from 301 individuals, Kilpiken et al. calculated that 21.4-45.8% of variance measured in immunocytochemical stainings, and 7.8%-22.8% of variance measured in cellular morphology is attributable to genetic variability between individuals. In a gene expression assay were 25.434 probes were tested, variation of 46.4% of the probes was explained by gene variability. CNVs, culture conditions, passage numbers or gender explained 23.4%, 26.2%, 2%, and 1.9% of the variance, respectively. Carcamo-Orive et al. found that 50% of the variance found in gene expression data was explained by genetic background. They added to this that several expression quantitative trait loci correlated with gene expression levels. It appeared that *cis*-regulatory variants contributed more to variance than shared environment and technical processing. The genes that varied most were related to developmental processes such as pattern specification,

regionalization, and organ and embryonic morphogenesis. This indicated that developmental pathways contributed at length to variability between lines.

In our studies, to correct for variability between lines, we made use of three lines from three different individuals in Chapter 2, where we tested the robustness of our neuronal differentiation protocol. For the remaining of our studies we made use of one line per condition. Despite the reported variability we found robust and reproducible results in our experiments. The reason we asked very specific questions that did not require comparison between different iPS lines may have been the reason for that: in Chapter 2 we were testing the robustness of our protocol which was confirmed by the development of neuronal cultures from three independent lines in several rounds of differentiation. In Chapter 3 we made use of one ES line, and one iPS line for investigating transcriptional control of the *BDNF* gene. In Chapter 4 we compared overexpression of different UBE3A protein isoforms. In Chapter 5 we evaluated the epigenetic characteristics of the *FMR1* promoter of several lines before and after reprogramming. As such our experiments did not suffer from inherent genetic variability between lines.

Whenever phenotypes between iPS lines are to be compared, an alternative study design would be the use of isogenic lines: lines theoretically only differing in the genetic perturbation to be studied.

In general three types of iPS-based studies are discernable: the study of fundamental biology irrespective of genetic variations or mutations, the study of a monogenic disease, or a multigenic disease. In each study the genetic background of the stem cell line to be used needs to be considered carefully. Commonly donor cells from a healthy subject are used to generate control stem cells lines. Control cell lines are used for fundamental studies unrelated to genetic variations and mutations as well as in 'patient vs control' designs. Yet the question remains what healthy subjects entail. Individuals could be free of disease at the moment of assessment, and develop disease later in life. Collection of donor cells from aged subjects may not be accommodating because of mutational load in older cells[10]. In this case selection of sufficient amount of lines would mitigate line-specific effects. This makes the selection of youthful donor cells-derived lines a possibility, in that way eliminating high mutation load derived from adult somatic cells. Retrospect check-up on donors when they reach the critical age for disease development is an alternative as well.

In case of studying monogenic diseases with strong effect sizes it may suffice to select unrelated controls[44–46] (such as in chapter AS) or at best healthy family members to compare with the patient-specific lines[47]. In case of smaller effect sizes variation between lines can partially be taken away by making use of isogenic lines, lines identical other than the mutation to be studied. This solution should mitigate the unintended genetic and epigenetic variability that remains between two unrelated stem cell lines[48]. Yet several points have to be taken into account. Commonly used procedures for gene editing entail zinc fingers, TALENs and CrisprCas9 where the use of the latter nowadays becomes standard-practice in labs. Nonetheless, albeit they are being improved, these techniques can unintentionally create mutations elsewhere

in the genome in the editing procedure[49]. Several economical high-throughput methods are being developed to check the genome for additional mutation. Yet these would have to become standard-practice as well in laboratories as often only *in silico* predicted off-targets are examined. Unfortunately there is accumulating evidence that the current algorithms have low prediction accuracy[49]. Taking into account the variability induced by the *in vitro* procedures, the chance on differences between derived isogenic lines in point mutations as well as aberrant epigenetic landscapes increases with every passage (see above).

In case of studying multigenic diseases where uncertainty remains on which genetic perturbation is responsible, patient and control selection is a delicate business. Some suggest that selecting patients and controls with clustered risk-scores may create enough power for phenotype detection[50]. Surely high numbers of lines are necessary to tease out the phenotype.

Nonetheless, where possible the use of isogenic lines seems the best way forward to reduce genetic and epigenetic variability.

IN VITRO DIFFERENTIATION-INDUCED VARIABILITY

Another source of variability seen in human brain modeling studies is the neural differentiation procedure itself. Most differentiation strategies are based on modulating naturally occurring specialization in the brain. In development, neurons are derived from the ectoderm, one of the three germ layers generated in early embryogenesis[51]. Ectoderm forms the neural tube, which gives rise to the brain and spinal cord. These developmental steps are triggered by the expression of morphogens in strict patterns along the rostral-caudal axis (Fibroblast Growth Factors (FGFs), Wingless/Int (WNTs), retinoic acid (RA)) and ventral-dorsal (WNTs, Bone Morphogenetic Proteins (BMPs), Sonic Hedgehog (Shh). Under influence of FGF and RA ectodermal tissue develops into neuroepithelia. Subsequently, a specific combination of morphogens in the neural tube triggers the neural stem cells in that area to develop into either neuronal or glial progenitors[52]. These progenitors differentiate into mature neurons or astrocytes and oligodendrocytes, respectively. In the differentiation process similar developmental stages are passed. An important step is the induction of neuroectoderm. From there on, neural stem cells and precursors continue on to differentiate into specific neuronal subtypes with or without addition of specific morphogens.

There are several ways to induce neuroectoderm *in vitro*. One way is the isolation of neuroectoderm from embryoid bodies[53]. Embryoid bodies (EBs) are three-dimensional cellular aggregates of IPSCs, obtained when cells are grown in suspension. This method allows the spontaneous differentiation of IPSCs to cells of the three germ layers. When treated with specific growth factors or morphogens such as RA, the proliferation of neuroectoderm is promoted. Subsequently, cells are plated in neuronal supporting media. However, the drawback of this system is that embryoid bodies can vary in size, which results in inconsistent yields

of neural progenitors. Since the inner cell layers of embryoid bodies are difficult to reach for morphogens, radial concentration gradients emerge which induces heterogeneous cell types.

As EBs are derived from iPS, their homogeneity also plays an important part. *In vivo* the pluripotent state is a transient one, such that *in vitro* critical media components are necessary to maintain the pluripotent state[1,54–57]. Yet iPS occasionally escape the pluripotent state and randomly differentiate, thereby reducing the line's overall pluripotency. Whenever these cells are used for targeted terminal differentiation the outcome is inevitably a mixed population of desired and undesired cells.

Kilpinen et al. tested over hundreds of lines from 301 individuals and found that 84% of them are classified as pluripotent by the Pluritest, a tool for pluripotency assessment by whole genome expression analysis[18]. In an average iPS line 18-62% of the cells co-express the pluripotency markers NANOG, OCT4 and SOX2. Whenever an iPS line was differentiated to one of the germ layers 70%, 84% and 77% of the cells in the line would express markers specific for respectively endoderm, mesoderm, and ectoderm. This indicates that roughly one-sixth of the lines are pluripotent, and that roughly one-fourth of the cells do not differentiate to the desired germ layer.

After neural induction of EBs from iPS, EBs are commonly plated and display neural rosettes. Series of radial migration of NPCs occur, yet however not all these NPCs are the same[58]. They are an ensemble of several neuronal precursors such as radial glia, intermediate progenitors, symmetrically and asymmetrically dividing NPCs[59], but also progenitors of oligodendrocytes, and astrocytes[60]. Depending on the question to be asked, treating them as one population may result in high batch-to-batch differences. Next to that, terminal neural differentiation highly depends on the composition of the original NPC population. As such, mixed NPC populations may lead to dissimilar terminally differentiated neural cultures if the ratios of the different types of progenitors are not the same. As terminally differentiated cultures derived from NPCs by dual-SMAD inhibition also go through a neural rosette stage, the same caution should be taken.

Apart from the cell type diversity of the NPC population, another point to take into account is the age of the cell. Regular passage of NPCs may contain migrated NPCs, as well as newly born NPCs. These cells represent different neurodevelopmental stages of NPCs: each of these cells may have a different temporal-spatial expression profile, such as in the brain. Practically, the first rounds of passaging of the NPC population deliver mostly neuronal precursors, where later populations produce more astrocytes. Whenever this tipping point occurs though is unknown. Yet in our studies we find this to be around 10 passages of the NPC population. During *in vivo* brain development different NPC populations co-exist simultaneously; however, for modeling neurodevelopmental diseases the fine balance of the population may be crucial.

The shortest protocols to differentiate neuronal cells with basic electrophysiological properties from a common neural progenitor need 6 weeks of *in vitro* culturing from a neuroectoder-

mal stage. As the protocols to generate neurons lengthen and several procedures ((sub)plating, refreshing, stable environmental factors) need to be performed, well-to-well variability is increased, such as differences in cell density and cellular heterogeneity. Volpato et al. tested the reproducibility of the Shi et al. protocol[61,62]. They used two lines: one control line, and one line containing a mutation in the *PSEN1* gene, in 5 different laboratories and examined the RNA and protein profiles of differentiated cortical neurons. Within each laboratory the differences between lines were visible (three independent neuronal induction). However across multiple laboratories differences in expression between the two lines were not consistently detectable. They indicated cell type heterogeneity as the major contributor to variability. Subsequently, they also collected 771 individual transcriptomes of cells in the culture. Here 4-5 subpopulations were identified within the neural cultures expressing neuronal, astrocytic, oligodendritic and microglial marker genes. They also found out that factors that hampered cross laboratory comparison were iPS passage number before differentiation, the number of passages before terminal differentiation plating, media volume changes, feeding at weekends, and use of frozen progenitors. They hypothesized that the factors may alter epigenetic and cellular programs that determine cell fate choice, eventually influencing the composition of the final culture.

For the majority of the studies we made use of the differentiation protocol described in Chapter 2. In this chapter we showed reproducible outcomes of neural cell composition. In Chapter 4 we made use of fluorescence-activated cell sorting (FACS) to enrich our NPC population, thereby increasing the homogeneity of the precursor population. To validate our results it may be worthwhile to test again the transcriptional activation of *BDNF* VIII-IX transcript and the methylation states of the different iPS lines used in chapter 3 and 5 respectively with FACS-sorted NPC populations to prove that the obtained results were not due to contamination of the cell population. Others have proved it to be a valuable technique to enrich the NPC population[63,64]. It should be taken into consideration though that even populations selected by canonical marker expression may still present a diversity within their own subclass.

We used real-time quantitative PCR and EB differentiation to test pluripotency of our iPS lines. However, determining when to call a line a pluripotent stem cell line remains a topic for debate. Several tests and assays are available yet none of these tests provides exclusive proof for all the genetic, epigenetic, transcriptional and translational assets of a stem cell[65]. Momentarily the community is moving towards validation of stem cell lines by gene expression data by the algorithm provided by Pluritest[66]. By computationally comparing the gene expression data of induced stem cells to *bona fide* stem cells, a cut-off score determines when an induced line is considered a stem cell line. However, once validated, stem cell lines need revalidation after passaging, manipulation and colony picking, making it practically impossible to control the exact composition of an iPS population. Here also studies would benefit from improved iPS maintenance protocols, such as methods to minimize contamination of pluripotency by spontaneous differentiation.

To define the different types of neural cells in the culture still remains challenging. However others are paving the way with single-cell analyses and systematic characterization on the basis of electrophysiological and transcriptomic profiles[67]. As the brain in development also depicts heterogeneity *in vitro* iPS-derived neural cultures may actually not stray too far from their *in vivo* counterparts[68]. Nonetheless, as pathways regarding cell fate decision and network formation in neuronal cultures are being explored and more knowledge on the different NPCs emerges, caution should be taken with treating every batch of NPCs and neuronal differentiation similar.

Reproducibility between labs is a concerning factor, as findings of previous researchers may prove non-repeatable and hence their results faulty. Comparable to iPS maintenance, protocols to generate reproducible neuronal cultures contain multiple steps, variables, and often, subjective judgment decisions. Clearer and more accurate experimental descriptions, improvement of induction protocols and pre-selection of NPC pools may benefit reproducibility between labs.

Overall, the extent to which acquired mutations in iPS lines, aberrant epigenomic markers, inherent genetic variability and heterogeneity in neural cultures hamper the ability to model human brain diseases seems to largely depend on the research question to be answered. Experimental conditions are to be tailored to these questions. Nonetheless, small effect sizes of genes, unidentified neural cell types, and the involvement of pathways in disease require the highest standard of iPS modeling. As such, the community at whole may do best to optimize iPS technology to its highest capacities to continue to unravel the molecular and cellular mechanisms underlying human brain disorders.

REFERENCES

1. Takahashi, K. *et al.* Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell* **131**, 861–72 (2007).
2. Carey, B. W. *et al.* Reprogramming of murine and human somatic cells using a single polycistronic vector. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 157–62 (2009).
3. Aasen, T. *et al.* Efficient and rapid generation of induced pluripotent stem cells from human keratinocytes. *Nat. Biotechnol.* **26**, 1276–84 (2008).
4. Wang, L. *et al.* Generation of integration-free neural progenitor cells from cells in human urine. **10**, (2013).
5. Zhou, T. *et al.* Generation of Induced Pluripotent Stem Cells from Urine. *J. Am. Soc. Nephrol.* **22**, 1221–1228 (2011).
6. Yan, X. *et al.* iPS Cells Reprogrammed From Human Mesenchymal-Like Stem/Progenitor Cells of Dental Tissue Origin. *Stem Cells Dev.* **19**, 469–480 (2010).
7. Li, J., Song, W., Pan, G. & Zhou, J. Advances in understanding the cell types and approaches used for generating induced pluripotent stem cells. *J. Hematol. Oncol.* **7**, 50 (2014).
8. Raab, S., Klingenstein, M., Liebau, S. & Linta, L. A Comparative View on Human Somatic Cell Sources for iPSC Generation. *Stem Cells Int* **2014**, 768391 (2014).
9. Tsai, S. Y. *et al.* Oct4 and Klf4 reprogram dermal papilla cells into induced pluripotent stem cells. *Stem Cells* **28**, 221–228 (2010).
10. Kennedy, S. R., Loeb, L. A. & Herr, A. J. Somatic mutations in aging, cancer and neurodegeneration. *Mech. Ageing Dev.* **133**, 118–26 (2012).
11. Milholland, B., Auton, A., Suh, Y. & Vijg, J. Age-related somatic mutations in the cancer genome. *Oncotarget* **6**, 24627–35 (2015).
12. Gore, A. *et al.* Somatic coding mutations in human induced pluripotent stem cells. *Nature* **471**, 63–67 (2011).
13. Ji, J. *et al.* Elevated Coding Mutation Rate During the Reprogramming of Human Somatic Cells into Induced Pluripotent Stem Cells. *Stem Cells* **30**, 435–440 (2012).
14. Mayshar, Y. *et al.* Identification and Classification of Chromosomal Aberrations in Human Induced Pluripotent Stem Cells. *Cell Stem Cell* **7**, 521–531 (2010).
15. Risques, R. A. & Kennedy, S. R. Aging and the rise of somatic cancer-associated mutations in normal tissues. *PLoS Genet.* **14**, 1–12 (2018).
16. De, S. Somatic mosaicism in healthy human tissues. *Trends Genet.* **27**, 217–223 (2011).
17. Haase, A. *et al.* Generation of induced pluripotent stem cells from human cord blood. *Cell Stem Cell* **5**, 434–41 (2009).
18. Kilpinen, H. *et al.* Common genetic variation drives molecular heterogeneity in human iPSCs. *bioRxiv* **546**, 055160 (2016).
19. Hussein, S. M. *et al.* Copy number variation and selection during reprogramming to pluripotency. *Nature* **471**, 58–62 (2011).
20. Laurent, L. C. *et al.* 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**, 106–118 (2011).
21. Taapken, S. M. *et al.* Karyotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. *Nat. Biotechnol.* **29**, 313–4 (2011).
22. Epstein, D. J. Cis-regulatory mutations in human disease. *Briefings Funct. Genomics Proteomics* **8**, 310–316 (2009).

23. Narlikar, L. & Ovcharenko, I. Identifying regulatory elements in eukaryotic genomes. *Briefings Funct. Genomics Proteomics* **8**, 215–230 (2009).
24. Mayshar, Y. *et al.* Identification and Classification of Chromosomal Aberrations in Human Induced Pluripotent Stem Cells. *Cell Stem Cell* **7**, 521–531 (2010).
25. Lowry, W. E. *et al.* Generation of human induced pluripotent stem cells from dermal fibroblasts. *Proc. Natl. Acad. Sci.* **105**, 2883–2888 (2008).
26. Chin, M. H. *et al.* Induced Pluripotent Stem Cells and Embryonic Stem Cells Are Distinguished by Gene Expression Signatures. *Cell Stem Cell* **5**, 111–123 (2009).
27. Goldberg, A. D., Allis, C. D. & Bernstein, E. Epigenetics: A Landscape Takes Shape. *Cell* **128**, 635–638 (2007).
28. www.geneimprint.org.
29. Nashun, B., Hill, P. W. & Hajkova, P. Reprogramming of cell fate: epigenetic memory and the erasure of memories past. *EMBO J.* **34**, 1296–1308 (2015).
30. Spivakov, M. & Fisher, A. G. Epigenetic signatures of stem-cell identity. *Nat. Rev. Genet.* **8**, 263–271 (2007).
31. Lister, R. *et al.* Hotspots of aberrant epigenomic reprogramming in human induced pluripotent stem cells. *Nature* **471**, 68–73 (2011).
32. Ohi, Y. *et al.* Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nat. Cell Biol.* **13**, 541–9 (2011).
33. Doi, A. *et al.* Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat. Genet.* **41**, 1350–1353 (2009).
34. Bar-Nur, O., Russ, H. A., Efrat, S. & Benvenisty, N. Epigenetic memory and preferential lineage-specific differentiation in induced pluripotent stem cells derived from human pancreatic islet beta cells. *Cell Stem Cell* **9**, 17–23 (2011).
35. Nazor, K. L. *et al.* Recurrent variations in DNA methylation in human pluripotent stem cells and their differentiated derivatives. *Cell Stem Cell* **10**, 620–634 (2012).
36. Mekhoubad, S. *et al.* Erosion of dosage compensation impacts human iPSC disease modeling. *Cell Stem Cell* **10**, 595–609 (2012).
37. Chamberlain, S. J. *et al.* Induced pluripotent stem cell models of the genomic imprinting disorders Angelman and Prader – Willi syndromes. (2010) Proc Natl Acad Sci,12;107(41):17668-73 (2010).
38. Stanurova, J. *et al.* Angelman syndrome-derived neurons display late onset of paternal UBE3A silencing. *Sci. Rep.* **6**, 30792 (2016).
39. Burnett, L. C. *et al.* Induced pluripotent stem cells (iPSC) created from skin fibroblasts of patients with Prader-Willi syndrome (PWS) retain the molecular signature of PWS. *Stem Cell Res.* **17**, 526–530 (2016).
40. Pólvora-Brandão, D. *et al.* Loss of hierarchical imprinting regulation at the Prader-Willi/Angelman syndrome locus in human iPSCs. *Hum. Mol. Genet.* **00**, 1–13 (2018).
41. Burrows, C. K. *et al.* Genetic Variation, Not Cell Type of Origin, Underlies the Majority of Identifiable Regulatory Differences in iPSCs. *PLoS Genet.* **12**, 1–18 (2016).
42. Kyttälä, A. *et al.* Genetic Variability Overrides the Impact of Parental Cell Type and Determines iPSC Differentiation Potential. *Stem Cell Reports* **6**, 200–212 (2016).
43. Rouhani, F. *et al.* Genetic Background Drives Transcriptional Variation in Human Induced Pluripotent Stem Cells. *PLoS Genet.* **10**, (2014).
44. Vanhauwaert, R. *et al.* The SAC1 domain in synaptojanin is required for autophagosome maturation at presynaptic terminals. *EMBO J.* **36**, 1392–1411 (2017).

45. Yeh, E. *et al.* Patient-derived iPSCs show premature neural differentiation and neuron type-specific phenotypes relevant to neurodevelopment. *Mol. Psychiatry* 1–12 (2017).
46. Rooney, G. E. *et al.* Human iPS Cell-Derived Neurons Uncover the Impact of Increased Ras Signaling in Costello Syndrome. *J. Neurosci.* **36**, 142–152 (2016).
47. de Vrij, F. M. *et al.* Candidate CSPG4 mutations and induced pluripotent stem cell modeling implicate oligodendrocyte progenitor cell dysfunction in familial schizophrenia. *Mol. Psychiatry* 1–15 (2018).
48. Germain, P. L. & Testa, G. Taming Human Genetic Variability: Transcriptomic Meta-Analysis Guides the Experimental Design and Interpretation of iPSC-Based Disease Modeling. *Stem Cell Reports* **8**, 1784–1796 (2017).
49. Tsai, S. Q. & Joung, J. K. Defining and improving the genome-wide specificities of CRISPR–Cas9 nucleases. *Nat. Rev. Genet.* **17**, 300–312 (2016).
50. Hoekstra, S. D., Stringer, S., Heine, V. M. & Posthuma, D. Genetically-Informed Patient Selection for iPSC Studies of Complex Diseases May Aid in Reducing Cellular Heterogeneity. *Front. Cell. Neurosci.* **11**, 164 (2017).
51. Lupo, G., Harris, W. a & Lewis, K. E. Mechanisms of ventral patterning in the vertebrate nervous system. *Nat. Rev. Neurosci.* **7**, 103–14 (2006).
52. Temple, S. The development of neural stem cells. *Nature* **414**, 112–7 (2001).
53. Dhara, S. K. *et al.* Human neural progenitor cells derived from embryonic stem cells in feeder-free cultures. *Differentiation.* **76**, 454–64 (2008).
54. Manuilova, E. S., Gordeeva, O. F., Grivennikov, I. A. & Ozernikov, N. D. Embryonic stem cells: spontaneous and directed differentiation. *Izv. Akad. Nauk. Seriia Biol.* 704–10
55. Sathananthan, A. H. & Trounson, A. Human embryonic stem cells and their spontaneous differentiation. *Ital. J. Anat. Embryol.* **110**, 151–7 (2005).
56. Levenstein, M. E. *et al.* Basic Fibroblast Growth Factor Support of Human Embryonic Stem Cell Self-Renewal. *Stem Cells* **24**, 568–574 (2006).
57. Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **126**, 663–76 (2006).
58. Otani, T. *et al.* 2D and 3D Stem Cell Models of Primate Cortical Development Identify Species-Specific Differences in Progenitor Behavior Contributing to Brain Size Article 2D and 3D Stem Cell Models of Primate Cortical Development Identify Species-Specific Differences in. *Stem Cell* **18**, 1–14 (2016).
59. Shi, Y., Kirwan, P., Smith, J., Robinson, H. P. C. & Livesey, F. J. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat. Neurosci.* **15**, 477–486 (2012).
60. Gunhanlar, N. *et al.* A simplified protocol for differentiation of electrophysiologically mature neuronal networks from human induced pluripotent stem cells. *Mol. Psychiatry* 1–9 (2017). doi:10.1038/mp.2017.56
61. Volpato, V. *et al.* Reproducibility of Molecular Phenotypes after Long-Term Differentiation to Human iPSC-Derived Neurons: A Multi-Site Omics Study. *Stem Cell Reports* **11**, (2018).
62. Shi, Y., Kirwan, P., Smith, J., Robinson, H. P. C. & Livesey, F. J. Human cerebral cortex development from pluripotent stem cells to functional excitatory synapses. *Nat. Neurosci.* (2012). doi:10.1038/nn.3041
63. Pruszak, J., Sonntag, K.-C., Aung, M. H., Sanchez-Pernaute, R. & Isacson, O. Markers and methods for cell sorting of human embryonic stem cell-derived neural cell populations. *Stem Cells* **25**, 2257–68 (2007).
64. Yuan, S. H. *et al.* Cell-surface marker signatures for the Isolation of neural stem cells, glia and neurons derived from human pluripotent stem cells. *PLoS One* **6**, (2011).
65. Franz-josef, M., Kiel, I. P. & Kiel, C. Assessment of human pluripotent stem cells with PluriTest. 1–17 (2012). doi:10.3824/stembook.1.83.1

66. Muller, F. J. *et al.* A bioinformatic assay for pluripotency in human cells. *Nat Methods* **8**, 315–317 (2011).
67. Bardy, C. *et al.* Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology. *Mol. Psychiatry* **21**, 1573–1588 (2016).
68. Silbereis, J. C., Pochareddy, S., Zhu, Y., Li, M. & Sestan, N. The Cellular and Molecular Landscapes of the Developing Human Central Nervous System. *Neuron* **89**, 248–268 (2016).