Does Epigenetic Memory Influence the Differentiation of MSNs from Fetal Ganglionic Eminence Derived iPS Cells?

Thesis submitted for the degree of Doctor of Philosophy

School of Biosciences, Cardiff University

2019

Oliver Jay Michael Bartley

Submission Declarations

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ……..…… (candidate) Date ……..…

STATEMENT 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD

Signed ……..…… (candidate) Date ……..…

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ……..…… (candidate) Date ……..…

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ……..…… (candidate) Date ……..…

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ……..…… (candidate) Date ……..…

Acknowledgements

"I see now that the circumstances of one's birth are irrelevant. It is what you do with the gift of life that determines who you are." *-Mewtwo, Pokémon: The First Movie*

First, I have to apologise to everyone I am about to mention. The few sentences I can write here cannot accurately portray how thankful I am to each of you.

But I will do my best.

I would like to thank my primary supervisor, Anne, for her unabating support of my research and vision throughout my PhD. I always left our meetings with a wider perspective, new ideas, and inspiration to carry me forwards. I still don't understand when you find time to sleep. I would also like to thank my other supervisors: Nigel, Sophie, and Ngoc-nga. Each of you has been pivotal to my success, and you have each taught me so much. In particular I have to thank both Sophie and Ngoc-nga for listening to my often inane (although occasionally rather brilliant) experiment ideas!

A big thank you to the many fantastic members of the BRG (both past and present), you've all been amazing and I wouldn't have wanted to do my PhD anywhere else. Most of you have helped me with things at one time or another, and without that support this whole endeavour would have been near impossible. So thank you all! I of course have to give a special mention to my fellow PhD office mates: Sus, David, Kyle, and Laura. Sus, I owe a special debt to you for the many hours you put in to teach me everything I needed for my animal experiment at a time when you had so little to spare, and of course for the laughter you brought to our office. David, I taught you about the glitter world, and you were reluctant for the whole journey - you're welcome ;) . Kyle and Laura. Together we survived the absolute *worst* buses in the world, discovered canchanchara's, and made a lifelong friend in Sonia. I doubt I'll ever be able to pronounce Radyr properly, at least not without thinking about it first, and I hope you will always enjoy that.

And where would anyone be without the support of their friends? First, to the Astoner's I haven't been able to see as much as I would have liked. Now this ordeal has come to an end I promise I'll come to parties and meals again! To my friends from BoA, especially Sally, Lucy and Vixy. You've been with me the longest and have supported me through some of the darkest times in my life. I love you all, here's to many *many* more years together. And of course, the friends I've made here in Cardiff:

Rachael, Jon, Kira, Rae, Aurelien, Ellen, Louick and Sefarrah. You already know it (and Louick and Sefarrah definitely don't need reminding), *but just to make sure you know*, you are all the best people. Thank you for these past years. I already miss you all.

I am also lucky to have the support of two families. Firstly, I have to mention my acquired family, the Hodgetts *et al.* Thank you for welcoming me into your lives with open arms. You are all wonderful people. A special mention goes to Mickey, for her unending interest in my research. Otherwise, I especially want to thank my first family: Mum and Aly (and allllll the cats of course). You have been by my side throughout my life and helped shape me into who I am today. We've been through it haven't we? It's not always been easy (in fact I'd dare to say difficult has been the default setting) and unfortunately I don't think we're through all the hard times yet. But thank you for being there, I love you both, and as ever, next year is going to be better.

Last, this is the spot where I thought I'd make a joke about how I've thanked everyone and then *"remember"* Lee, haha, yes, very funny. But honestly, I can't bring myself to even joke about his support these past few years. I'm pretty sure I would have quit this whole thing a while back without him. So Lee, thank you for being weird, thank you for being loud, and thank you for being you. I love you more than you know.

Summary

Cell replacement therapy for Huntington's Disease requires a transplantable source of authentic medium spiny neuron (MSN) progenitors from a renewable cell source. Human pluripotent stem cells (hPSC) theoretically fit the required criteria, but without further research it is unlikely we will realise their potential for this purpose, as presently we cannot reliably produce sufficient authentic MSN progenitors with these cells. There is a growing body of work that has demonstrated human induced-PSCs (hiPSC) retain an epigenetic memory of their tissue of origin, which can enhance their differentiation towards cellular phenotypes similar to that tissue of origin. As such, it is possible that hiPSCs derived from authentic MSN progenitors could retain an authentic epigenetic memory of their previous cell type that enhances their differentiation towards an MSN fate, and in this thesis I aimed to test this hypothesis.

In Chapter 3, I generated hiPSCs from the lateral ganglionic eminence (LGE), medial ganglionic eminence, and cortex of the human fetal brain, as well as fibroblast tissues. These hiPSCs were derived using a non-integrating virus, and from tissues collected from the same fetal donor to ensure these hiPSC lines were isogenic. I then validated these cells as pluripotent, and provided evidence that these hiPSCs could be defined by an epigenetic signature indicative of their tissue of origin.

In Chapter 4, I explored the differentiation potential of these cells *in vivo* and *in vitro*. Here, I found some indication that hiPSCs derived from neural tissues undergo neuronal fate commitment at a faster rate than hiPSCs derived from non-neural tissues. However, I also found evidence that the epigenetic memory of hiPSCs derived from neural tissues may conflict with the differentiation methods used here, as generally these cells underperformed compared to hiPSCs derived from fibroblasts. As such, further research will be required to determine if this epigenetic memory can be used to enhance MSN differentiations.

In Chapter 5, I conducted genome wide DNA methylation analysis on MSNs derived from these hiPSCs, and various controls including primary ganglionic eminence tissues. Here I found evidence that LGE derived hiPSCs may have an epigenome more similar to primary LGE tissue. However this similarity was found to be greatly overshadowed by a generally strong epigenetic difference between hPSC derived MSNs and the authentic primary tissues they are supposed to replace.

The work contained in this thesis has important implications for the generation of MSNs from hPSCs, specifically regarding the importance of starting cell source and the understudied role of the epigenome in fate specification.

Abbreviations

27K	Infinium Human Methylation 27K bead chips
450K	Infinium Human Methylation 450K bead chips
6-OHDA	6-hydroxydopamine
AA	Ascorbic Acid
Activin	Activin A
ANR	Anterior neural ridge
AP	Anterior-Posterior
BMP	Bone morphogenetic proteins
BSA	Bovine serum albumin
CAG	Cytosine-Adenine-Guanine
CGE	Caudal ganglionic eminence
CpG	Cytosine-phosphate-guanine
CRT	Cell replacement therapy
CRL	Crown-Rump Length
DAB	Diaminobenzidine
DKK1	Dickkopf1
DMEM+Dα	Wash media
DMP	Differentially Methylated Probe
DNA	Deoxyribonucleic acid
Dnmt	DNA methyltransferases
DPBS	Dulbecco's phosphate-buffered saline
DV	Dorsal-Ventral
Dα	Dornase Alfa
E8	Essential 8 medium
E8+	Essential 8 Flex medium

EB	Embryoid body
ESC	Embryonic stem cell
FGF	Fibroblast growth factor
GE	Ganglionic eminence
HD	Huntington's disease
hESC	Human embryonic stem cell
hiPSC	Human induced pluripotent stem cell
hPSC	Human pluripotent stem cell
hPSC-MSNs	MSNs derived from hPSC cultures
HTT	Huntingtin (gene and protein)
ICM	Inner cell mass
iPSC	Induced pluripotent stem cell
LDN	LDN193189
LGE	Lateral ganglionic eminence
LMCF	LGE, MGE, cortex, fibroblast
MDS	Multidimensional scaling plot
MEA	Multi-electrode Array
MEM-NEAA	MEM Non-Essential Amino Acids Solution
MGE	Medial ganglionic eminence
mHTT	Mutated Huntingtin
MOI	Multiplicities of infections
MSN	Medium spiny neuron
NDM	Neural Differentiation Media
OSKM	OCT4, SOX2, KLF4, C-MYC
p.c.	Post conception
PD	Parkinson's Disease

PDL	Poly-D-lysine
PFA	Paraformaldehyde
PFC	Prefrontal cortex
PS	Penicillin-Streptomycin
PSC	Pluripotent stem cell
QA	Quinolinic acid
RA	Retinoic Acid
RGF Matrigel	Growth factor reduced Matrigel
RNA	Ribonucleic acid
ROI	Region of interest
SB	SB431542
SEM	Standard error of the mean
SeV	Sendai Virus
SHH	Sonic Hedgehog
SN	Substantia nigra
TBS	Tris buffered saline
TGFβ	Transforming growth factor beta
TNS	Tris non saline
TXTBS	Triton X-100 in TBS
VM	Ventral mesencephalon
WGE	Whole ganglionic eminence
WNT	Wingless
XAV	XAV-939
α-FP	Alpha-Fetoprotein
α-SMA	Alpha-Smooth muscle actin

Contents

Chapter 1: Introduction

Cell replacement is a potential therapeutic for various neurodegenerative diseases. It follows the principle that lost function resulting from neurodegeneration may be restored by replenishing degenerated cell populations. By its nature, cell replacement therapy requires a reliable and authentic source of cells matching the phenotype of those lost during the process of a given neurodegenerative disease. Whilst there is evidence of both safety and efficacy following the transplantation of fetal neural tissues, this remains an ethically contentious source of cells that additionally suffer from serious logistical issues. Subsequently, these cells are inappropriate for widespread clinical use and an alternative source is required. There has been considerable effort to progress pluripotent stem cells as an alternative cell source for cell replacement therapy. Theoretically these cells can be directed towards any cellular phenotype in the body, and we are now aware of, and able to manipulate, several key developmental pathways required to differentiate pluripotent stem cells towards desirable neuronal phenotypes. Yet there are many aspects which mediate cell fate commitment that remain understudied including epigenetic components such as DNA methylation, which plays a vital role in mammalian development and is reflective of cellular phenotypes. Furthermore, there is now a body of evidence that demonstrates that the fate commitment of pluripotent stem cells is partially mediated by such epigenetic mechanisms. In this thesis, Huntington's Disease is used as the neurodegenerative model, and the effect of epigenetic variation is examined in pluripotent stem cells undergoing directed differentiation towards a striatal medium spiny neuron phenotype (the cells most affected in Huntington's disease).

1.1 Huntington's Disease

1.1.1 Historical Context and Clinical outline

In 1872, George Huntington published a detailed description of a form of Chorea unique for its late age of onset, strong heritability, and the sufferer's tendency for insanity: '*It is attended generally by all the symptoms of common chorea, only in an aggravated degree, hardly ever manifesting itself until adult or middle life, and then coming on gradually but surely, increasing by degrees, and often occupying years in its development, until the hapless sufferer is but a quivering wreck of his former self*.' This was the first detailed account of Huntington's disease (HD), a rare autosomal dominant neurodegenerative disorder that typically manifests between 30 to 40 years of age.

Disease onset is insidious and in the few years prior to formal clinical diagnosis, a gene positive individual may exhibit signs of cognitive and behavioural disturbances including disinhibition, apathy and irritability. Disease onset is formally diagnosed with the occurrence of clear motor symptoms, often including chorea, and eventually the condition progresses into a more pervasive array of cognitive, motor, and psychiatric symptoms, most notably impairment of executive functions, anxiety and depression. Over time these symptoms progress, and bradykinesia and rigidity, which usually co-exist with chorea from the earliest stages, often become more prominent in the later stages. Once diagnosed, individuals with HD have a life expectancy of approximately 20 -30 years. Pneumonia, heart disease and suicide are all more common in HD than the general population, and are among the most common causes of death (Craufurd and Snowden, 2002; Walker, 2007; Ross and Tabrizi, 2011; Bates, Tabrizi and Jones, 2014).

In 1983, the gene was localised to the short arm of chromosome 4 (Gusella *et al.*, 1983), but it wasn't until a decade later that the Huntingtin gene (HTT) was identified and the precise genetic aberration determined (MacDonald, *et al.*, 1993). Specifically, it was found HD occurs in individuals with expansion of the cytosine-adenine-guanine (CAG) repeat at the 5' end of the HTT gene, and that this is translated as an expanded polyglutamine stretch of the HTT protein (mHTT). The severity of the disease and age of onset have been found to inversely correlate with the length of this trinucleotide expansion (Andrew *et al.*, 1993). The typical length in healthy individuals is around 17-21 CAG repeats, however an individual can have up to 34 CAG repeats without any detrimental effects (Wexler et al., 2004). 35-39 CAG repeats are associated with incomplete penetrance but repeats of 40 or over are considered fully penetrant. Longer repeats, especially over 60 are associated with early onset and often a more rapid course. The CAG repeat is unstable, and can expand between generations (known as

anticipation); usually with paternal inheritance, and typically resulting in earlier disease onset than the parent.

Due to its late onset and heritable nature, the prevalence of HD varies drastically from country to country, as such between 1-12 cases of HD are estimated to occur for in every 100,000 people depending on location (Bates, Tabrizi and Jones, 2014). It is estimated that up to 10% of all cases of HD are spontaneous expansions. To date, there are no disease modifying interventions available for individuals with HD, and current drug treatments instead focus on temporarily relieving HD associated symptoms, but are limited in their effectiveness and are often associated with adverse side effects (Ross and Tabrizi, 2011). Furthermore, treatment of symptoms does not improve the overall prognosis or life expectancy of individuals who receive them compared to those who are untreated (Walker, 2007). A variety of disease modifying interventions are being actively explored, for example, a recently published phase 1/2a clinical trial exploring the use of antisense oligonucleotides has demonstrated dose-related reduction of mHTT in the cerebrospinal fluid, suggesting that it is indeed able to lower mHTT in the brain (Tabrizi *et al.*, 2018, Tabrizi *et al.*, 2019), although proof of a positive clinical effect will have to await further studies. The advancement of this trial is in large part due to the progression of our understanding of the mechanisms that underpin HD.

1.1.2 Clinical pathology and disease mechanisms of HD

HTT is expressed ubiquitously, although it is more highly expressed in the brain and testes than elsewhere (Cattaneo *et al.*, 2005). Its biological function is thought to involve regulation of gene transcription, vesicular transport and RNA trafficking, although its precise functional mechanisms are not well defined (DiFiglia *et al.*, 1995; Zuccato *et al.*, 2001; Sadri-Vakili and Cha, 2006; Savas *et al.*, 2010). Yet in spite of this wide spread expression, specific regions seem to be more susceptible to mHTT expression than others. Whilst there is considerable neurological loss of cortical grey matter (approximately 23% less volume than healthy controls), the most striking neurological loss associated with the disease is of the putamen and caudate nucleus or neostriatum (striatum), which in HD brains are observed to reduce in volume by up to 64% compared to healthy brains (Monte, Vonsattel, and Richardson, 1988; Halliday *et al.*, 1998). The volume loss within the striatum is primarily caused by a selective loss of medium spiny neurons (MSN), which comprise approximately 75% of the human striatum, and this specific neurological loss produces the motor and cognitive symptoms of HD. In the early stages of the disease there is greater loss of MSNs within the indirect

pathway of the basal ganglia (66% less MSNs than healthy controls), than those MSNs associated with the direct pathway (29% less), but by later stages there are considerably fewer MSNs in both the indirect and direct pathways (<10% compared to controls). Loss of MSNs from the indirect pathway is thought to be primarily associated with choreatic movements observed earlier in the disease, and loss of MSNs from the direct pathway to result in bradykinesia and rigidity observed in the later stages (Deng *et al.*, 2004, Starr *et al.*, 2008).

Although mHTT has been determined as the underlying cause of HD, and interventions that target it are in development, exactly how this mutation brings about HD pathology is not yet fully understood. The inverse correlation between CAG length and age of onset suggests that the mutant gene product exerts its effects predominantly through a gain of function, rather than loss of function mechanisms. This is further supported by the observation that although complete knock out of HTT is embryo lethal in rodents (Nasir *et al.*, 1995), yet a single copy of the mHTT gene can provide rescue and allow normal development to occur, suggesting that the mutant protein retains most of its normal functions and that its determinantal effects are due to a gain of toxic function. Similarly, in humans, individuals homogeneous for the mutant huntingtin gene appear to develop normally (Wexler *et al.*, 1987; Myers *et al.*, 1989) and the condition is not dramatically more severe than in heterozygous individuals. However, there is still some evidence for a degree of loss of gene product function. For example, BDNF has been shown to be essential for the differentiation and survival of striatal MSNs (Mizuno, Carnahan and Nawa, 1994; Ventimiglia *et al.*, 1995), and in these cells BDNF is upregulated by HTT but not by mHTT (Zuccato *et al.*, 2001), which demonstrates the focal loss of MSNs could be in part due to a loss of BDNF regulated activity.

Considering the advancing state of preventative interventions such as the antisense oligonucleotide clinical trial outlined above, a complete understanding of how the polyglutamine expansion of mHTT alters physiological functions and ultimately causes HD are imperative. However, it remains to be seen if such disease modifying treatments can ever completely prevent HD onset. Furthermore, even if preventative treatments become available, reparative interventions are likely to be required for those diagnosed after neural atrophy has taken place which is especially likely in sufferers with spontaneous mutations. One reparative strategy is to restore lost function by replenishing the ablated MSN population (Introduction 1.3.1; Peschanski, Cesaro and Hantraye, 1995; Clelland, Barker and Watts, 2008). To date, this has been explored largely by transplanting developing fetal striatal tissues (Introduction 1.3.2; Lelos *et al.*, 2016; Precious and Kelly, 2017).

1.2 Striatal Development

1.2.1 Early neural development

In mammalian embryogenesis, early neural fate commitment is largely dictated by gradient concentrations of critical morphogens along an anterior-posterior (AP) axis, and dorsal-ventral (DV) axis (Bronner-Fraser and Fraser, 1997; Tabata, 2001; Darnell and Gilbert, 2017). This begins during gastrulation with the formation of the primitive streak that defines the AP axis and leads to the induction of the three germ layers (endoderm, mesoderm and ectoderm; Leptin, 2005). Shortly following this, the notochord forms across the AP axis on the dorsal side of the mesoderm layer, and releases key morphogens including NOGGIN and CHORDIN which act to inhibit BMP and TGFβ signalling (Lamb and Harland, 1995; Rubenstein *et al.*, 1998; Anderson *et al.*, 2002). This drives the adjacent mesoderm to form the somites, and the adjacent ectoderm layer to begin neurulation. During neurulation the ectoderm forms the neural plate, which then folds towards the notochord to form the neural tube, and separates from the remaining ectoderm (Darnell and Gilbert, 2017). At this point, the neural tube is positioned between the notochord, epidermis, and the somites, which together coordinate regional patterning across the DV axis of the developing neural tube. The notochord releases the morphogen Sonic Hedgehog (SHH), which patterns cells within the ventral neural tube and forms the ventral plate (Rubenstein *et al.*, 1998). BMP4 is released from the epidermis, resulting in the formation of the roof plate and time sensitive cascades of TGFβ signalling on the dorsal side of the neural tube (Lee *et al.*, 2000; Tabata, 2001). Lastly, retinoic acid (RA) is released from the surrounding somites and patterns part of the DV axis morphogen gradient (Pierani *et al.*, 1999). At the most rostral point of the neural tube the anterior neural ridge (ANR) develops, which, through the expression of fibroblast growth factor 8 (FGF8) and the Wingless (WNT) pathway antagonist Dickkopf1 (DKK1), mediates anterior patterning of the neural tube (Gunhaga *et al.*, 2003). This induces signalling that results in the formation of three primary vesicles that will later develop into the encephalon (Tabata, 2001; Darnell and Gilbert, 2017). Of these vesicles, the most anterior/rostral of these is the prosencephalon, which can be further subdivided along the AP axis into the telencephalon (anterior) and diencephalon (posterior).

1.2.2 Development of the Fetal Striatum

The dorsal telencephalon (named the pallium) gives rise to the neocortex, whereas the ventral telencephalon (named the subpallium) contributes primarily to the various structures of the basal ganglia, but through migratory streams also contributes

to the neocortex. As the subpallium develops, two intra-ventricular bulges form, the lateral and medial ganglionic eminences (LGE & MGE, respectively). The LGE extends further along the caudal axis than the MGE, and its caudal tail end is referred to as the caudal ganglionic eminence (CGE). Collectively, these structures are known as the whole ganglionic eminence (WGE). In human, the MGE arises one day before the LGE (respectively during Carnegie stages 14 and 15; Freeman *et al.*, 1995). Its progeny contribute interneurons to the caudate and putamen (collectively the striatum), but are also understood to contribute to the cortex, globus pallidus, amygdaloid body, and olfactory bulb (Anderson *et al.*, 1997; Campbell *et al.*,1995; Olsson, Björklund and Campbell, 1998; Parnavelas, 2000). The ventral LGE (vLGE) produces neuronal populations that later form the MSNs of the striatum, whereas the dorsal LGE (dLGE) contributes primarily to the olfactory bulb and cortex (Deacon *et al.*, 1994; Olsson *et al.*, 1995; Stenman *et al.*, 2003). Finally, the CGE contributes calretinin positive interneurons that migrate to both the striatum and cortex (Brazel *et al.* 2003; Pilz et al., 2013; Pauly et al., 2013).

These substructure divisions are thought to be mediated initially by the morphogens that influenced the initial formation of the telencephalon, and later by more precise and regulatory transcription factors that induce gene cascades to result in highly specific cellular subtypes. Evidence for this comes from studies of the first gene distinctions that separate the dorsal and ventral regions of the telencephalon. For example, the pallium is associated with expression of genes such as PAX6, NGN2, and EMX1/2, whereas the subpallium is instead more associated with genes including GSX2, DLX1/2, MASH1, OLIG2 and NKX2.1 (Straccia *et al.*, 2015). These genes are directly mediated by the morphogens that defined early patterning of the neural tube, such as SHH and WNT pathways.

As previously mentioned, SHH is associated with ventral patterning of the neural tube, and its expression continues later into telencephalon development. Together with GLI3, SHH continues to form a DV gradient that mediates DV patterning in early telencephalon fate decisions. Rallu *et al.*, 2002 demonstrated that knockout mouse models of SHH develop a sparsely populated and abnormal ventral telencephalon with reduced expression of entopic subpallium markers including DLX2 and GSX2, and increased ectopic expression of pallium marker PAX6 in the subpallium. Conversely, knockout mouse models of GLI3 exhibit the reverse: a disformed pallium with reduced entopic PAX6 expression, and increased ectopic expression of DLX2 and GSX2 in the pallium. Similarly, WNT signalling also plays a key role in DV patterning, and as mentioned WNT inhibition through DKK1 expression is an important anterior marker that is understood to play a role in the telencephalon vesicle formation. Backman *et al.*,

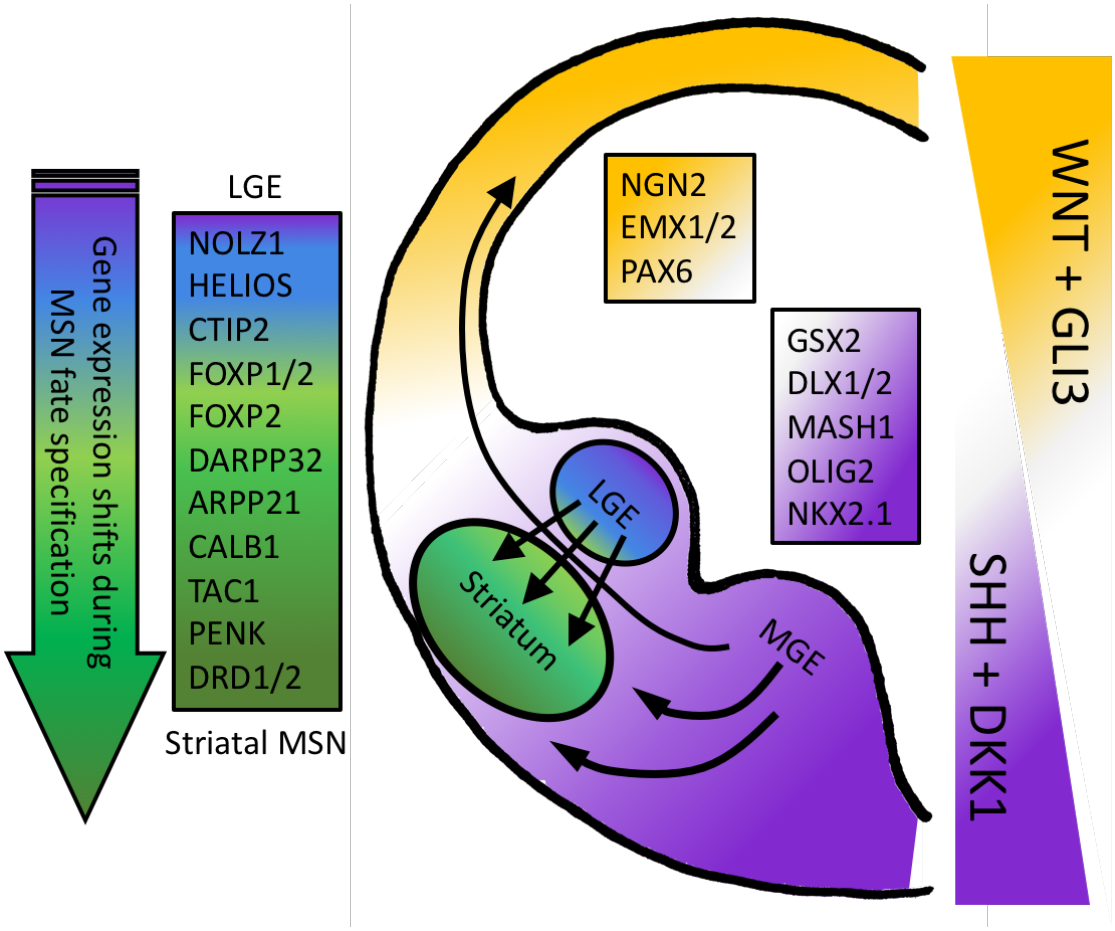


Figure 1.1 The developing telencephalon

The early neural tube is divided into dorsal (yellow) and ventral (purple) regions by differential expression of genes regulating the SHH and canonical WNT pathways. Following vesicle formation, the dorsal region is identified by its expression of PAX6, NGN2, and EMX1/2, whereas the ventral region expresses GSX2, DLX1/2, MASH1, OLIG2, and NKX2.1. This gradient gives rise to distinct regions, of particular interest here: the LGE. The LGE exhibits weak expression of NKX2.1 and PAX6 due to its regional positioning between the dorsal and ventral planes. As the LGE develops, it begins expressing genes such as NOLZ1 and HELIOS, and these cells begin migrating from the proliferative subventricular zone of the LGE, towards the developing Striatum, where they express a combination of early MSN markers such as CTIP2 and FOXP1/2. As these cells terminally differentiate into MSN they express mature markers, including DARPP32, ARPP21, CALB1, TAC1, PENK and DRD1/2.

2005 demonstrated that the inhibition of canonical WNT in the developing mouse pallium lead to the deformation of the pallium, reduced entopic expression of NGN2, EMX1, and EMX2, and caused ectopic expression of GSX2, DLX2, and MASH1 in the pallium. In contrast, increased expression of canonical WNT in the subpallium reduced the entopic expression of GSX2, DLX2, MASH1, OLIG2 and NKX2.1, and caused ectopic expression of PAX6 and NGN2 in the subpallium. As such, the combined increase in SHH and inhibition of the canonical WNT pathway is required to define early subpallium fate. However, these morphogens work in a gradient manner, as evidenced by interactions between these early telencephalon regional markers. For example,

PAX6 is downregulated by GSX2 (Toresson, Potter and Campbell, 2000), and as such is more highly expressed in the developing cortex where GSX2 is not expressed. In the LGE, low levels of PAX6 expression are identifiable, and none is present in the MGE. The reverse is true of the ventral marker NKX2.1, which is highly expressed in the MGE, expressed at a lower level in the LGE, and not present in the developing cortex, emphasising the gradient from dorsal to ventral positions (Gulacsi and Anderson, 2006; Hansen *et al.*, 2013; Pauly *et al.*, 2013; Onorati *et al.*, 2014).

As the LGE develops, two primary populations of cells arise, divided along the DV axis into the dorsal and ventral sides of the LGE (dLGE and vLGE respectively). Whilst GSX2 and DLX1/2 are expressed by both of these populations, differential expression of ISL1, SP8, and ER81 defines these separate regions. The dLGE contains cells strongly expressing SP8 and ER81, which later contribute primarily to the olfactory bulb and cortex, whereas the vLGE that goes on to contribute to the striatum, expresses ISL1 and weak levels of SP8 (Stenman *et al.*, 2003; Ma *et al.*, 2013; Xu *et al.*, 2018). Then, unique to the vLGE, downstream effects of GSX2 and DLX1/2 induce expression of ZNF-503 (NOLZ1) and IKZF2 (HELIOS) both of which are involved in cell cycle exit and the initiation of neuronal differentiation, indicating they are amongst the earliest genes that can be used to identify LGE cells patterned towards an MSN fate (Jensen, Björklund and Parmar, 2004; Urbán *et al.*, 2010; Martín-Ibáñez *et al.*, 2012).

1.2.3 The MSN Phenotype and Function

Eventually GSX2 and DLX1/2 become downregulated, as do NOLZ1 and HELIOS. Instead vLGE cells which are committed to a striatal MSN phenotype can be identified by a plethora of genes including BCL11B (CTIP2), FOXP1, and FOXP2, the expression of which are maintained into a mature phenotype and are required for normal MSN function (Arlotta *et al.*, 2008; Martín-Ibáñez *et al.*, 2012; Pauly *et al.*, 2013; Precious *et al.*, 2016). It is worth noting that identification of progenitor MSNs requires the colocalization of these genes, as each also occurs independently in the cortex (Onorati *et al.*, 2014). Over time, these MSN progenitors terminally differentiate into mature MSN phenotypes. MSNs are GABAergic inhibitory neurons, so named because they have a medium sized cell body, with projecting spiny dendrites. MSNs are typically identified by their near ubiquitous expression of of dopamine- and cAMP-regulated neuronal phosphoprotein (DARPP32, also PPP1R1B) and CTIP2 (Arlotta *et al.*, 2008; Arber *et al.*, 2015), and the colocalization of these two genes is considered the gold standard for the accurate identification of MSNs. Other mature MSN markers are not

expressed as uniformly, but include CALB1 (Calbindin1) and ARPP21 (Ivkovic and Ehrlich, 1999; Straccia *et al.*, 2015).

Importantly, whilst MSNs are morphologically indistinguishable, they are not a homogeneous cell population. There exists various methods of defining subtypes of MSN, the foremost of which defines at least two subtypes divided by their unique expression of the dopamine receptors D1 (encoded by the gene DRD1) or D2 (encoded by the gene DRD2), though there may also be a small proportion of MSNs which express both receptors. DRD1 positive MSNs (D1-MSN) also express TAC1 (Substance P), and can be further subdivided by gradient expression of PCHD8 and FOXP1; whereas DRD2 positive MSNs (D2-MSN) express PENK (Proenkephalin) and can be further subdivided by gradient expression of HTR8 and SYNPR (Ferré *et al.*, 2010; Straccia *et al.*, 2015; Gokce *et al.*, 2016; Xu *et al.*, 2018). These differences are understood to have functional consequences on these discreet MSN subtypes, primarily driven by the manner these two dopamine receptors respond to dopamine: D1-MSNs, which are excited by dopamine, are typically more associated with the direct pathway of the striatum, which acts to inhibit the neurons of the interior globus palidus and substantia nigra pars reticulata and initiate movement; conversely, D2-MSNs, which are instead inhibited by dopamine, are primarily associated with the indirect pathway of the striatum, which acts to inhibit the exterior globus palidus and supress movement (Albin *et al.*, 1989; Delong, 1990; Valjent *et al.*, 2009; Kravitz & Kreitzer, 2011). MSNs can also be divided by their position within the striatum, and there is evidence that MSNs of the dorsolateral and ventralmedial striatum are responsible for different neural functions (Doherty *et al.*, 2004; Yin *et al.*, 2004; Atallah *et al.*, 2007). Finally, the MSNs of the striatum can also be divided into striatal compartments, composed of striosomes (or patch) embedded within striatal matrix (Gerfen, 1984). MSNs of both the striosomes and matrix have been identified as projecting to key neural structures associated with the sensorimotor functions of the direct and indirect pathways. However, only MSNs of the striosomes also project to the substantia nigra pars compacta, which is the primary source of dopamine within the striatum, and thus these MSNs may be more involved in wider striatal function through feedback loop regulation of dopamine expression (Jimenez-Castellanos & Graybiel, 1989; Gerfen, 1989; Eblen & Graybiel, 1995; Kincaid and Wilson, 1996; Tokuno et al., 2002; Levesque & Parent, 2005; Fujiyama et al., 2011).

1.3 Cell Replacement Therapy

1.3.1 Background and Current Progress

One potential therapeutic approach for conditions that result in neurological loss is cell replacement therapy (CRT). CRT seeks to restore function by replenishing cell populations that have succumbed to a patient's disease, and thus could be offered to patients after disease onset and following neural cell loss (Peschanski, Cesaro and Hantraye, 1995; Clelland, Barker and Watts, 2008). CRT was suggested as a potential clinical therapeutic for sufferers of Parkinson's Disease (PD) following a series of landmark papers that explored the effect of neural transplants in rats in which substantia nigra (SN) dopaminergic neurons were ablated with 6-hydroxydopamine (6-OHDA). Incredibly, it was found that the motor deficits caused by this lesion could be reversed by direct engraftment of developing fetal rat ventral mesencephalon (VM) into the striatum, which is the projection target of the SN. Cells from grafted tissues innervated the neural tissues surrounding the transplant site, and later experiments demonstrated that when the grafted tissue was surgically removed, the restored function was once again lost (Björklund and Stenevi, 1979; Perlow *et al.*, 1979; Björklund *et al.* 1980).

Later, the theory behind this work was applied to other neurodegenerative diseases including HD. Similarly, recovery has been observed following transplantation of fetal WGE (which gives rise to striatal tissues; Introduction 1.2.2) into the quinolinic acid (QA) lesioned rat striatum (Dunnett and Björklund, 1994). QA is potent neurotoxin which causes apoptosis of cells with NMDA receptors, and as such, when it is directly administered to the striatum it causes selective ablation of MSNs, but other striatal cells such as interneurons are preserved (Schwarcz, Whetsell and Mangano, 1983; Schwarcz and Kohler, 1983). This effectively mimics the selective MSN loss observed in HD pathology, which results in motor and cognitive deficits in rodent models (Beal *et al.*, 1986; Guillemin, 2012). WGE derived allografts in the QA lesioned rat striatum result in integrated grafts rich for striatal tissue phenotypes that can restore both motor and cognitive functions resulting from the QA lesion (Dunnett and Björklund, 1994; Nakao *et al.*, 1996; Fricker, *et al.*, 1997; Watts *et al.*, 2000; Dunnett and White, 2006). Furthermore, there is evidence that the transplantation of human WGE into the QA lesioned adult rat striatum can also form grafts that exhibit striatal morphology, with projections to key striatal targets (Grasbon-Frodl *et al.*, 1996; Pundt *et al.*, 1996b; Grasbon-Frodl *et al.*, 1997). These grafts were able to induce functional recovery (Pundt *et al.*, 1996a; Sanberg *et al.*, 1997) comparable to that produced by transplanting rat WGE (Lelos *et al.*, 2016). Thus, human to rodent xenografts are suitable for pre-clinical research advancing CRT in HD.

This body of work in rodent paved the way for transplantation of human fetal WGE in several small human safety and feasibility studies (Rosser and Bachoud-Levi, 2012; Cisbani and Cicchetti, 2014). Critically, assessments following these trials suggested no acceleration of HD disease progression (Rosser *et al.*, 2002; Bachoud-Levi *et al.*, 2006), and only rare instances of graft overgrowth (Keene, *et al.*, 2009). To date evidence of clinical benefit has been mixed, although this is perhaps not surprising as no studies to date have been adequately powered to assess efficacy. Furthermore, many of the studies have used different methods of tissue selection, dissection, tissue handling, and engraftment, which will have reduced the likelihood of observing improvements (Freeman *et al.*, 2011), as it is well know that the age, quality, and content of grafted tissue plays a critical role in transplantation outcome in rodent models (Fricker, *et al.*, 1997; Watts *et al.*, 2000; Zietlow *et al.*, 2005; Anderson *et al.*, 2007; Zietlow *et al.*, 2012). A small number of studies have reported positive outcomes of human fetal WGE transplantation. For example, Hauser *et al.*, 2002 detailed the outcomes of three such transplants, including histological analysis of post-mortem in one subject. They determined that fetal tissue survived transplantation, exhibited typical morphology of the developing striatum, and was innervated by host derived dopaminergic fibers, however observed no clinical functional recovery. A second study detailed the transplantation of fetal tissues into 5 patients, and reported convincing clinical evidence of functional stability following engraftment, with signs of increased metabolic function in the striatum (Bachoud-Levi *et al.*, 2000). However, long term assessment of these patients revealed that the benefits of these grafts faded overtime (Bachoud-Levi *et al.*, 2006). Similar findings of improved function and metabolic activity were observed in another trial, but again these benefits faded after a number of years (Gallina *et al.*, 2008; Gallina *et al.*, 2010; Gallina *et al.*, 2014). Collectively, these few studies indicate proof of principle that human WGE transplants can improve function in HD. However, it is also clear that there remains much to learn about the processes that facilitate this recovery, and significant basic research is required to continue progressing cell replacement therapy mediated brain repair.

1.3.2 The Future of CRT

Work up to this point has clearly defined the potential of CRT for the treatment of neurodegenerative diseases, but has also demonstrated that significant further work is required to bring this treatment to the clinic. Future advancement requires a good model of neurodegeneration in which to understand and optimise CRT, and HD has many advantages as a model in this respect (Rosser and Dunnett, 2003). First, although CRT in PD is more advanced in terms of clinical application of CRT than HD,

it can be argued that PD is a less useful model for learning how to best achieve circuit repair, something that will be necessary for repair of most neurodegenerative diseases. This is because in PD, normal circuitry is not presently restored as donor cells are transplanted ectopically into the primary region the SN projects to, the striatum (Boronat-García, Guerra-Crespo and Drucker-Colín, 2017). By contrast, in HD, the donor cells are effective when transplanted into their entopic position within the striatum, thereby facilitating the restoration of the normal anatomical circuitry. Additionally, because the cause of HD is known, it is easier to create useful preclinical animal models, including in large brained mammals and primates, essential for future transition (Morton, 2018; Snyder and Chan, 2018). With regards to early clinical human trials there are also advantages for HD over PD that result from its autosomal dominant, monogenic, and highly heritable nature. This allows for easy identification of high-risk populations prior to disease onset and the predictability of penetrance by expansion length, which together allow for more refined human trials of CRT (Rosser and Dunnett, 2003).

However, whilst fetal WGE is at present the gold-standard cell source for CRT for HD, fetal tissue is not suitable for widespread clinical application of CRT for any disease for a multitude of reasons. Firstly, multiple donors are likely required to provide enough tissue to transplant into each hemisphere of the adult striatum, yet at present, there are no means by which to store or expand fetal tissue whilst retaining its potency as a transplantable cell source (Zietlow *et al.*, 2005; Anderson *et al.*, 2007; Zietlow, *et al.*, 2012). Subsequently, to meet tissue demands multiple donors at the correct developmental age would be required to become available in the same brief period of time and in the same geographical area, which is challenging due to the relative scarcity of these tissues. The alternative of undertaking repeat surgeries for each patient is not a practical solution due to the significant burden this places on the patient. Furthermore, tissue age and quality between donors is highly variable and thus their use would require extensive quality control and validation measures to be conducted on each donor used, that require a significant investment of time which as mentioned above, is not available for fetal tissues as there are presently no long term storage or expansion options. Finally, there are also ethical issues with the use of fetal tissues, which are exacerbated by the potential requirement of multiple donors for each surgical procedure. Therefore, alternative sources must be considered (Precious and Kelly, 2017). To date, the most promising and widely researched of these are pluripotent stem cells, in particular embryonic stem cells.

1.4 Pluripotent Stem Cells in CRT

1.4.1 Pluripotent Stem Cells

Pluripotent stem cells (PSC) are capable of forming all three germ layers of the developing embryo (mesoderm, endoderm, ectoderm), and therefore can theoretically become any cell in the adult body (Niwa, 2007; Mahla, 2016). There is only a brief window during embryogenesis where mammalian cells are considered pluripotent: during the blastocyst stage, as the inner cell mass (ICM) divides in preparation for gastrulation. These cells are known as embryonic stem cells (ESC), and were first isolated and cultured from the mouse blastocyst in 1981 (Evans and Kaufman, 1981; Martin, 1981). Upon reintroduction into a blastocyst, mouse ESCs can contribute to the formation of chimeric animals, indicating their authentic capacity for complete mammalian development (Bradley *et al.*, 1984; Beddington and Robertson, 1989).

Nearly two decades after the isolation of mouse ESCs, human embryonic stem cells (hESC) were also successfully isolated and cultured (Thomson *et al.*, 1998), which offered unparalleled potential to model and study the mechanisms underpinning human development, the so-called “disease in the dish”, and as a possible new donor cell source for regenerative medicines including CRT. hESCs can be maintained in a pluripotent state indefinitely, which is possible by manipulating an inherent cycle of pluripotency that is regulated by the genes NANOG, OCT4, and SOX2, and by preventing exposure to critical environmental cues that normally trigger differentiation (Boyer *et al.*, 2005; Vallier *et al.*, 2005; Masui *et al.*, 2007, Fong *et al.*, 2008). Whilst they share much of the same functional traits as mouse ESCs, hESCs demonstrate different morphological traits to those seen in mouse, more indicative of ESCs derived from other primates, and the later isolated mouse epiblast stem cells (Brons *et al.*, 2007). Specifically, they exhibited a high nucleus to cytoplasm ratio, with prominent nucleoli, and formed flat uniform colonies with tightly packed, but individually distinct, cells (Thomson *et al.*, 1995; Thomson *et al.*, 1998).

More recently a new form of PSC was discovered, the induced pluripotent stem cell (iPSC). This was achieved originally for mouse cells, by Takahashi and Yamanaka in 2006, who then one year later went on to demonstrate that the same principles could be applied to generate iPSCs from human tissue (hiPSCs) (Takahashi and Yamanaka, 2006; Takahashi *et al.*, 2007). Correctly reprogrammed iPSCs are remarkably similar to ESCs, in that they share all of their defining features, including an unlimited capacity for self-renewal and the ability to differentiate towards all three germ layers (and in mouse to contribute to chimera formation), yet they can theoretically be derived from any cell in the body (Shi *et al.*, 2017). In the first of their landmark papers, Takahashi

and Yamanaka determined that by inducing expression of 24 genes found to be expressed in pluripotent cells, somatic cells could undergo a phenotypic shift to become similar to ESCs. Successive rounds with the systematic removal of candidate genes led to the discovery that this induced form of pluripotency required the inclusion of four genes: OCT4, SOX2, KLF4, and C-MYC, collectively termed the Yamanaka factors or OSKM. The resulting iPSCs expressed critical pluripotency genes including NANOG, formed teratomas when injected subcutaneously into immunocompromised mice, and contributed to different tissues of developing embryos upon blastocyst injection. However, these first iPSCs were not equivalent to ESCs as they were described as unable to produce viable chimeras. Since then many labs have repeated this work, and further refinement of selection processes has allowed for the identification of fully reprogrammed iPSCs which are capable of all features associated with ESCs (Okita, Ichisaka and Yamanaka, 2007; Maherali *et al.*, 2007; Wernig *et al.*, 2007), including contributing to animals formed entirely of iPSCs (Boland *et al.*, 2009).

In the years since their discovery, iPSCs have captured the imagination of many researchers for various reasons. For instance, being able to generate them from any somatic cell means they offer a unique opportunity to study disease in the dish by generating them directly from tissue samples collected from disease sufferers (Park *et al.*, 2008) Additionally, they are an intriguing option as a potentially autologous cell source for regenerative medicines, as iPSCs could be created from a patient's own cells, thereby overcoming issues with immune rejection (Nishikawa, Goldstein and Nierras, 2008). As such, there have been interesting advancements in the repertoire of methods that can be used to deliver OSKM factors to cells for reprogramming, which have now expanded from the initial genome integrating retrovirus to include non-integrating and non-viral options (Malik and Roe, 2013). Furthermore, new combinations of factors to induce pluripotency have also been uncovered (Yu *et al.*, 2007; Maekawa *et al.*, 2011). We now understand significantly more about the biological mechanisms that underpin the induction of pluripotency brought upon by the Yamanaka factors (Polo *et al.*, 2012). For example, there is a requirement to reverse chromatin modifying mechanisms that define non-pluripotent cell states, such as DNA methylation which needs to be removed at specific developmental gene promotors and returned to the vast majority of the genome (Papp and Plath, 2013). It is through the study of the mechanisms of DNA methylation in iPSC induction that one of the critical differences between iPSCs and ESCs was uncovered. Specifically, that early passage iPSCs have been shown to retain some of their previously held chromatin structures, including tissue specific DNA methylation which can alter their differentiation abilities (Kim *et al.*, 2010; Polo *et al.*, 2010); this difference is discussed in more detail below (Introduction 1.5.4). Further comparison of iPSCs and ESCs has been conducted to try

to establish precisely how similar these types of pluripotent cells are. Initially it was found that there were gene expression differences between these cell types which was understood to indicate that iPSCs were not as equivalently pluripotent as ESCs (Chin *et al.*, 2009). However whether this is a true effect remains to be determined, as there is also evidence that PSCs express different genes depending on the lab conditions under which they are cultured (Newman and Cooper, 2010); genetic differences could also be the cause of this variability (Kyttälä *et al.*, 2016); and the previously identified differences in DNA methylation have also been attributed to this gene expression difference (Kim *et al.*, 2010; Polo *et al.*, 2010). Nonetheless, these differences have not prevented the progress of iPSCs for clinical trials in regenerative medicines, with the first efficacy and safety trials already underway. For instance, in Japan the first transplantation of autologous hiPSCs differentiated into retinal pigment epithelial cells to treat age related macular degeneration has successfully taken place (Mandai *et al.*, 2017).

1.4.2 hPSCs as a cell source in CRT

As previously discussed, fetal tissues are not suitable for widespread clinical application because of unavoidable issues with logistical limitations, ethical concerns, and inconsistent developmental age and quality between donors (Introduction 1.3.2). By comparison, hPSCs have the potential to overcome these issues. First, as hPSCs have an unlimited capacity for self-renewal, they can be maintained indefinitely before being prepared for transplantation, thereby overcoming the difficulty of sourcing sufficient numbers of cells. Whilst hESCs are still considered an ethically contentious source of tissue, they are comparatively less challenging because only one donor is required to generate entire populations of cells to transplant into multiple patients. Furthermore, hiPSCs are not derived from an embryo and therefore their use bares far fewer ethical concerns. Finally, hPSCs can theoretically overcome issues with developmental age and quality variation, as it is possible to have direct control over the developmental stage of these cells and to harvest them at the appropriate time best suited for their purpose. Furthermore, it would be possible for batches of cells to undergo quality assurance assays prior to clinical use, a practice that currently isn't possible with the limited supply of fetal tissues. Therefore, compared to human fetal tissues, hPSCs appear to be a superior choice for the future of CRT, and the only challenge preventing them from entry to clinic is the determination of the exact methodology required to direct hPSCs towards an authentic terminally differentiated phenotype (MSNs in the case of HD). This theoretically requires the understanding and

mimicking of the developmental processes that take place within the human embryo allowing for the development of MSNs in the normal striatum (Introduction 1.2).

1.4.3 Directed differentiation of hPSCs toward a neuronal phenotype

Much of the progress in differentiating hPSCs towards an MSN fate suitable for transplantation comes from our advancing understanding of the defining features of neural development. Early protocols relied on culturing PSCs as 3D embryoid bodies (EB), and allowing these EBs to undergo the initial stages of embryogenesis. Zhang *et al.*, 2001 demonstrated that exposure of hESC derived EBs to FGF2 induced neuralisation and the formation of neural rosettes, although this did not produce a pure population as cells of other lineages were also present. However, Zhang *et al.* also demonstrated that the neural rosettes could be isolated from these other lineage derivatives, and after doing so, the neural rosettes could be terminally differentiated towards various neural fates including neurons, astrocytes, and oligodendrocytes by removing FGF2 from culture. This provided clear evidence that hESCs could be used to produce neuronal populations *in vitro*. Yet, there is an unavoidable increase in cell population heterogeneity in 3D based cultures, which variably reduces the consistency of final cell yields. Isolating a pure neuronal population is required, but challenging (Zhang *et al.* reported 96% of cells were neuronal following isolation methods), and the location of cells within an EB is critical, since cells on the surface and inside an EB are exposed to varying gradients of patterning factors, which is well established to play a significant role in regional specification (Bronner-Fraser and Fraser, 1997; Tabata, 2001; Darnell and Gilbert, 2017). As such, it is challenging to control the final fate specification of such cultures, and it is therefore possible that 3D and EB based methods may not be suitable for CRT applications.

More recently, 2D monolayer cultures of hPSCs have become increasingly explored, as theoretically exposure to morphogen gradients can be more closely controlled across monolayer cell populations. Neural induction of 2D monolayer cultures was first reported in 2009 when Chambers *et al.* used a “dual SMAD inhibition” protocol to demonstrate that hPSCs cultured as a monolayer could be induced to a neural forebrain progenitor fate by inhibiting BMP and TGFβ signalling. This was originally achieved using NOGGIN and SB431542 (SB) in combination, although a year later it was demonstrated that LDN193189 (LDN) could successfully mimic NOGGIN for the purposes of neural induction, and Dorsomorphin could block additional BMP related developmental pathways (Boergermann et al., 2010). The dual SMAD inhibition protocol functions on multiple levels; first breaking the pluripotency gene cycle of PSCs

by rapidly downregulating NANOG, whilst simultaneously promoting neurulation and downregulating differentiation towards mesoderm or endoderm pathways (Chambers *et al.*, 2009, Boergermann et al., 2010). The resulting cells are positive for PAX6 indicating that they default to a dorsal forebrain phenotype (Chambers *et al*., 2009). However, and importantly in the context of this thesis, through exposure to additional regional morphogenic factors these progenitors can be directed to the various other phenotypes found across the telencephalon including cortical and striatal interneurons (Cambray *et al.*, 2012; Noakes *et al.*, 2019) and striatal MSNs (Delli Carri *et al.*, 2013; Nicoleau *et al.*, 2013; Arber *et al.*, 2015).

1.4.4 Directed differentiation of hPSCs toward an MSN phenotype

Once hPSCs have been patterned to a general neural progenitor phenotype, further signalling is required to continue regional specification towards an LGE phenotype. Initial success was achieved in 2008 by Aubry *et al.*, who used an EB based method to derive neural rosettes and treated them with AP and DV gradient morphogens to facilitate commitment to a ventral telencephalon fate. Specifically, they treated neural rosettes with SHH and DKK1 to ventralise their neuronal population, and combined this with BDNF to promote anterior specification. Whilst, their protocol did not produce pure neuronal populations (only 22% of the cells expressed the neuronal marker MAP2), the majority of these co-labelled with the MSN marker DARPP32 (53% of total MAP2 positive cells), supporting the use of these factors to induce an LGE progenitor fate. The use of SHH and DKK1 to induce an MSN phenotype has been further validated and improved upon since this report, and there are now several protocols that describe their use in combination with an EB based method to generate DARPP32 expressing neurons (Li *et al.*, 2009; Ma *et al.*, 2012). Of particular note, Ma *et al.*, 2012 reported a dose dependent response to SHH in hPSC derived forebrain progenitors, finding an increased dose of SHH (500 ng/ml) resulted in reduced GABA positive neuronal populations with minimal DARPP32 expression, compared to a reduced dose (200 ng/ml) which resulted in increased GABA positive neuronal populations, of which 89% were DARPP32 positive. They also demonstrated a dose dependent effect on the progenitor markers PAX6 (dorsal telencephalon) and NKX2.1 (ventral telencephalon), finding that increased concentrations of SHH directly reduced PAX6 expression, and increased NKX2.1 expression. However, as previously mentioned, EB based methods are undesirable for CRT application due to widespread heterogeneity, and with the advance of monolayer cultures additional methods have been explored.

Both Delli Carri *et al.*, 2013 and Nicoleau *et al.*, 2013 applied SHH and DKK1 based protocols to monolayer cultures of hPSCs that had been patterned towards a forebrain progenitor fate using dual SMAD inhibition as outlined by Chambers *et al*., 2009. After 80 days of maturation, Delli Carri *et al.* reported 51% of their cell population expressed MAP2a/b, of which 60% were positive for the MSN precursor marker CTIP2, and an estimated 20% were positive for DARPP32 (indicating approximately 10% of the total population were DARPP32 positive neurons). Nicoleau *et al.*, optimised SHH and WNT inhibition exposure through examination of subpallium markers in dose dependent conditions of SHH, DKK1 and the WNT antagonist XAV939 (XAV). Using optimised SHH and XAV dosages, they achieved an estimated 58% MAP2 positive population, of which an estimated 28% were DARPP32 positive. Despite some success using SHH and WNT inhibition to induce an LGE progeny cellular fate, optimised conditions have consistently failed to produce a homogeneous population of DARPP32 positive MSNs. It is possible this is due to an incomplete understanding of the developmental signals that result in LGE progenitors, and there is evidence that SHH may not actually be required for the specification of an LGE phenotype. For example, during optimisation experiments, Nicoleau *et al.* found that high levels of SHH resulted in reduced expression of the LGE progenitor marker CTIP2, corresponding to a relative increase in expression of the MGE marker NKX2.1, corroborating earlier findings by Ma *et al*.. Furthermore, inclusion of Cyclopamine, a SHH antagonist, did not affect the CTIP2 expressing population.

There is some evidence that TGFβ family signalling occurs in the developing subpallium (Feijen, Goumans and van den Eijnden-van Raaij, 1994; Maira et al., 2010), and Activin A (Activin), a TGFβ family protein, has been used to induce a CGE interneuron fate in hPSCs (Cambray *et al.*, 2012). In 2015, Arber *et al.* used dual SMAD inhibition to induce a neural forebrain phenotype in 5 hPSC lines. They then used Activin to rapidly induce an LGE phenotype in forebrain progenitor precursors, reporting that 50% of cells expressed CTIP2 after 9 days of treatment, compared to <4% in untreated controls, corresponding to a 40 fold relative increase in CTIP2 RNA expression. Furthermore, they exposed forebrain progenitors to either Activin or SHH, or both Activin and SHH, and compared these with untreated controls. They observed that the expression of the LGE progenitor markers CTIP2 and NOLZ1 were unaffected by the addition of SHH, and only increased following the introduction of Activin, whereas MGE progenitor markers NKX2.1 and LHX8 showed the opposite trend, only increasing in conditions that included SHH. Collectively this data suggests that Activin and SHH function in distinct pathways to elicit LGE and MGE phenotypes respectively, and that SHH is therefore not required to induce an LGE phenotype. It is therefore possible that previous success using SHH was due to the inhibition of dorsal fates rather than a direct

induction of an LGE phenotype. However, even under optimal conditions, the Activin based Arber *et al.* protocol did not produce a pure MSN population, finding a maximum of 50% DARPP32 positive neurons in hPSC derived cultures.

1.4.5 Transplantation of hPSC derived MSNs into Pre-Clinical HD Models

Deriving MSN progenitors from hPSCs suitable for CRT in HD requires good *in vitro* protocols that reliably produce an authentic MSN phenotype. However, *in vitro* examination of a cell population is not necessarily indicative of how they perform outside of defined culture conditions. Therefore, many of the studies outlined above have also examined how their cells perform *in vivo* by transplanting into pre-clinical models of HD. Aubry *et al.* transplanted unilaterally into the QA lesioned rat striatum. They used cells harvested at five time points across their MSN differentiation protocol, to determine at what developmental stage hPSC derived MSNs should be transplanted. The earliest of these resulted in teratoma formation, indicating that the hESCs had not yet sufficiently committed to a neural cellular fate to be suitable for transplantation. The latest time point also produced disappointing results, finding that the transplantation of mature, committed cells resulted in poor graft survival. They observed good graft survival and populations of DARPP32 positive neurons when transplanting progenitor cells at the mid-stage of their protocol, following long term exposure to SHH and DKK1. However, in a second, long-term study of the *in vivo* potential of cells cultured to the mid-stage of their protocol, graft overgrowth became apparent, which resulted in the deformation of the non-grafted, contralateral hemisphere (Aubry *et al*, 2008). Thus, it appears that the most suitable hPSCs for grafting will be populations that have committed to an MSN phenotype and are mature enough that they will not cause graft overgrowth, but not so mature that they lose their plasticity and viability following transplantation. Following refinement of their differentiation protocol, the same lab group transplanted MSN progenitors; Nicoleau *et al.* demonstrated the resulting grafts maintained their regional patterning for an MSN fate as they contained the MSN markers FOXP1, CTIP2, and DARPP32. Furthermore, there was evidence of innervation with host TH fibers, indicating the potential for functional interaction between the host tissue and graft, although no functional assessments were reported (Nicoleau *et al.*, 2013).

Both Ma *et al.* and Delli Carri *et al.* have reported a degree of functional recovery following transplantation of their cells into the QA lesioned striatum. First, Ma *et al.* transplanted hESC derived forebrain progenitors into the QA lesioned mouse striatum, and compared these animals to control populations which received either spinal GABA

progenitors or sham graft surgery. They observed a large proportion of GABA positive neurons in both graft groups, and a high proportion of DARPP32 positive neurons in the forebrain progenitor derived grafts. Furthermore, they observed significant recovery in animals receiving forebrain progenitors compared to both control groups at 16 weeks post-transplantation across a variety of motor tests. This supports the notion that hPSC derived DARPP32 positive neurons can restore function in models of HD, however at this time behaviour was not restored to the same functional levels observed prior to the lesion and transplant. It is possible a longer study could have allowed for the observation of further recovery, as animals were not maintained until a functional plateau was reached, and there is evidence from the transplantation of human fetal WGE that further recovery can be seen at 20 weeks post transplantation (Lelos *et al.*, 2016). However, it was also evident that these cells resulted in graft overgrowth. Reportedly, approximately 50,000 cells were transplanted into each animal, but at week 16 an average of 3,466,667 ± 648,931 cells were observed in forebrain progenitor derived grafts, equating to an estimated 70 fold increase in total cell population. Delli Carri *et al.*, also reported some functional recovery following transplantation. They transplanted bilaterally into rats with a unilateral QA lesion, and culled animals at three, six, and nine weeks post transplantation. Transplanted animals undergoing apomorphine induced rotations showed reduced rotations compared to lesioned controls. However, the inclusion of a graft into the un-lesioned hemisphere in the grafted groups undermined this observation, as the number of induced rotations occurs as a direct result of the interaction between the lesioned and intact hemispheres. Furthermore, post mortem histology revealed <1% DARPP32 positive cells across the resulting grafts at 9 weeks, suggesting that this behavioural change was not due to the introduction of DARPP32 expressing MSNs. Analysis of the resulting grafts indicated widespread expression of Ki67 and continued graft overgrowth.

Arber *et al.* transplanted hESC derived LGE progenitor cells following 9 days of treatment with Activin into the QA lesioned rat striatum. Unlike previous experiments, there was no graft overgrowth observed and Ki67 was not expressed within the engrafted cells at 16 weeks. Immunohistological analysis of the grafts indicated the expression of LGE and MSN markers FOXP2, calbindin, GABA, substance P and enkephalin, furthermore, 50% of the transplanted cells expressed DARPP32. They also reported graft innervation from host TH fibres, and host tissue innervation from graft derived NCAM fibres which extended from the graft towards the host globus pallidus and midbrain. Collectively, this seems to indicate that an Activin based differentiation protocol can successfully pattern hPSCs towards an LGE progenitor fate, that maintains this commitment *in vivo* and consistently gives rise to DARPP32 expressing populations. However, animals engrafted with these cells did not show any signs of

functional recovery at 16 weeks, although it is possible this was due to the very small graft sizes.

Whilst these studies each represent significant progress in closing the gap between hPSCs and clinical application, it is clear that the current methods of deriving MSNs from hPSCs are imperfect and require further refinement. Until we are able to reliably direct hPSCs towards a phenotype that produces a rich population of MSNs that restores function in pre-clinical models, these cells cannot be considered for clinical application. Therefore, the future of hPSCs as a cell source in CRT for HD depends on overcoming the various barriers that have thus far prevented their pre-clinical success.

1.4.6 Overcoming the barriers to clinical application of hPSCs in CRT for HD

There are various possibilities to explain the relatively poor functional improvements following engraftment of hPSC derived MSNs in animal models of HD to date. First, it has been proposed that the relative success of fetal tissue transplants over hPSC derived MSNs could in part be due to the heterogeneous cell population contained within WGE (LGE plus MGE) tissues (Reddington *et al.*, 2014; Precious and Kelly, 2016). It is well established that transplantation of LGE progenitors gives rise to rich populations of DARPP32 expressing graft neurons (Watts *et al.* 2000), whereas transplantation of age-matched fetal rat MGE does not, demonstrating the requirement of an LGE phenotype for DARPP32 positive populations. However, there is also evidence that transplantation of WGE (both LGE and MGE together) resulted in the largest volume of striatal neurons, surpassing that observed when the same numbers of LGE cells were transplanted (Watts *et al.* 2000), suggesting that the MGE derived interneurons could play a role supporting the LGE population during engraftment, potentially providing critical developmental signals and neurotrophic factors that enhance the engrafted LGE population. Thus, it is possible that a more diverse range of cells need to be generated from hPSCs to be transplanted together (LGE and MGE progeny), and work addressing this issue continues (Noakes *et al.*, 2019).

Another possibility is that PSC derived MSNs may not yet be authentic MSNs. Although MSN differentiation methods have improved substantially and reliably yield neurons that express known MSN markers, our own understanding of what defines a functional MSN is still somewhat limited (Introduction 1.2.3). As such, PSC derived MSNs may still lack some presently undefined feature of an authentic MSN, and thus be rendered unable to function in the same manner that authentic fetal tissue derived MSNs can. If so, this likely due to the methods currently used in MSN differentiation

protocols. For instance, whilst 2D monolayer cultures are desirable because of the unparalleled control over cell fate and potential to generate single phenotypes of cells, it must be recognised that this is unlike normal development which occurs in 3D space and in diverse heterogeneous populations (Duval *et al.*, 2017). And as previously mentioned, there exists evidence that heterogeneity in cell populations during development confers a beneficial effect on LGE derivatives (Watts *et al.* 2000) which may be a requirement for developing the full functionality of an authentic MSN. Furthermore, the majority of focus thus far has been on shaping hPSCs using widely expressed morphogens and the few known critical genes that induce the relevant cascades of genes required to form an MSN (e.g. SHH and Activin). Yet, the precise mechanisms that underpin MSN development remain unclear, and therefore our understanding of how to best guide hPSCs towards an MSN fate is currently limited. For example, key morphogenic factors considered critical to MSN differentiation are also fundamental in many other differentiations, with SHH known to play various patterning roles across the entire developing neural tube (Danesin & Soula, 2017), and late exposure of Activin on neuronal precursors is used in interneuron differentiations (Cambray *et al.*, 2012). As such, precise timing and dosage of these factors is considered fundamental to properly differentiate an MSN, which increases the challenge of generating pure MSN populations *in vitro* and will likely continue to do so until we fully understand MSN fate specification. It is also plausible that further late stage signalling refinement is necessary for enhanced MSN fate specificity, as until relatively recently, the role of Activin in the mediation of LGE regional fate specification was unknown (Arber *et al.*, 2015), and so there may exist other unused morphogens essential to MSN development. This is especially relevant considering the known diversity of MSN subtypes (Introduction 1.2.3), as it remains unclear what mechanisms drive this diversity, or even how much subtype specificity has to be defined in an MSN progenitor population prior to transplantation. Additionally, it is unknown whether earlier developmental processes (such as dual SMAD inhibition which results in high PAX6 expression, associated with more dorsal regions than the LGE) have a more restrictive impact on cellular potency than current protocols allow for, which prevent large subpopulations of cells from differentiating towards the desired MSN phenotype. Indeed, there is ongoing work examining the impact of early genes that define telencephalon regionality on later stage developmental cascades. For example, the roles of inducible GSX2 and EBF1 have been examined in a SHH and DKK1 based MSN differentiation, and there is evidence that forcing the expression of these two early developmental genes improves the purity of the resulting MSN populations both *in vitro* and *in vivo* (Faedo *et al.*, 2017).

Each of the above issues may drive a difference between PSC derived MSNs, and authentic MSN progenitors derived from primary fetal tissues which have undergone a complex and highly refined developmental process. But, there are various other features of development that are understood to play integral parts in cell fate specification that have not yet been considered. For example, the role of epigenetic mechanisms such as DNA methylation, which are known to undergo precise temporal and tissue specific changes during development (Cedar and Bergman, 2012), remain relatively unexplored, and to date there have been no attempts to compare the developing DNA methylome of authentic MSNs derived from fetal tissues, with those derived from PSCs. However, with the increasing availability of genome wide array technology, it is easier than ever to begin exploring such features, and subsequently to begin applying them to PSC based culture of MSNs.

1.5 The Relevance of the Epigenome to Development and Differentiation

1.5.1 Epigenetics and DNA methylation

The usage of the terms “Epigenetics” and “Epigenome” are a subject of debate within the scientific community, as rather than referring to a single measurable physical feature they instead refer to a concept, one which has undergone various revisions as our understanding of cell biology and the techniques we have to study the underlying mechanisms have advanced (Henikoff and Greally, 2016). For the purposes of this thesis, epigenetics will be considered by a more liberal definition proposed by Adrian Bird, referring to: “the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states” (Bird, 2007). This definition includes various mechanisms that act on a cells chromatin structure to alter gene expression and maintain cellular phenotype, including DNA methylation. DNA methylation refers to the addition of methyl groups to DNA nucleotides in order to change the local genetic activity without altering the underlying DNA sequence (Moore, Le and Fan, 2013). It is catalysed by a family of enzymes called DNA methyltransferases (Dnmts), which methylate DNA both during replication and in a *de novo* manner. Dnmt1 plays a critical role during cell mitosis, acting on the newly synthesised daughter strand of DNA, to replicate the methylation status of the parental DNA strand, thereby preserving the DNA methylation status of the parent cell (Bestor 1992; Yoder *et al.*, 1996; Pradhan *et al.*, 1999). Conversely, Dnmt3a and Dnmt3b function as *de novo* methylators, as they are used to methylate otherwise unmethylated DNA molecules when required (Okano *et al.*, 1998, 1999; Hsieh 1999).

Two of the four bases that comprise DNA can by methylated; cytosine and adenine, which when methylated become the *hyper*methylated 5-methylcytosine and N^6-methyladenosine respectively (when not methylated they are instead described as *hypo*methylated). The role of differentially methylated adenine in mammals remains unclear, although recently there has been an increase in research interest (Wu, *et al.*, 2016; Iyer, Zhang and Aravind, 2016). By comparison, methylated cytosine as part of a CpG dinucleotide pairing is one of the most researched epigenetic mechanisms known, and its role as a modifier of gene transcription and a maintainer of terminally differentiated cellular phenotype is well established (Bird, 2002). CpG dinucleotides are underrepresented in the genome, occurring at approximately only 21% of the expected frequency (Lander *et al.*, 2001). Furthermore, when they do occur, they tend to cluster together forming CpG islands, most of which occur in gene promotor regions (including transcription start sites), and an estimated 72% of gene promotor regions contain CpG

islands (Saxonov, Berg and Brutlag, 2006). The methylated states of these CpG islands directly corresponds to the expression of these genes, with methylated CpG islands in promotor regions preventing transcription, and essentially silencing that gene. However, CpG island methylation is not a requirement of gene silencing, and the majority of CpG islands that reside in gene promotor regions are not methylated in adult somatic tissues (Antequera and Bird, 1993).

The regions that directly border CpG islands are known as CpG shores (<2kb flanking CpG islands), and CpG shelves (<2kb flanking CpG shores), and beyond these the remaining genome with only a sparse population of CpGs is known as the opensea. These regions are much less densely populated with CpGs, and their role is less well understood, however with the introduction of technology that allows the study of genome wide DNA methylation, there has been an increasing interest in the wider roles played by such DNA methylation (Gupta, Nagarajan, and Wajapeyee, 2010). For example, many of these more sparse CpG regions also occur in gene bodies and promotor regions, and have been recognised to play a differential role in gene regulation compared to CpG islands. Specifically, these tend to correspond to tissue specific genes, and therefore may control the regulation of genes that trigger cascades of expression that define specific tissues and cell phenotypes (Cedar and Bergman, 2012; Slieker *et al.*, 2013; Lokk *et al.*, 2014). Additionally, DNA methylation and other epigenetic mechanisms in adult brains are understood to be more active than elsewhere in the body (Bjornsson, 2015) and recently we have begun to understand the role the epigenome plays in the adult brain. For instance, the robust learning and memory model of contextual fear conditioning is now understood to be associated with a DNA methylation response in the hippocampus in rodents (Halder *et al.*, 2016; Duke *et al.*, 2017). Such discoveries have prompted the emergence of the field of neuroepigenetics, which examines how the epigenome is involved in learning, memory, synapse connectivity (Cholewa-Waclaw *et al.*, 2016; Sweatt, 2016; Kim and Kaang, 2017). Interestingly, these technological advancements have also shed light on additional functionality of CpG islands, for instance, approximately 25% of CpG islands do not reside in gene promotor regions, but instead within gene bodies and until recently, their assumed purpose has been as alternative gene promotor regions as they typically correlate with increased gene expression in highly expressed genes (Laurent *et al.*, 2010). However, it is now understood that many of these regions instead play a role in alternative gene splicing (Laurent *et al.*, 2010; Lev Maor *et al.*, 2015). As evidenced by these examples, it has become possible in recent years to refine our understanding of the role that DNA methylation plays in an array processes that were previously much more challenging to study, and it is becoming clear that DNA methylation plays a more dynamic role than previously understood. It therefore follows

that we may be able to use such technology to advance our understanding of the role DNA methylation plays in development, particularly how the fetal methylome contributes to MSN fate specification which could help further refine protocols for differentiating hPSCs towards an authentic MSN phenotype.

1.5.2 The Role of DNA Methylation in Development

It is estimated that only 10% of the human genome is functional (Eddy, 2013; Graur *et al.*, 2013), and as such, it has been argued that the majority of DNA methylation may therefore also play no significant biological role (Edwards *et al.*, 2017). Whilst it remains to be determined exactly how much of the methylome is of functional relevance, at least some DNA methylation plays a significant role in various biological processes. Most widely established are its critical and well understood roles in genomic imprinting (Li *et al.*, 1993), X chromosome inactivation (Heard *et al.*, 1997), and transposon repression (Walsh *et al.*, 1998), which will not be discussed here. More relevant to this thesis are its roles in embryogenesis, fetal development, and tissue specific gene transcription, understanding of the latter two having advanced considerably in recent years.

The role that DNA methylation plays in development begins shortly after oocyte fertilisation, during which there is a near complete loss of DNA methylation across the genome, with the exception of the maintenance of the methylation status of imprinted genes (Li *et al.*, 1993). By the time the blastocyst forms and undergoes implantation, there has been a second global methylation event that induces methylation across the vast majority of the genome, as pluripotency is conferred to the ICM (Reik, Dean and Walker, 2001; Arand, *et al.,* 2015). Only well-defined CpG islands are resistant to this methylation, a process understood to preserve the transcription of important housekeeping cellular functions. This is assumed to be the baseline state of the methylome, and it is from here throughout development that the mature methylome develops (Cedar and Bergman, 2012). This is further validated by the discovery that successful reprogramming of iPSCs requires the recruitment of global methylation to ‘reset’ somatic cells to a pluripotent state (Papp and Plath, 2013). The maintenance of methylation through mitosis via Dmnt1 mediated activity is essential in hESCs, as disruption of these mechanisms results in cell death (Liao *et al.*, 2015). Curiously, this is not the case for mouse ESCs, as they are able to replicate in culture without Dnmt enzymes (Tsumura, *et al.*, 2006). Yet, Dmnts are required in mouse for germ layer specification, as Dnmt knock out mice models cannot produce viable embryos (Li, Bestor, and Jaenisch, 1992; Okano *et al.*, 1999), and Dnmt knock out mouse ESCs

exhibit differentiation defects *in vitro* (Boland, Nazor, and Loring, 2014). This distinction in mouse indicates that these are likely two independent roles for DNA methylation, and implicates DNA methylation as a necessary process for early embryogenesis and germ layer specification, and therefore early lineage commitment.

The precise role of DNA methylation following the formation of the blastocyst is less well understood because of the vast number of processes that occur between blastocyst formation and fully developed organism which still remain undetermined even on the genetic level. However, it is established that there is a transition during development from the generally globally hyper-methylated pluripotent ICM, to the comparatively less methylated adult organism, which occurs gradually in a tissue specific manner as and when relevant genes are required, or silenced, by the developing tissues (Cedar and Bergman, 2012). Of note, these same mechanisms have been observed *in vitro* in PSC based differentiations (Nazor *et al.*, 2012). This tissue specific methylation signature continues into adulthood and defines tissue phenotype (Nagae, *et al.*, 2011; Fernandez *et al.*, 2012), and whilst there are various other factors that also influence DNA methylation in adult tissues, including age and environment, tissue specific methylation is largely preserved throughout adult life. For example, Lokk *et al.*, 2014 studied the global methylation of 17 tissue types collected from 4 post-mortem donors using the Infinium Human Methylation 450K bead chips (Infinium 450K), and determined that 6.4% of global methylation variation was predicted by the donor (including age, environment, and genetic variability), whereas 51.2% was predicted by tissue type. Furthermore, this tissue specific methylation was found to be highly present within gene promotors and gene bodies, implicating a continued role of DNA methylation for gene silencing, gene promotion, and splicing throughout adult life.

Whilst the advancements in genome wide DNA methylation sequencing have been utilised and explored for a variety of adult tissues, fetal tissues still remain relatively underexamined. However, the few studies that have been conducted indicate that DNA methylation could play an even more critical role during fetal development, which will be discussed specifically regarding the brain below. In brief, there is evidence that DNA methylation is more closely conserved between fetal tissues donors, especially relative to comparisons of similar child and adult tissues (Numata *et al.*, 2012). The variability that does occur is primarily associated with developmental age, and shows tissue wide hyper- or hypo-methylation shifts across these developmental ages, which is unlike child or adult tissues (Numata *et al.*, 2012, Spiers *et al.*, 2015), and these methylation changes are tissue specific across fetal tissues (Roost *et al.*, 2017). Collectively, this implies, though does not prove, that DNA methylation is a highly regulated process in fetal tissues, and it is used to mediate gene expression at critical

developmental times. There is direct evidence from adult stem cell populations that DNA methylation is the preferred epigenetic mechanism to regulate gene expression and thereby cellular phenotype in highly proliferative and rapidly developing cells, including in the epidermis (Sen *et al.*, 2011), and in the epithelium of the adult intestine (Sheaffer *et al.*, 2014). This is credited to the relatively high degree of malleability and cellular heritability afforded to DNA methylation compared to other chromosomal modifications.

1.5.3 The Methylome of the Human Fetal Brain

As mentioned, examination of the global methylation profile of tissues has only recently become available, and due to the scarcity of human fetal tissue, there are limited studies examining the role this mechanism plays in fetal neural development.

Numata *et al.* 2012 examined the methylation status of samples of prefrontal cortex (PFC) of fetal, child, and adult post mortem brains using Infinium Human Methylation 27K bead chips (27K), and sought variability first between donor samples within their respective age group, based on age of donor, before comparing between age groups. They observed the expected variability within both adult and child groups, that was presumably mediated in part by donor associated variability. However, when examining fetal PFCs, they found that between donors there was far less variability. Specifically, fetal PFCs varied between samples in this group at 861 CpGs (≈3% of CpGs examined), whereas child and adult PFCs varied at 5506 (≈19%) and 10578 (≈38%) CpGs sites respectively, indicating that the methylation profile is highly conserved across the developing fetal PFCs, but not as much in child and adult PFCs. Strikingly, in spite of having the fewest significantly different CpGs within their age group, the variability of methylation between fetal PFC samples were also the most drastic, and importantly, these changes seemed to correspond to fetal age (observed DNA methylation β-value shift where ≈100% change indicates a complete reversal in site methylation across one year: average shift between fetal donors ≈80%, child ≈1%, adult ≈0.1%). Furthermore, when comparing DNA methylation between age groups (fetal, child and adult), they observed a reversal in methylation across a small subset of 131 CpGs between fetal and child PFCs, a trait which was not observed between child and adult brains. Together this implicates a highly conserved and dynamic functional role for DNA methylation in the brain across development and into early childhood. This is further supported by later research by Spiers *et al.* in 2015, who examined the methylation of fetal whole brain tissues using the Infinium Human Methylation 450K bead chips (450K). Similar to Numata *et al.*, they found no large scale

differences in global methylation between subjects, but they did observed a small subset of CpG regions (28,718 probes, ≈7% of CpGs examined) that differed significantly by age. Supporting the notion that DNA methylation is highly conserved during development, and the variability in the methylome at this time is strongly associated with development.

There is also evidence that different fetal tissues and neural cell types have varying methylomes. Lister *et al*., 2013 studied the methylation of human adult and human fetal frontal cortex, and sorted cells within these samples into NeuN positive and negative populations. They found significant methylation changes between fetal and adult brains, and observed that NeuN positive and negative populations had unique methylation signatures, indicating that methylation in fetal cells varies by cellular phenotype, although this difference was more pronounced in adult than the developing fetal cortex. Lastly, Roost *et al.*, 2017 collected tissue from 21 human fetal organs from 18 different fetuses, including whole brain, and conducted genome wide methylation analysis on these organs using a 450K array. They found all fetal organs tested had specific methylomes indicative of their germ layer and phenotype, and demonstrated the high degree of consistency between donors for each organ including whole fetal brain.

Of particular note, Roost *et al.*, 2017 went on to generate hiPSCs from six of these organs and demonstrated that hiPSCs derived from fetal whole brain tissues retained an organ specific methylation signature that improved induction to a neuronal phenotype compared to skin derived hiPSCs. This trend observed by Roost *et al.* is not unique, but rather an established feature of iPSC generation that is the unintended result of Dnmt1 mediated DNA methylation memory in proliferative cells, and incomplete DNA methylation reprogramming. In iPSCs this occurrence has been termed an *epigenetic memory of tissue of origin* and there is evidence that it can be used to enhance PSC based differentiations (Kim *et al.*, 2010; Polo *et al.*, 2010; Bar-Nur *et al.*, 2011; Tian *et al.*, 2011; Hargus *et al.*, 2014).

1.5.4 Epigenetic memory in hiPSCs

As previously discussed (Introduction 1.4.1), iPSCs are highly similar to ESCs in both appearance and functionality, however, one critical difference between them is their epigenome, in particular their DNA methylation profile. Generally, iPSCs and ESCs share a very similar methylome, indeed, it is now widely accepted that the generation of a fully reprogrammed iPSC requires the ablation of a cells previous epigenetic structure through global hypermethylation across the genome and hypomethylation of

pluripotent specific genes, resulting in a methylome comparable to ESCs (Papp and Plath, 2013). Comparisons of the methylomes of iPSCs and ESCs have revealed this similarity makes them indistinguishable when compared to terminally differentiated cells, however, high powered examination of the methylomes of ESCs and iPSCs reveals that these cell types do appear to have a distinct DNA methylation pattern (Deng *et al.*, 2009). For instance, Doi *et al.*, 2009 compared the DNA methylomes of 3 iPSC lines to the fibroblasts from which they were derived, and found 2179 differentially methylated regions (DMR) between these cell types, whereas a second comparison between those 3 iPSC lines and 3 ESC lines revealed only 71 DMRs between these cell types. Doi *et al.* went on to examine these DMRs more closely and found that the majority of the differences were driven by additional hypomethylation in ESCs (51 DMRs), although a small subset were hypomethylated in hiPSCs (20).

This difference raised an interesting question: are these DNA methylation differences largely stochastic depending on the variability between different rounds of iPSC generation? Or is there some underlying cause for these differences? Doi *et al.* compared the DNA methylation of their fibroblast derived iPSCs, to the fibroblasts they were derived from and found cases of increased methylation levels in iPSCs, beyond the level observed in ESCs or the fibroblasts from which they were derived. As this aberrant DNA methylation pattern was unique to the iPSCs, it is likely a product of the reprogramming process itself. This aberrant DNA methylation has since been examined by several labs, and is now somewhat understood. It tends to take the form of additional hypermethylation beyond that observed in the tissue of origin and hESCs (although some limited evidence indicates this can lead to aberrant hypomethylation too: Lister *et al.*, 2011). Additionally, it may in part be due to Dnmt3a/b mediated activity, which have been shown to have a central role in the methylation shift observed in iPSC generation (Stadtfeld *et al.*, 2008; Nishino *et al.*, 2011; Ruiz *et al.*, 2012; Huang *et al.*, 2014). There is some evidence that this aberrant DNA methylation is influenced by methods of culture, as over time these aberrant methylations seem to fade. It is therefore possible that this aberrant methylation is simply an indication that iPSC reprogramming is a longer process than previously thought (Nishino *et al.*, 2011; Tessarova *et al.*, 2016). However, this is still an issue of debate, as some have argued that this aberrant methylation persists through differentiation and seems to be defining features of individual lines (Nazor *et al.*, 2012).

In addition to the aberrant methylation in iPSCs, Doi *et al.* also determined that some of the differential methylation observed between their iPSCs and ESCs were shared between the iPSCs and fibroblasts, indicating that some of these differences between ESCs and iPSCs are related to the iPSC tissue of origin. It has since been

proposed that this is due to incomplete or inefficient recruitment of genome wide methylation processes during the reprogramming process, which fail to fully methylate the open chromatin observed in the tissue of origin (Kim *et al.*, 2010). This has been dubbed an epigenetic memory of tissue of origin (Summary of key papers in Table 1). Like the aberrant DNA methylation formed during reprogramming, this also seems to fade over time as it is most prevalent in early passage (<30) iPSCs (Kim *et al.*, 2010; Polo *et al.*, 2010; Nazor *et al.*, 2012). But, whilst still present in an iPSC line, this residual epigenetic memory has been shown to affect an iPSC lines differentiation propensity, biasing the line towards cell fates and phenotypes related to the differentially methylated genes, or simply put, back towards the tissue type of origin. This effect seems to be more routinely consistent in mouse iPSCs, but there are emerging trends in hiPSCs too. Interestingly, whilst both aberrant and epigenetic memory derived DNA methylation in iPSCs are largely considered barriers to iPSCs that need to be overcome (Byrne, 2013; Tapia and Schöle, 2016), some lab groups have instead attempted to use this tissue specific epigenetic memory to enhance differentiations of cells from PSC sources with varying levels of success. For example, there have been multiple attempts to use iPSC epigenetic memory to enhance hepatic differentiation, however, there has been no observed improvement in efficiency, gene expression, or *in vivo* effect that can be attributed to epigenetic memory (Ohi, *et al.*, 2011; Liu *et* al., 2011; Heslop et al., 2017). Conversely, there has been success in using iPSC epigenetic memory to enhance differentiations towards specific cellular fates including haematopoietic fates (Kim et al., 2010; Pfaff et al., 2012), cardiac fates (Xu et al., 2012; Sanchez-Freire et al., 2014), insulin producing β-cells (Bar-Nur et al., 2011) corneal/ocular fates (Sareen et al., 2014; Hiler et al., 2015) and neural fates (Tian, *et al.*, 2011; Hargus *et al.*, 2014; Roost *et al.*, 2017), many of which result in enriched target gene expression and improved *in vivo* activity.

Particularly relevant to this thesis are the papers that observed epigenetic memory in hiPSCs derived from neural tissues. First, Tian *et al.*, 2011 examined hiPSCs derived from mouse astrocytes (A-iPSC) and fibroblasts (F-iPSC) and differentiated them towards a neuronal precursor and then dopaminergic phenotype using an EB method in combination with an N2 based medium and SHH, FGF8, and L-Ascorbic Acid (AA). They observed no differences in the initial induction to neural precursor cells, but found a significant improvement for neural differentiation (via βIII-tubulin expression) and positive TH expression in cultures derived from A-iPSCs compared to F-iPSCs. Next, Hargus *et al.*, 2014 generated hiPSCs from fetal neural stem cells (fNSC) that had been collected from the fetal forebrain (NS-iPSC), cord blood (CB-iPSC), and fibroblast (F-iPSC), and examined the potential of these cells to differentiate into midbrain neural precursors using an EB based method in N2B27

medium supplemented at various time points with CHIR 99021, purmorphamine, dorsomorphin, SB, and AA. They identified several effects of epigenetic memory in their NS-iPSCs that were not shared by CB-iPSCs or F-iPSCs. For example, NS-iPSCs exhibited expression of non-midbrain marker DLX1, which was also exhibited by the fNSCs from which they were derived, but is not expected under the culture conditions they used and was not exhibited by other PSC lines. Additionally, exploratory gene expression analysis further indicated enriched expression of additional neural genes important for neural development in NS-iPSCs and fNSCs that were not highly expressed in other cell lines. However, they also observed shared gene expression between CB-iPSC, F-iPSCs and fNSCs that were not observed in NS-iPSCs, yet of these genes only one has a known neural function. Last, when directed towards a midbrain neural precursor fate and transplanted *in vivo,* they found evidence of increased cell survival from NS-iPSCs compared to F-iPSCs, and evidence that F-iPSCs exhibited increased cell survival when transplanted in a mixed population with NS-iPSCs (CB-iPSC were not examined). The last publication exhibiting evidence of a retained epigenetic memory in hiPSCs derived from neural tissues was published by Roost *et al.* in 2017. They generated hiPSCs from fetal whole brain (WB-iPSC) and fetal fibroblast (F-iPSCs), and, using an EB based method and a Stemdiff Neural Induction medium from StemCell Technologies, they found that WB-iPSCs exhibited higher expression of βIII-tubulin and GFAP that F-iPSCs at earlier time points, indicating an increased capacity to differentiate towards a neuronal phenotype and at a faster rate. Collectively, these studies demonstrate a consistent effect of an epigenetic memory of tissue of origin in hiPSCs derived from neural tissues.

It therefore follows that it might be possible to achieve similar results in MSN differentiations by generating hiPSCs from tissues known to have an epigenome suitable for an MSN phenotype. Previously, our lab began examining this possibility, by generating hiPSCs from human fetal WGE, and patterning the resulting cells towards an MSN phenotype (Choompoo, 2015). This work has demonstrated that human fetal WGE tissues can be used to create hiPSCs, and that these cells are capable of differentiating towards a DARPP32 positive MSN phenotype *in vitro* and expressing a wide variety of other MSN markers. Furthermore, when transplanted into the QA lesioned rat striatum they produce grafts exhibiting DARPP32, and there is provisional evidence that this is at a greater efficiency than hESC derived MSNs, cultured under the same protocol. However, in animals grafted with MSN progenitors derived from both hESC and hiPSC origin we observed graft overgrowth and in some cases teratoma formation in this study. Overgrowth is not uncommon in SHH/DKK1 based MSN differentiation protocols, but has not been seen in those based in Activin (Introduction 1.4.4), and therefore this study needs to be repeated using more refined MSN protocols

to prevent overgrowth and examine whether hiPSCs derived from tissues known to have an epigenome suitable for an MSN phenotype can be used to enhance striatal MSN differentiations. Additionally, our previous study, and the three studies outlined above did not completely control for the influence of the hiPSC genome, and therefore it remains to be determined that this effect is independent of genetic variability in hPSC lines.

Table 1.1 Key studies of epigenetic memory in iPSCs

Study	Animal and Tissue of Origin	Did epigenetic memory:		Summary
		persist in iPSCs?	effect differentiation?	
Kim *et al.*, 2010	Mouse. Bone marrow progenitors (B-iPSC), and dermal fibroblasts (F-iPSC)	Yes (direct evidence from DNA methylation analysis)	Yes	First systematic study of epigenetic memory in iPSCs. Generated four lines: B-iPSC, F-iPSCs, ESCs, and nuclear transfer stem cells (N-ESC). Found B-iPSCs more readily formed haematopoietic colonies than F-iPSCs, whereas F-iPSCs more readily formed osteogenic colonies. ESC and N-ESCs were as efficient as each other. DNA methylation revealed three main clusters of cells separating B-iPSCs, F-iPSCs, and ESCs and N-ESCs.
Polo *et al.*, 2010	Mouse. Tail tip fibroblasts, splenic B cells (S-iPSC), bone marrow granulocytes, and skeletal muscle precursors	Yes (direct evidence from DNA methylation analysis)	Yes	First evidence in truly isogenic lines, as iPSCs were all derived from an iPSC mouse chimera with doxycycline inducible OSKM genes. Though four lines were used, they were run in pairs in separate experiments. iPSCs derived from different lines expressed genes related to their cell type of origin, this was found to correspond to DNA methylation, and a lines propensity to differentiate towards cells of related origin (e.g. S-iPSCs readily differentiated towards macrophage phenotypes).
Ghosh *et al.*, 2010	Human. Foreskin fibroblast, adipose stem cells, neonatal fibroblast, and keratinocytes	*Yes* (inferred from indirect evidence)	Not tested	First evidence in human, though differentiation capacity was untested. Transcriptional analysis revealed gene expression was more similar between hiPSCs and their tissue of origin, than all other lines and a hESC line. However, between hESCs and hiPSCs gene expression was consistently more similar than between hiPSCs and their tissue of origin. Of all the hiPSCs, foreskin fibroblasts were most similar to hESCs. No differentiation was conducted.
Tian *et al.*, 2011	Mouse. Astrocytes (A-iPSC) and fibroblasts (F-iPSC)	*Yes* (inferred from indirect evidence)	Yes	First evidence in neuroectoderm derived cells. A-iPSCs displayed slower EB formation and reduced proliferation than F-iPSCs and ESCs. However, when differentiating towards a neuronal and dopaminergic fate, A-iPSCs expressed significantly more βIII-tubulin and TH than F-iPSCs.

Table 1.1 Key studies of epigenetic memory in iPSCs (Continued)

Study	Animal and Tissue of Origin	Did epigenetic memory:		Summary
		persist in iPSCs?	effect differentiation?	
Ohi *et al.*, 2011	Human. Hepatocytes, fibroblasts, and melanocytes.	Yes (direct evidence from DNA methylation analysis)	No (unpublished data)	Provisional evidence that partially reprogrammed DNA methylation could be required for successful induction to pluripotency rather than random, as knock down of one of these genes reduced reprogramming efficiency. Otherwise, similar findings to earlier studies, hiPSCs retain a small degree of transcription and DNA methylation based memory to their tissue of origin. Differentiation was carried out on the cells, but no differences were observed.
Bar-Nur *et al.*, 2011	Human. β-pancreatic (β-cell), non-β pancreatic, and fibroblast	Yes (direct evidence from DNA methylation and histone analysis)	Yes	First evidence that epigenetic memory also influences differentiation in hiPSCs, also suggest that this might be because the differentiation protocols are insufficient in hESCs, and are therefore enhanced by epigenetic memory. Demonstrated β-cell derived hiPSCs maintained more open chromatin structures at critical β-cell genes, and expressed these genes significantly more than all other tested hiPSCs and hESCs both *in vitro* and *in vivo*.
Xu *et al.*, 2012	Mouse. Ventricular myocytes (V-iPSC) and fibroblasts (F-iPSC)	Yes (direct evidence from DNA methylation analysis)	Yes	Found that during chimera formation, V-iPSCs exhibited a bias to contribute to heart formation, due to their epigenetic memory, and spontaneously formed beating cardiomyocytes 2 days before F-iPSC and ESC controls. Further, they exhibited a bias towards ventricular myocyte differentiation in protocols that typically result in mixed ventricular/atrial myocyte populations.
Hargus *et al.*, 2014	Human. Fetal neural stem cells (NS-iPSC), cord blood (CB-iPSC), and fibroblast (F-iPSC)	Yes (direct evidence from DNA methylation analysis)	Yes	First evidence that hiPSCs derived from fetal brain tissue retain an epigenetic memory. NS-iPSCs were enriched for neural genes and retained a methylation signature more similar to their tissue of origin than CB-iPSCs and F-iPSCs. Differences were observed when transplanting *in vivo* to the mouse cortex, indicating improved graft survivability of NS-iPSCs.

1.6 Summary and Thesis Aims

Considering the research outlined above, it is clear that CRT is a viable and worthwhile therapeutic strategy to restore the neurological resources and function lost in the process of many neurodegenerative diseases, including HD, which is a useful model for the continued advancement of CRT. However, there are various barriers that currently prevent the introduction of CRT for HD to clinic, foremost of these being the lack of a reliable and renewable source of cells capable of differentiating into authentic MSNs. The most widely researched of the alternative cell sources capable of producing striatal MSNs are pluripotent stem cells, including hESCs and the more recently discovered hiPSCs. However, at present these stem cells require complete direction towards the correct cellular phenotype, and with our presently limited knowledge of the various aspects that facilitate this process, this is currently difficult to achieve reliably. It therefore follows that we need to continue building on our understanding of the developmental mechanisms required to produce an authentic MSN phenotype. An aspect of development that has thus far been relatively neglected is the epigenetic mechanisms that underpin cell fate decisions. The most well understood of these is perhaps DNA methylation, which has been demonstrated to play various roles in cellular phenotype specification and regulation. Interestingly, iPSCs have been shown to retain some of the epigenetic mechanisms associated with the tissues from which they were derived, which is known as an epigenetic memory of their tissue of origin. It therefore follows that we may be able to use this epigenetic memory of tissue of origin to enhance our MSN differentiations and potentially begin the exploration of this avenue of development. In this thesis I aim to explore this possibility by generating hiPSCs from fetal LGE tissues, the origin of MSN progenitors, and comparing their differentiation capacity to isogenic hiPSC controls derived from other tissues.

Chapter 2: Materials and Methods

2.1 Cell Culture

2.1.1 General Cell Culture Techniques

All cell culture was conducted in a Class II Safety Cabinet under sterile conditions. Cells were grown on cell culture treated plastic Nunc™ multi-well plates (Thermo Scientific), and maintained in incubators at 37°C and 5% CO_2. Cells were maintained in volumes of media proportionate to their culture plate (Table 2.1). Cell counts were conducted manually using Trypan Blue (Invitrogen) and a haemocytometer, under dilution factors between 2-250, to achieve an average count of ≈200 cells.

Table 2.1 – Media volume and Plating density by well size.

Plate diameter (Well Size)	Volume of media per well	Plating density:
6 well plate (9.6cm^2)	3 ml – 4 ml	N/A*
12 well plate (3.5cm^2)	2 ml – 2.5 ml	400,000/well
24 well plate (1.9cm^2)	1 ml – 1.5 ml	200,000/well

**only PSCs were cultured in 6 well plates and were not plated according to cell number but according to ratio of a split at each passage.*

2.1.2 Acquisition, dissection and dissociation of human fetal tissues

Primary human fetal tissues were acquired from the South Wales Initiative for Fetal Transplantation (SWIFT) research tissue bank which has full ethical approval to obtain fully informed consent to collect fetal tissues from maternal donors undergoing an elective termination of pregnancy. Throughout this thesis donor tissues and their derivatives are referred to by their SWIFT number (e.g. SWIFT 2285), further details for all samples are described in Table 2.2. Primary human fetal tissues and their relevant derivatives were handled under the Cardiff University Human Tissue License, which is in compliance with the HTA guidelines.

Primary human fetal tissues were dissected according to Roberton *et al*., 2018 (Figure 2.1), and maintained in Hibernate E media (Gibco) at 2-8°C for a maximum of 7 days until prepared for dissociation. The tissue was harvested from Hibernate E media and washed in a 15 ml falcon tube by adding 500 µl wash media (DMEM+Dα:

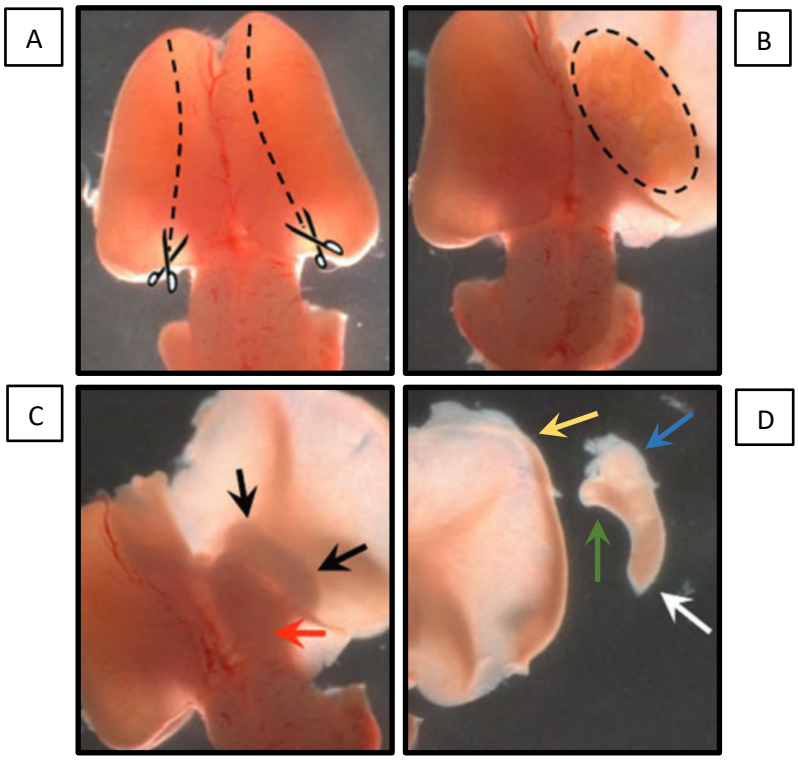


Figure 2.1 – Dissection of fetal tissues (fibroblast not shown)
Images retrieved and adapted from Roberton et al., 2018.

A. The cortex is cut close to the midline in a caudal-to-rostral manner as indicated by dashed black lines.
B. The cortex is opened to reveal the choroid plexus (circle of dashed black lines).
C. The choroid plexus is removed to expose the WGE (black arrows) lying on the cortical surface beside the thalamus (red arrow).
D. The WGE is removed by cutting underneath the structure. Then it is divided into LGE (blue arrow), MGE (green arrow), and CGE (white arrow). Cortical tissue can then be collected (yellow arrow).

Dulbecco's Modified Eagle Medium: Nutrient Mixture F-12 (DMEM/F12, Gibco), with 2%Pulmozyme Dornase Alfa (Dα, Genentech)), then disturbing the tissue manually using a pipette, and then removing the media. This was repeated 2-3 times to ensure removal of Hibernate E media. Following the final wash, 500 µl of TrypLE Select (Gibco) +2% Dα was added to the tissue sample, and incubated at 37°C with 5% CO_2. Neural tissues were incubated for 10 minutes, and fibroblast tissues were incubated for 40 minutes and inverted every 10 minutes to achieve full dissociation. Following this incubation period 5 ml of DMEM+Dα wash media was added to the falcon tube and the tissue was centrifuged at 1000 rpm for 3 minutes. The excess media was carefully removed, and replaced with 200 µl fresh DMEM+Dα, and the tissue was triturated 20-25 times using a 200µl pipettor to dissociate the cells into a single cell suspension, which was then counted to determine total viable yield.

Table 2.2 – SWIFT tissue donation details.

SWIFT number	Crown to rump length (CRL)	Age post conception (age p.c.) (days)	Dissected by	Tissues	Purpose
2242	25 mm	50	Sophie Rowlands	LMCF*	GFP SeV test
2285	42.1 mm	61	Rachel Hills	LMCF*	hiPSC generation
2230	≈55 mm	65	Sophie Rowlands	WGE	Methylation analysis
2240	40 mm	62	Sophie Rowlands	WGE	Methylation analysis
2415	53.4 mm	69	Oliver Bartley	LGE	Methylation analysis
2451	37.8 mm	70	Rachel Hills	LGE	Methylation analysis
3005	39.5 mm	59	Victoria Roberton	WGE	Methylation analysis

**LMCF = LGE, MGE, cortex, and fibroblast*

2.1.3 Plating and culturing of human fetal tissues

Plastic multiwell plates were prepared with Poly-D-lysine (PDL, Gibco) prior to plating of primary human fetal tissues. 500 µl of PDL solution (10 µg/ml of PDL/dH_2O) was added to each well, and incubated for one hour. The PDL solution was then removed, and the plates were washed using dH_2O, and dried under UV for 45 minutes. Dissociated primary human fetal tissues were then plated as a monolayer onto the PDL-treated plates to achieve ≈100,000 cells/cm^2 (Table 2.1). Neural tissues were cultured in Neural Differentiation Media (NDM: DMEM/F12, +2% B27^{TM} Supplement 50x (B27; Gibco), +1% fetal bovine serum (FBS; Gibco), +1% Penicillin-Streptomycin (PS; Gibco)), and Fibroblasts were cultured in Fibroblast media (DMEM/F12, +10% FBS, +1% MEM Non-Essential Amino Acids Solution 100x (MEM-NEAA; Gibco), +1% L-Glutamine (Gibco), +1% PS. Cells were allowed to settle over night before being washed with Dulbecco's phosphate-buffered saline without calcium or magnesium (DPBS; Gibco) and given fresh culture media. Media was changed every 3 days, and cells were monitored using a microscope daily.

2.1.4 Transduction of Human Fetal Tissue with Sendai Viral Vectors

Fetal tissues were transduced with either the Cytotune™ EmGFP Sendai Fluorescence Reporter Virus (SeV reporter virus; Invitrogen) or the CytoTune™ -iPS 2.0 Sendai Reprogramming Kit (SeV-iPS; Invitrogen). Fetal tissues to be transduced were plated and cultured as described above. Following 3-6 days of maintenance, the cultures were transferred to a viral suite for transduction (Day 0). A sample well for each tissue type was harvested using TrypLE Select reagent and counted on a haemocytometer using trypan blue to estimate number of cells per well. Viral vector volumes were calculated for each tissue type as follows:

$$\text{Volume of virus (µl)} = \frac{\text{Desired MOI (cell infectious units/cell) x number of cells}}{\text{titre of virus (cell infectious units/ml) x } 10^{-3} \text{ (ml/µl)}}$$

Fresh NDM and fibroblast media was warmed in a copper-sulphate treated water bath at 37°C in separate falcon tubes for each tissue type. The viral vectors were thawed by submerging in the water bath for 5-10 seconds, and then allowed to thaw at room temperature. Once thawed they were centrifuged briefly and immediately placed on ice. For each tissue type media was supplemented with the calculated volume of each viral vector and mixed by titration. The SeV-media was then immediately applied to the tissues to be transduced, and returned to an incubator at 37°C with 5% CO_2. Following a 24 hour incubation in the media containing SeV, the cells were washed with DPBS and fresh NDM/fibroblast media was added to the cultures, which were then maintained according to human fetal tissue culturing methodologies outlined above (Methods 2.1.3).

3 days after fetal tissue cultures were transduced with a SeV reporter virus, they were fixed and imaged. Whereas, 5 days after fetal tissue cultures were transduced with a SeV-iPS virus, they were washed with DPBS and maintained in Essential 8™ medium (E8, Gibco) which was replaced daily. Between days 8-14 developing cell colonies were manually picked from each infected tissue type and cultured individually on 24 well plates pre-treated with Geltrex™ Matrix (Gibco; Geltrex treated plates prepared as recommended by Gibco). Once a colony was picked from the infected cultures, it was cultured and expanded as PSCs (Methods 2.1.5).

2.1.5 Culturing and expansion of human PSCs

The ESC line H9 used throughout this thesis was purchased from Wicell. PSCs (both ESC and iPSC) undergoing expansion were cultured on 6 well plates pre-treated with either Geltrex or Matrigel™ Matrix (Corning; Matrigel treated plates prepared as recommended by Corning). Monolayer cultures were maintained in E8 or Essential 8 Flex™ medium (E8+; Gibco) (note: E8 and E8+ were not used in conjunction). E8 was replaced daily, E8+ was replaced every other day on <50% confluent cultures and daily on >50% confluent cultures. The volume of media was varied according to culture vessel size (Table 2.1). PSC cultures that were at 80% confluency, or those that were expected to surpass 90% confluency overnight were passaged at a 1:3-6 ratio. A cycle of expansion typically took 5 days, but varied by cell line, clone number, split ratio, passage method and passage number. Cultures were passaged using one of two methods:

If cultures were free of differentiating cells then they were passaged using an EDTA (Sigma) scrape: cultures were washed with DPBS and then EDTA was applied to the well (enough to cover the base of the vessel, e.g. 1 ml/well of a 6 well plate). The cultures were incubated for 2 minutes until cells exhibited signs of detaching. The EDTA was next aspirated and fresh E8/E8+ was applied to the well. Immediately afterwards cells were detached from the well by scraping the base of the well with a sterile 10 ml stripette (Sarstedt). The cells were collected and centrifuged at 1500 RPM for 3 minutes, before being resuspended in fresh E8/E8+ media for plating.

Alternatively, if differentiating cells were present in the cultures then the cells were passaged using ReLeSR™ (StemCell Technologies): cultures were washed with DPBS and then ReLeSR was applied to the well (enough to cover the base of the vessel). The cultures were incubated for 5 minutes until cells exhibited extensive signs of detaching throughout the colonies. Fresh E8/E8+ was applied to the well, and the plate was tapped repeatedly for 30 seconds to dislodge the PSCs from the plate. The cells were then collected and centrifuged at 1500 RPM for 3 minutes, before being resuspended in fresh E8/E8+ media for plating.

2.1.6 Spontaneous differentiation of PSCs

PSC cultures were grown between passages 8-12 as a monolayer in a 24 well plate on glass cover slips under PSC culture conditions described above until plates were 40% confluent, at which time spontaneous differentiation was initiated (Day 0). On day 0, E8/E8+ media was removed and cultures were washed with DPBS, and were instead cultured in KSR medium (KnockOut™ DMEM (Gibco), +10% KnockOut™

Serum Replacement, +1% L-glutamine, +1% MEM-NEAA, +1% PS, +0.1% β-mercaptoethanol (Gibco)). Cultures were maintained for 5 days in KSR medium, before being processed for Immunocytochemistry.

2.1.7 Directed differentiation of PSCs towards a Medium Spiny Neuron fate

PSCs were patterned towards an MSN phenotype using a protocol adapted from previously published works (Arber *et al.*, 2015, Repair HD; Figure 2.2A). For the experiments described in this thesis, MSNs were differentiated using two protocols: Protocol 2 and Protocol 3 (Figure 2.2B-C). The critical difference between these protocols is the timing of the second passage, which is described below.

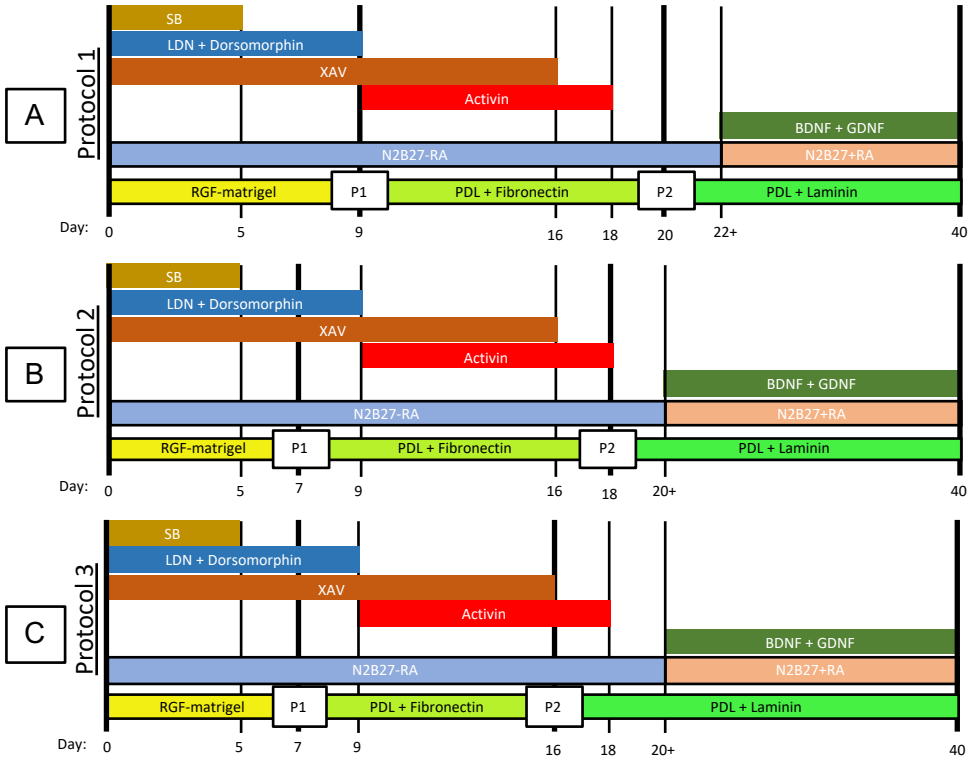


Figure 2.2: MSN differentiation protocols
A-C. Schematics of the different versions of the MSN differentiation protocol used in this thesis. Protocol 1 (**A**) was not used for experimental design, and therefore is not described here. Protocol 2 (**B**) and Protocol 3 (**C**) are used for separate experiments. The critical difference between these protocols is the timing of passage 2 (P2).

For both Protocol 2 and 3: PSCs were passaged as described above (Methods 2.1.5), and plated as a monolayer culture in E8/E8+ on 12 well plates treated with Growth Factor Reduced (RGF) Matrigel™ (Corning; diluted 1:15 in DMEM-F12 and coated for 1 hour at 37°C). When PSCs were >70% confluent, cultures were considered

to be at Day 0 of the differentiation, and medium was changed to N2B27-RA medium (Base media 2:1 ratio of DMEM/F12 and Neurobasal medium (Gibco), +1% L-glutamine, +0.6% N2 supplement 100x (N2; Gibco), +0.6% B27^{TM} without retinoic acid supplement 50x, (B27-RA; Gibco), +0.2% MycoZap^{TM} plus CL (Lonza), +0.1% β-mercaptoethanol. Between day 0-5, N2B27-RA was supplemented with 10 μM SB-431542 (SB; Tocris), and between days 0-9 N2B27-RA was supplemented with 200 nM dorsomorphin (R&D),100 nM LDN-193189 (LDN; Tocris) and 2 μM XAV-939 (XAV; Tocris). Cells were washed with DPBS on day 1, and given fresh medium. From that day onwards medium was changed by removing half the spent medium and replenishing with fresh medium every other day. Cells were monitored using a microscope daily.

On day 7, cultures were passaged (P1). This passage was achieved by collecting the culture medium (retained for later use) and adding fresh N2B27-RA (+ dorsomorphin, LDN, and XAV) supplemented with 10 μM Y-27632 (ROCK inhibitor; Millipore). Cells were then incubated at 37°C for one hour in this medium. The medium was then collected (retained for later use), and the cells were washed with DPBS and incubated at 37°C in EDTA for 1 minute. EDTA was then removed, and the ROCK inhibitor supplemented medium was returned to the cultures. Immediately following this, the cultures were gently scraped from the culture vessel using a 10 ml stripette, and resuspended in all previously collected medium. This mixture was then topped up to allow 2 ml per well following the passage, and split at a 2:3 ratio onto 12 well plates treated with PDL and Fibronectin (Millipore; 15 μg/ml in dH_2O, coated for 1 hour at 37°C). Cells were allowed to settle back to a monolayer culture overnight. The following day, cultures were washed with DPBS and given fresh N2B27-RA medium (+ dorsomorphin, LDN, and XAV).

Following P1, cultures completed forebrain progenitor patterning, and began LGE specific regional patterning. This was achieved by supplementing N2B27-RA medium from day 10-18 with 25 ng/ml Activin A (R&D) and 2 μM XAV.

For cells undergoing Protocol 2, passage 2 (P2) was conducted on day 18 of the differentiation. For cells undergoing Protocol 3, P2 was conducted on day 16. P2 was conducted by washing cultures with DPBS, then incubating them at 37°C in Accutase (Sigma) for 5 minutes. Following this, each well was manually disturbed by trituration, and then incubated for an additional 5 minutes. Cells were then broken up into a single cell suspension by trituration, collected, and suspended in a 15 ml falcon tube with fresh N2B27-RA. The cells were then centrifuged at 1500 RPM for 3 minutes, resuspended in 1 ml fresh N2B27-RA and counted for total cell yield. Following counts, cells were either frozen for future use at 5,000,000 cells per vial (Methods 2.1.8), or

seeded onto plates treated with PDL and laminin (Sigma; 10 µg/ml dH_2O, coated for 1 hour at 37°C).

After day 20, cells were cultured in N2B27+RA medium (as N2B27-RA above, but B27-RA is replaced with B27), and supplemented with 10 ng/ml BDNF (Peprotech) and GDNF (Peprotech). This medium was refreshed every third day, and cells were matured in this medium for a maximum of 60 days.

2.1.8 Freezing, storage, and thawing of PSCs and MSN progenitors

To freeze, cells were passaged as described above (PSC: Methods 2.1.5; MSN progenitors: Methods 2.1.7), but following centrifugation the cells were resuspended in a freezing medium of either relevant culture medium (PSC: E8/E8+; MSN progenitors: N2B27-RA) with 10% DMSO (Sigma) or in Cryostor CS10 freezing media (StemCell Technologies), and transferred to a cryovial (Starlabs) suitable for temperatures below -80^{o}C. The vials were then gradually cooled in a -80^{o}C freezer, in a Cool Cell Freezing Container (BioCision), and stored for later use. Once at -80^{o}C, cells destined for long term storage were transferred to liquid nitrogen.

Before thawing PSCs, thawing media was first prepared by adding 1% RevitaCell^{TM} supplement 100x (Gibco) to E8/E8+. MSN progenitors were thawed in N2B27-RA without additional supplements. To thaw cells, vials were removed from storage, and thawed using an automated cell thawing platform (ThawSTAR, Biocision). 1 ml of thawing media was then added dropwise to the vial, to complete thawing, and the diluted solution transferred dropwise to a 15 ml falcon tube. The cells were then centrifuged at 1000 RPM for 3 minutes, before being resuspended in fresh thawing media by gentle trituration. The cells were then plated in conditions appropriate for their cell type (PSC: Methods 2.1.5; MSN progenitors: Methods 2.1.7) in thawing media and allowed to recover overnight. The next day, cells were washed with DPBS and cultured as previously described.

2.1.9 Preparation of PSC derived MSN progenitors for transplantation

Two days prior to surgery, PSC derived MSN progenitors to be transplanted were thawed from -80^{o}C storage as described (Methods 2.1.8), and cultured on 24 well plates treated with PDL and laminin in N2B27-RA media, supplemented with Activin A and XAV (day 19 of the MSN protocol, Methods 2.1.7). The day after plating, cells were washed with DPBS to remove debris, and given fresh media. On the day of surgery, cells were prepared by washing with DPBS and then dissociated with Accutase as

described for passage 2 of the MSN differentiation protocol (Methods 2.1.7). Cells were then counted, and suspended in DMEM/F12 at 250,000 cells/µl, and then transported on ice to surgery.

2.2 Cell analysis

2.2.1 Immunocytochemical Staining and analysis

Cells to be stained were washed once with phosphate buffered saline (PBS; gibco) and fixed in 3.75% PFA (Millipore) for 15 minutes at room temperature. Once fixed, cells were washed three times with PBS, and stored in PBS at 4°C until required.

Cells were permeabilised in 100% molecular grade ethanol (Sigma) for 2 minutes, and washed three times with PBS before undergoing blocking for one hour at room temperature in PBS containing 3% normal serum (Gibco) and 3% BSA (Sigma). Primary antibodies were added to PBS (concentrations detailed in Table 2.3) with 3% goat serum and 3% BSA, and allowed to bind to antibody targets at 4°C overnight. In contrast, controls were maintained in blocking solution without primary antibodies. The following day, cells were washed three times with PBS, and exposed to fluorescent secondary antibodies at 1:200 in PBS with 3% goat serum, and 3% BSA for 2 hours in darkness at room temperature. Following three PBS washes, cells were stained with Hoechst for 5 minutes, and then washed a final three times in PBS. Cells adhered to coverslips were mounted onto glass slides using a PBS and glycerol (Fisher Chemicals) solution (1:1 ratio), and fixing with clear nail varnish. Cells kept in multi-well plates were instead submerged in PBS, +1% PS. All cells were stored in darkness at 4°C to maintain fluorescence.

Stained cells mounted on slides were imaged using a ZEISS Axio Imager 2 upright fluorescent microscope in conjunction with Axiovision software, whereas cells in multi-well plates were imaged using a Leica DMI6000B inverted fluorescent microscope in conjunction with LAS X software. In both cases, cell counting was conducted manually on ImageJ software, using 5 images taken at fields of view at 40x magnification. Fields of view were initially selected randomly, then assessed by Hoechst staining only to avoid uncountable large clusters of cells. Quality control was undertaken in all experiments prior to image acquisition, by determining true fluorescent staining using a control well that had not been exposed to the primary antibody, and then adjusting exposure and gain to eliminate all false staining in experimental wells. Scale bars were applied to images post-processing using ImageJ.

Table 2.3 – Antibody details

Antibody	Company	Species	Concentration
OCT4	Abcam	Rabbit	1:1000
TRA-1-60	Abcam	Mouse	1:100
SOX2	Abcam	Rabbit	1:100
NESTIN	Neuromics	Mouse	1:1000
NANOG	Abcam	Rabbit	1:100
SSEA4	Abcam	Mouse	1:100
DESMIN	Abcam	Rabbit	1:1000
α-SMA	Sigma	Mouse	1:500
α-FP	Abcam	Mouse	1:100
VIMENTIN	Millipore	Mouse	1:1000
GSX2	Abcam	Rabbit	1:100
βIII-tubulin	Sigma	Mouse	1:1000
FOXP1	Abcam	Rabbit	1:500
DARPP32 (human)	Abcam	Rabbit	1:1000*
MAP2a/b	Sigma	Mouse	1:200
HuNu	Millipore	Mouse	1:1000
CTIP2	Abcam	Rat	1:500

*For both immunocytochemistry and immunohistochemistry

2.2.2 Gene expression analysis

2.2.2.1 RNA extraction and cDNA synthesis

Cells were harvested either during a passage, or by applying accutase to detach and dissociate cells for RNA extraction. Samples were then either transferred to RNA later (Ambion) and stored for later use at -20^{o}C for up to 3 weeks before RNA extraction, or immediately processed for RNA extraction. RNA extraction was achieved using the Qiagen RNeasy Mini Kit, following the manufacturer protocols, and including the optional DNase treatment steps. RNA was then stored at -80^{o}C for up to one month, before undergoing cDNA synthesis.

Defrosted RNA was measured using a Nanodrop (NanoDrop™ one, ThermoFisher Scientific) following thawing to determine quality and quantity. Next first strand cDNA synthesis was conducted by incubating up to 1 µg RNA with 50 µM random primers (Eurofins Genomics) and 10 mM dNTP buffer (Invitrogen) for 5 minutes at 65°C. Next, the RNA was kept on ice for 1 minute, then mixed with SuperScript™ IV First Strand buffer x5 (Invitrogen), 100 mM DTT, RNaseOUT™ Recombinant RNase Inhibitor (Invitrogen), and SuperScript™ IV Reverse Transcriptase (Invitrogen). Negative controls were mixed with Nuclease free water (Ambion) instead of SuperScript IV reverse transcriptase. The RNA mixture was then incubated at 23°C for 10 minutes, 55°C for 10 minutes, and then 80°C for 10 minutes. cDNA was then measured using a Nanodrop, and diluted to achieve a final concentration of 10 ng/µl and stored at -20°C until required.

2.2.2.2 Reverse Transcription PCR

Reverse transcription polymerase chain reactions (RT-PCR) were conducted by amplifying cDNA using a basic PCR mix comprising 10X Bioline NH4 Reaction Buffer, 50mM MgCl2, 10mM dNTP, BioTaq DNA polymerase (all Bioline), forward and reverse target primer pairs at 10pmol each, and nuclease free water to bring the final reaction to 25 µl. The reaction was denatured at 95°C for 1 minute, and then underwent cycles of denaturing at 95°C for 45 seconds, annealing at 60°C for 1 minute and extension at 72°C for 1 minute. Following 30 cycles, the reaction sample underwent a final extension at 72°C for 10 minutes. PCR products were analysed by electrophoresis on a 1% agarose gel and visualised with SafeView™ (ABM).

2.2.2.3 Quantitative Reverse Transcription PCR and the ΔΔCt method

Quantitative Reverse Transcription PCR (qPCR) was conducted using PowerUp™ SYBR™ Green (Invitrogen). A master mix of each gene to be analysed was created by combining PowerUp SYBR Green master mix x2, with both forward and reverse primers of the target gene (Table 2.4), and Nuclease free water (Invitrogen) at sufficient volumes for 19 µl/well. This was then frozen in small aliquots, to standardise the master mix between cell lines. For each gene and each cell line to be analysed, 10 ng of relevant cDNA was added to three of every four wells of MicroAmp™ Fast Opical 96-well reaction plates (Applied Biosystems), and a negative cDNA control was included in the fourth well. 19 µl of the relevant gene master mix was then mixed with the cDNA, and the plate was sealed with MicroAmp™ Optical Adhesive Film (Applied Biosystems), and centrifuged for 20 seconds. Plates were then frozen at -20°C for a maximum of two days until required. Plates were analysed in a StepOnePlus™ Real-Time PCR System (Applied Biosystems) using StepOne™ Software under the following

amplification conditions: 95ºC for 15 minutes, and 40 cycles of 95ºC, 60ºC, and 72ºC for 30 seconds each. Melt curves were generated from readings every 0.5ºC between 53ºC and 95ºC.

To avoid variability caused by efficiency of primer pairs, cycle threshold (Ct) values were set at a standardised level for all genes, which coincided with exponential growth of expression across all genes examined. Expression within a given sample of cDNA for each gene was then examined for quality control, Ct values within a technical triplicate were compared within the triplicate, and samples where Ct values varied by more than 0.3 from other samples were removed as outliers. A final Ct value for each gene was then determined by averaging the remaining triplicate Ct values. The relative expression for each gene within each experimental condition was then calculated by generating a delta Ct value for that experimental condition. This was generated by first averaging the Ct value of two house-keeping genes (β-ACTIN and GAPDH), and then subtracting the average house-keeping Ct value from the candidate gene Ct value. Next, the expression for each gene within each experimental condition was made relative to a control condition by generating a delta-delta Ct value (ΔΔCt). This was calculated by first averaging the delta Ct values of a control population, and then subtracting this value from the delta Ct value from each sample within an experimental condition. Finally, fold change expression values (compared to a control condition) were determined by calculating 2^-(ΔΔCt). Fold change expression values were used for statistical analysis.

Table 2.4 – PCR primer details

Gene	Forward Primer (5'->3')	Reverse Primer (5'->3')
β-ACTIN	GCTCTTTTCCAGCCTTCCTT	CGGATGTCAACGTCACACTT
GAPDH	GTCGGAGTCAACGGATTTGG	ATGGAATTTGCCATGGGTGG
OCT4	GTGGAGGAAGCTGACAACAA	ATTCTCAGGTTGCCTCTCA
SOX2	CAAGATGCACAACTCGGAGA	GCTTAGCCTCGTCGATGAAC
DLX2	AGCAGCTATGACCTGGGCTA	TCCTTCTCAGGCTCGTTGTT
NKX2.1	ATGTCGATGAGTCCAAAGCA	CTCCATGCCCACTTTCTTGT
NOLZ1	CTGCAGCCCCTGCCTTCCAC	GTCGGGCTTCCCGATCTGCG
CTIP2	CCATCCTCGAAGAAGACGAG	ATTTGACACTGGCCACAGGT
FOXP1	CGATCCCTTCTCTGATTTGC	CATGCATAATGCCACAGGAC
FOXP2	CCACGAAGACCTCAATGGTT	GCTCTTCCTTGACGTGGATT

DARPP32	CTGGGCAAAAGACAACCTGT	GGTCTTCCACTTGGTCCTCA
CALB1	GATACTGACCACAGTGGCTTC	TGCACTGGTAGTAACCTGGC
DRD1	AGCGAAGTCCACATTCCAAG	ATGTCTTCTCGCTCCTCCAA
DRD2	TCCTGAACTTGTGTGCCATC	GATGGAGATCATGACGGTGA

2.2.3 Multi-Electrode Array Analysis

Spontaneous electrical activity of cultured MSNs was examined using multi-electrode array (MEA) analysis. hPSCs undergoing an MSN differentiation were plated onto platic 24 well MEA plates with 12 gold 100 µm electrodes per well (Multichannel Systems) and cultured as described in 2.1.7. Recordings were conducted every third day up to day 55 using a Multiwell MEA headstage (Multichannel Systems) and Multiwell-Screen software (Multichannel Systems). Data was visualised and processed using Multiwell-Analyzer software (Multichannel Systems).

2.3 In Vivo Methods

2.3.1 Animal care and Immunosupression

All animal experiments were performed in compliance with local ethical guidelines and approved animal care according to the UK Animals (Scientific Procedure) Act 1986 and its subsequent amendments. Adult female Sprague-Dawley rats (Charles River) weighing 200-250 g were housed in cages of up to 4 animals in a natural light-dark cycle with access to food and water ad libitum. Animals were health checked on a daily basis, with weekly visits from a named veterinary surgeon. Starting at least 1 day before unilateral striatal transplantation surgery (Methods 2.3.4), all animals underwent immunosuppression to promote xenograft survival. They received daily intra-peritoneal injections of Cyclosporin A (Novartis) at 10 mg/kg. Weights were updated weekly, and the dose of Cyclosporin A was adjusted for each animal to accurately reflect weight changes across the course of the study.

2.3.2 General Surgical Procedures

Isoflurane was used to induce anaesthesia for all surgeries. This was initially achieved in an induction chamber with 0.8 L/minute oxygen and 5% isoflurane. During surgery, anaesthesia was maintained with passive inhalation of isoflurane at 1.5-2.5% L/minute in a mixture of oxygen at 0.8 L/minute and nitrous oxide at 0.4 L/minute. Anaesthetic depth was monitored during surgery with hind limb withdrawal reflex tests and breath rate monitoring, the isoflurane dosage was adjusted accordingly to maintain a depth of anaesthesia appropriate for surgery. All surgeries were performed on a stereotaxic frame, animals were fitted to the frame using a nose bar and ear bars to keep the head secure and level. To maintain regular body temperatures animals were kept in plastic sleeves on protective padding on a heat mat.

During surgery, all animals received a subcutaneous injection of Metacam (30 µl) for analgesia and 0.09% saline glucose (5 ml) for hydration. Directly following surgery, they were placed in a recovery chamber that maintained a 38°C temperature until they were active, at which time they were transferred to a recovery cage overnight. The next morning animals were returned to their home cages and monitored for weight and health for a minimum of 3 days post-surgery.

2.3.3 Surgery: unilateral striatal lesion via Quinolinic Acid

Before surgery, quinolinic acid (QA) was dissolved in 0.1 M phosphate-buffer to make a 15 mg/ml (90 mM) solution and pH adjusted to pH 7.4. Animals were anaesthetised for surgery, and fitted to the stereotaxic frame. A cannula needle was attached to a 10 µl Hamilton syringe driven by a mechanical pump at a setting delivering 0.28 µl/minute. An incision was made to expose the skull, and bregma was located and measured. Two burr holes were drilled into the skull at two sites above the right striatum at the co-ordinates -0.4/+1.2 anterior-posterior and -3.7/-2.9 medial-lateral from bregma. QA was injected via the canula at depths of -5 mm and -4 mm below dura mater for 56 seconds at each depth, delivering approximately 0.26 µl QA per depth. The needle was left in place at the upper depth for an additional 3 minute diffusion. This was repeated in both target sites.

Following the second diffusion the cannula was withdrawn and the incision was cleaned and sutured. Before moving the animal to the recovery chamber 150 µl of diazepam was administered subcutaneously to prevent lesion induced seizures.

2.3.4 Surgery: Unilateral striatal transplantation

Unilateral striatal transplantation surgery was conducted between 1-3 weeks following QA lesion surgery. Before surgery, cells for transplantation were prepared as described above (Methods 2.1.9). Animals were anaesthetised for surgery, and fitted to the stereotaxic frame with the animals head level. The skull was exposed and a hole was drilled directly between the two previously measured holes from QA surgery. A Hamilton syringe was loaded with 2 µl of the prepared cell mix (250,000 cells/µl) and 1 µl was delivered manually over the course of 1 minute at each depth of -5 mm and -4 mm below dura mater (500,000 total cells per animal). The syringe was left in place for an additional 3 minute diffusion, before being removed and the incision cleaned and sutured. The animal was then moved to a recovery chamber.

2.3.5 Apomorphine Induced Rotations and Analysis

Lesions were assessed between 1-2 weeks following QA surgery by use of apomorphine induced rotations, and graft induced recovery was examined every fourth week from the 12th week after transplantation. Prior to testing, apomorphine (Sigma) was dissolved in saline at 0.75 mg/ml, using an opaque bottle on a shaker for 10 minutes. Animals were injected subcutaneously with the prepared apomorphine proportional to their body weight at 1 µl/g. Immediately following this, they were placed

into automated rotometer cylinders (Rotorat) and the frequency of clockwise and anticlockwise rotations were recorded for 60 minutes. Apomorphine induced rotations were calculated first by subtracting all ipsilateral rotations from all contralateral rotations, then averaging across the 60 minute testing period.

2.3.6 Perfusion

Animals were terminally anaesthetised by overdose of sodium pentobarbital (Euthatal, Merial), delivered as an intra-peritoneal injection. Upon cessation of breathing and negative limb and eye reflex tests the animals were transcardially perfused with a prewash solution (1.8% di-sodium hydrogen phosphate and 0.9% sodium chloride, in dH_2O, at pH7.3) for 2 minutes followed by 4% PFA solution (4% PFA, in prewash at pH7.3) for 5 minutes. The brains were removed and transferred for post-fixing in a 4% PFA solution for an additional 4 hours. Following this brains were washed with prewash, and transferred to 25% sucrose in prewash solution until they sank and were ready for tissue sectioning.

2.4 Graft Analysis

2.4.1 Tissue Sectioning

Once perfused brains had sunk in 25% surcrose in prewash, they were sectioned in a 1:12 series and cut coronally using a freezing stage microtome at a thickness of 30µm. Sections were then stored in anti-freeze solution at -20°C for future analysis.

2.4.2 Immunohistochemistry

For each stain conducted, 1 in 12 series of brain sections were washed with tris buffered saline (TBS, pH 7.4), then quenched with 10% hydrogen peroxide and 10% methanol in ddH_2O for 5 minutes. Sections were then washed three times with TBS, and blocked for 1 hour in 3% appropriate normal serum in 0.2% Triton X-100 in TBS (TXTBS, pH 7.4). Next, sections were exposed to the primary antibody (concentrations detailed in Table 2.3) in 1% appropriate normal serum in TXTBS, and left to bind overnight at room temperature. Sections were then washed three times with TBS, and next exposed to a secondary antibody at 1:200 in 1% appropriate normal serum in TBS for 2 hours. Sections were next washed three times with TBS, and strepdavidin ABC (A and B both at 1:200 dilution in 1% appropriate normal serum in TBS, prepared 30 minutes before use) was added for a further two hours. Next the sections were washed 3 times with TBS, and then two washes in 0.05 M tris non saline (TNS, pH 7.4) and left in TNS overnight at $4^{o}C$. The following day, positive staining was visualised using diaminobenzidine (DAB) at 0.5 mg/ml in fresh TNS with 12 µl hydrogen peroxide (brown). Once sufficiently stained, the reaction was ended by diluting in TNS, and washing three times in TNS and a further two times in TBS. Next, sections were mounted on gelatinized microscope slides, and left to dry overnight. Once dry, they were dehydrated in increasing levels of alcohol (70%, 95%, and 100% IMS) for 5 minutes each, then cleared in xylene for at least 5 minutes. Finally, coverslips were mounted on slides using DPX.

2.4.3 Immunofluorescence

For each stain conducted, 1 in 12 series of brain sections were washed with TBS (pH 7.4), then permeabilised with 10% hydrogen peroxide and 10% methanol in ddH_2O for 5 minutes. Sections were then washed three times with TBS, and blocked for 1 hour in 3% appropriate normal serum in 0.2% TXTBS (pH 7.4). Next, sections were exposed to the primary antibody (concentrations detailed in Table 2.3) in 1%

appropriate normal serum in TXTBS, and left to bind overnight at room temperature, and washed three times with TBS the following day. From this point onwards sections were kept under low light conditions. Sections were exposed to a secondary antibody at 1:200 in 1% appropriate normal serum in TBS for 2 hours, then washed three times with TBS. They were then mounted on gelatinized microscope slides, and left to dry overnight. Once dry, they were dehydrated in increasing levels of alcohol (70%, 95%, and 100% IMS) for 5 minutes each, then cleared in xylene for at least 5 minutes. Finally, coverslips were mounted on slides using DPX.

2.4.4 Stereological analysis

Sections were visualised and imaged on a Leica DM6 B upright brightfield and fluorescent microscope. Stereological analysis was conducted using Visiopharm Stereology software, by tracing regions of interest (ROIs) across a 1:12 series, and counting cells in random areas within the ROI. A medium sized graft was used to calculate step distance between random areas to achieve an average cell count of 150 cells per animal, and this step distance was used uniformly across all grafts to reduce error between grafts. Viable grafts were identified by positive HuNu staining. Due to considerable migration from the grafted site, both the dense graft core and striatal periphery were measured.

Estimated graft volume (mm^3) was calculated using the following formula:

$$\frac{(\Sigma a * m * f)}{1{,}000{,}000{,}000} = V$$

Where a = area (μm^2), m = section thickness, f = frequency of sampled sections (1:12 = 12), and V is graft volume (mm^3).

Estimated total number of cells was calculated using the following formula:

$$\left(\frac{\Sigma a}{(r * g)} \right) {}^{*n*f} = C$$

Where a = area (μm^2), r = counting frame size, g = number of grids counted, n = total cells counted, f = frequency of sampled sections (1:12 = 12), and C = estimated total cells.

Estimated cell density per mm^3 was calculated using the following formula:

$$\frac{C^* \left(\frac{t}{t+s} \right)}{V} = D$$

Where C = estimated total cells, t = section thickness, s = average size of cell, V is graft volume (mm^3), and D is estimated cell density per mm^3.

2.5 Methylation Analysis

2.5.1 Sample Preparation and Methylation Quantification

Sample DNA was extracted using the QIAamp DNA mini kit (Qiagen) according to the manufacturer protocols. DNA then underwent bisulphite conversion, which converts unmethylated cytosine to uracil. This was achieved using the EZ DNA Methylation Gold™ kit (Zymo), according to the manufacturer protocols. Bisulphite converted DNA was then processed and quantified using Infinium HumanMethylation450™ BeadChips (Illumina) and an iScan system (Illumina). This was done according to the manufacturer protocols, and was conducted independently by Lesley Bates.

2.5.2 Methylation Analysis

Methylation data analysis was conducted in r software (R Core Team) using the chip analysis methylation pipeline (ChAMP; Bioconductor; Morris *et al.*, 2014; Aryee *et al.*, 2014).

2.5.2.1 Quality control, probe filtering, and probe normalisation

All data presented underwent quality control, filtering, and normalisation (QC) before analysis. First, samples with more than 10% failed probes were removed from the data set. Next, probes that failed in the remaining samples were removed from all samples within the data set. Following this, all probes that are associated with non CpG sites (2803 probes), known SNPs (49,804 probes; Zhou *et al.*, 2016), known to target multiple CpG sites (7125 probes; Nordland *et al.*, 2013), or are located on the X or Y chromosomes (10,165 probes) were filtered out of the data set. The data was then manually examined for abnormalities using hierarchical clustering and a heatmap comparison of the 1000 most variable probes. Following this, it was determined that there were 827 probes that were uniquely differentially methylated in SWIFT 2285 derived samples, which were then filtered from analysis (detailed in Chapter 5, 5.2.1). Finally, data was normalised to reduce technical variance, type 2 probe density bias, and type 1 probe enrichment bias, by applying the BMIQ correction (Teschendorff *et al.*, 2012).

2.5.2.2 Non-statistical explorative methods

The complete methylome of samples was explored using unsupervised hierarchical clustering expressed as a dendrogram based on all probes following QC. The cluster points and distances between them were calculated using complete-linkage

cluster analysis as part of the ChAMP package. The most variably methylated probes were explored using heatmaps and multidimensional scaling (MDS) plots. Heatmaps describe the β-value of CpGs with the highest range of β-values across samples or groups. MDS plots were generated by calculating an Euclidean distance matrix between each sample based on the sum differences between CpG β-values, and using that matrix to map samples to 2D space.

2.5.2.3 Differentially Methylated Probe Analysis

Differentially methylated probe (DMP) analysis was conducted using the ChAMP pipeline DMP analysis, an established method which implements the limma package to calculate the *p*-value of CpG β-values using linear modelling. The *p*-value was then corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate correction. The *p*-value correction was set at 0.01 for all DMP analysis, except for the DMP analysis used to determine SWIFT 2285 specific DMPs, for which the *p*-value correction was set at 0.05, as these DMPs would be removed from future analysis, so a more liberal *p*-value would reduce likelihood of later type 2 errors.

2.5.2.4 Gene Enrichment Analysis

Gene enrichment analysis was conducted using DMP analysis outputs. First, to increase the likelihood that genes associated with significant DMPs were functionally different between groups, the DMPs were filtered to only include those with a Δβ-value difference between groups of 0.5, representing a 50% shift in methylation status between groups. Next, the remaining DMPs were sorted into two groups to reflect the shift in methylation status between groups, i.e. more methylated in group A than B, and less methylated in group A than B. These DMPs were then queried for associated genes using the REFSEQ library, to generate gene lists. These gene lists were then input to Enrichr, and relevant libraries were explored for significant associations. Associations with only one gene were disregarded.

2.6 Statistical Analysis

Statistical analysis was conducted in SPSS (IBM) for all non-methylation data, and in R for methylation data. In Chapter 3, qPCR data were analysed using a one way ANOVA and Dunnet's T test post hoc to compare hiPSCs to the hESC control. In Chapter 4, stereological outputs were analysed using one way ANOVAs. Apomorphine rotation data, qPCR data, and cell counts generated from immunocytochemistry were all analysed using a two way repeated measures ANOVA and Sidak's correction for multiple comparisons. For these tests, Week/Day was the repeated within subjects factor, and Group/Line was the between subjects factor. In Chapter 5, DMPs were analysed in ChAMP which uses liner modelling via the limma package (Smyth, 2004; Wettenhall and Smyth, 2004), with an FDR of either 0.05 or 0.01 as described in section 2.5.2.3. Gene enrichment data was analysed using Enrichr (Chen *et al.*, 2013; Kuleshov *et al.*, 2016), which uses an adapted form of Fisher's exact test to determine the significance of term enrichment, and control for multiple comparisons.

Chapter 3: The Generation and Validation of hiPSCs from Human Fetal Tissues

3.1 Introduction

The generation of authentic MSNs from pluripotent cell sources has various scientific and clinical applications, including the study of basal ganglia development, related disease mechanisms, drug discovery, and as a source of transplantable MSN progenitors in CRT for HD. However, despite significant advances in our understanding of the developmental cues MSNs undergo during development, we are yet to observe reliable functional recovery in animal models of HD grafted with MSNs derived *in vitro* from pluripotent cell sources, leading us to question their authenticity (Introduction 1.4.6).

hiPSCs are one such pluripotent cell source, and can theoretically be derived from any somatic cell in the body. Since their discovery, a body of work has built indicating that at early passages iPSCs retain some degree of the epigenome associated with the cell type from which they were derived, known as an epigenetic memory of tissue of origin (Kim *et al.*, 2010; Polo *et al.*, 2010). This partially retained epigenome has been shown to influence the differentiation and fate specification of iPSCs towards phenotypes associated with their tissue types of origin (Introduction 1.5.4, Table 1.1). It therefore follows that differentiation of pluripotent cells towards an MSN phenotype *in vitro* could potentially be enhanced by deriving hiPSCs from tissues already patterned towards an MSN fate.

To explore this possibility, it is necessary to first consider which tissues are most suitable for the derivation of hiPSCs which may be primed to develop an MSN fate. Logically, terminally differentiated MSNs would be the best candidate, as they carry authentic and mature MSN epigenomes. However, sourcing viable primary human adult MSNs would be extremely challenging, as would maintaining them in sufficient numbers *in vitro*. Alternatively, human fetal tissue is comparatively more accessible and easier to culture than adult MSNs, and the whole ganglionic eminence (WGE) of the human fetal brain is capable of developing mature MSNs both *in vitro* and following transplantation *in vivo* without additional guidance (Watts *et al.* 2000).

The WGE is a heterogeneous tissue, and striatal MSN progenitors first arise within the LGE, thus this is the fetal structure most likely to harbour an advantageous epigenome for MSN differentiation (Introduction 1.2.2; Deacon *et al.*, 1994; Olsson *et*

al., 1995). However, it is also important to bear in mind that the LGE progeny can be further divided into subtypes; of particular note, the dorsal $Sp8^+$ and $Er81^+$ LGE progeny have been found to migrate throughout human development (up to gestational week 24) to both the cortex and the olfactory bulb, and lack striatal specific gene expression patterns (Stenman *et al.*, 2003; Ma *et al.*, 2013).

The MGE neighbours the LGE and thus is exposed to many similar developmental cues, but it progeny are largely interneurons and glial cells that from early stages exhibit a unique gene expression pattern that distinguishes them from the LGE (Introduction 1.2.2; Parnavelas, 2000; Ma et al., 2013). Like the LGE, the MGE contributes cells to the basal ganglia and olfactory bulb, but its progeny also migrate from the subventricular zone to the cerebral cortex (Brazel *et al.* 2003; Ma *et al.*, 2013). Cells arising in the MGE express genes associated with early subpallium patterning, including the DLX and GSX gene families, both of which are key genes expressed during early stages of MSN fate patterning. Therefore, whilst in typical development the MGE does not give rise to large bodies of MSNs, hiPSCs derived from this tissue may still be epigenetically primed to differentiate towards a fate suitable for *in vitro* MSN specification.

Having established two tissue types that will hypothetically confer a beneficial epigenome on hiPSCs for differentiation towards an MSN phenotype, it is important to validate any effect of tissue of origin by deriving hiPSCs from tissues that will hypothetically confer epigenomes which are less suitable for an MSN differentiation. The cells in the developing cerebral cortex separate from the structures of the ganglionic eminences early during development, and do not contribute to the formation of the adult basal ganglia structures. However, these cells are still derived from ectodermal progenitors and are largely neuronal, therefore hiPSCs derived from these tissues may still give rise to neuronal phenotypes more readily than hiPSCs derived from non-neural tissue types, but will not have been subject to the same WGE regional specification cues during their development. By comparing hiPSCs derived from fetal LGE, MGE, and cortical tissues, it should be possible to make a preliminary assessment of the degree to which differences in the epigenome of adjacent fetal brain regions affect the generation of a specific mature phenotype, and specifically whether if any retained epigenome from these structures influences differentiation of an improved mature MSN phenotype.

Conversely, dermal fibroblasts are descended from a different germ layer (mesoderm) to LGE, MGE and cortex (ectoderm); a lineage distinction made during gastrulation, one of the earliest features of embryogenesis (Leptin, 2005). Assuming the retained epigenome has a significant impact on differentiation, hiPSCs derived from

fibroblasts should be better suited to mesoderm lineage differentiations and should therefore have an epigenome that does not enhance, or even potentially hinders their differentiation efficiency towards an MSN phenotype. By including them in a comparison with hiPSCs derived from LGE, MGE and cortical tissues, it should be possible to determine how much of an effect the residual epigenome of fetal derived hiPSCs has on an MSN differentiation, or if it is perhaps negligible. Additionally, due to their accessibility, fibroblasts are one of the most widely reported sources for iPSC generation. By using hiPSCs derived from fibroblasts as a negative control, any results from this work will have implications for a wide range of hiPSC-related studies.

Taking into consideration that outlined above, by generating hiPSCs from human fetal LGE, MGE, cortex and fibroblast tissues (collectively LMCF), any effect of a residual epigenome maintained through induction to pluripotency on *in vitro* MSN differentiation should be systematically observable. However, to add confidence to this comparison, it is also important to minimise other factors that could alter cell behaviours. First, it will be essential to include MSNs derived from hESCs alongside those derived from hiPSCs to ensure any observed effects are due to the retained tissue specific epigenome, rather than any benefits or difficulties that arise from using hiPSCs instead of hESCs. Furthermore, as the majority of *in vitro* MSN differentiations are conducted using hESCs, by including a hESC comparison any findings from this project will be more relevant and comparable to previous works. Next, by generating hiPSCs from tissues from the same fetal donor, all subsequent lines will be genomically identical, thus removing genomic variation as a variable between those lines. Finally, the various methods of iPSC induction were considered with regards to their potential effect on the cell's epigenome. The most common methods of generating iPSCs rely on the use of viruses that integrate directly at random sites across the infected cell's genome to deliver reprogramming factors. Whilst this is an efficient method of induction, considering the aims of this work and the planned use of genomically identical tissues, the direct genomic interaction is undesirable. Random inclusions of reprogramming transgenes across the genome could alter the cell lines in unpredictable ways leading to downstream effects on cell behaviour. Furthermore, it is feasible (although not yet proven), that direct interaction with the cell's genome during reprogramming could disrupt the epigenetic apparatus we are intending to preserve and study. To circumvent these potential issues, it is important to use a non-integrating method of induction. The Sendai virus (SeV) is a negative-sense single-stranded RNA virus, that is restricted to the infected cell's cytoplasm, thereby preventing direct interaction with the host cell's genome. Additionally, commercially available SeV kits are replication-deficient and have a temperature sensitive mutation that breaks down the machinery at temperatures routinely used to culture human cells. This allows the easy deactivation and removal of

viral vectors from infected cells by exploiting typical cell culture temperatures and the virus's cytoplasmic nature.

3.1.1 Aims

In this Chapter I aimed to derive and validate hiPSC lines from tissues with promising epigenetic backgrounds for subsequent MSN differentiations, so as to facilitate the future exploration of the epigenetic memory of tissue of origin retained by these hiPSCs. To achieve this, I generated hiPSCs from human fetal-LMCF tissues, harvested from the same fetal donor (to maintain genomic unity across the lines), using an RNA restricted Sendai virus to deliver Oct3/4, SOX2, Klf4 and c-Myc (OSKM) factors to the target tissue populations. Following induction, hiPSC lines generated underwent a pluripotency assay to confirm reprogramming to a pluripotent state, and the methylation status of these lines was examined to test whether there is epigenetic variation between the cell lines indicative of an epigenetic memory of tissue of origin.

3.2 Results

3.2.1 The Sendai virus is able to infect primary human fetal neural tissues

Whilst the infection mechanisms of the SeV make it a desirable choice for factor delivery within the remit of this work and it is a well validated method to infect fibroblast cultures, there was no published evidence that it could be used to deliver OSKM factors into human fetal neural tissues. To assess the feasibility of delivering OSKM factors into fetal neural tissues using a SeV, I first attempted to infect tissues of interest (LMCF) with a GFP expressing SeV reporter virus and used the opportunity to test different multiplicities of infections (MOI).

Primary fetal LMCF tissues were collected from SWIFT 2242 (age p.c.: 50 days, CRL: 25 mm; Methods: 2.1.2). Each tissue was dissociated and cultured across five wells of a twenty-four well plate under primary fetal tissue monolayer conditions (Methods: 2.1.3) for 3 days. Following this, three wells from each tissue type were infected with the reporter SeV (Methods: 2.1.4). One well from each tissue type was infected at an MOI of 3, 5, and 9 respectively. A GFP signal was visible in all tissues under each MOI condition as early as 1 day following infection, and across all cell types the strongest GFP signal was visible in cells infected at an MOI of 9, and the weakest in those infected at an MOI of 3. GFP expression peaked on day 4 following infection, showing wide coverage across all infected samples, particularly of the conditions with an MOI of 5 and 9. This demonstrates the suitability of the SeV to introduce transgenes to fetal LMCF tissues (Figure 3.1, data shown for LGE tissues under MOI 5).

Although not quantified, there was no observable impact on cell health under any of the MOI conditions, indicating the potential feasibility for all tested MOIs for future SeV work. Due to the satisfactory coverage observed in the MOI of 5 condition 4 days following infection for all the cell types tested (Figure 3.1) an MOI of 5 was chosen for future hiPSC generation using the SeV-iPS construct.

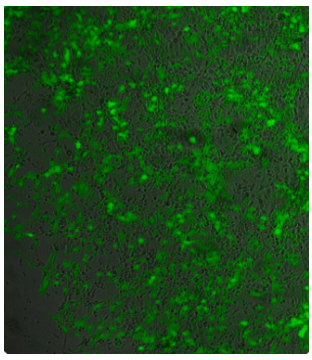


Figure 3.1 Infection of LMCF tissues with a GFP reporter SeV
LGE fetal tissues four days post infection with a GFP expressing SeV (MOI: 5). Positive GFP expression was detectable in ≈50% cells (estimated by eye) indicating successful uptake of the viral machinery.

3.2.2 Generation of hiPSCs from human fetal tissues using a Sendai virus

Following the success of the GFP expressing SeV reporter virus, I proceeded to generate hiPSCs using a SeV to deliver OSKM factors into fetal LMCF tissues. Primary fetal LMCF tissues collected from SWIFT 2285 (age p.c.: 61 days, CRL: 42.1 mm) were dissected, dissociated and cultured across 6 wells of a 24 well plate under primary fetal tissue monolayer conditions (Methods: 2.1.3) for 5 days before infecting with SeV vectors delivering OSKM factors at an MOI of 5 for all viral vectors (Methods: 2.1.4).

As early as five days post infection notable morphological shifts were observed within all infected cultures as individual cells began to swell and become more uniformly circular (Figure 3.2 A-D: black arrows), a feature that was not present in uninfected control cultures (Figure 3.2 E-H). From six days post infection, some of the aforementioned cells had begun to divide, forming immature colonies that started to resemble pluripotent colonies exhibiting traditional pluripotent morphology (Figure 3.2 B-D: red arrows). New colonies continued to emerge the following week across all infected cultures.

Between days 8-14 post infection these colonies grew to sizes sufficient to enable picking for individual culturing and screening for successful reprogramming (Figure 3.3 A). 24 separate colonies were picked from each tissue type for further selection, although there were many more colonies generated in all tissue types.

3.2.3 hiPSCs derived from human fetal tissues exhibit morphology typical of pluripotent cells

The initial identification of successfully reprogrammed cells was made by observing the morphology of the newly generated cells across passages 3-6, systematically selecting for morphological traits typical of hPSCs (Introduction 1.4.1). Cell colony formation was examined between passages 3-6 to identify lines that produced flat and rounded colonies, with tightly packed cells and low levels of intracellular space (Figure 3.3 B); and individual cell morphology within those colonies was examined to identify lines composed of cells with a circular cytoplasm and high nucleus to cytoplasm ratio (Figure 3.3 C). Lines were removed if they exhibited inconsistent or aberrant features, in brief: doming colonies; irregular and unpredictable colony shape and growth; regular spontaneous differentiation; high levels of apoptosis; highly variable individual cell morphology; or an inconsistent colony periphery (Figure 3.3 D).

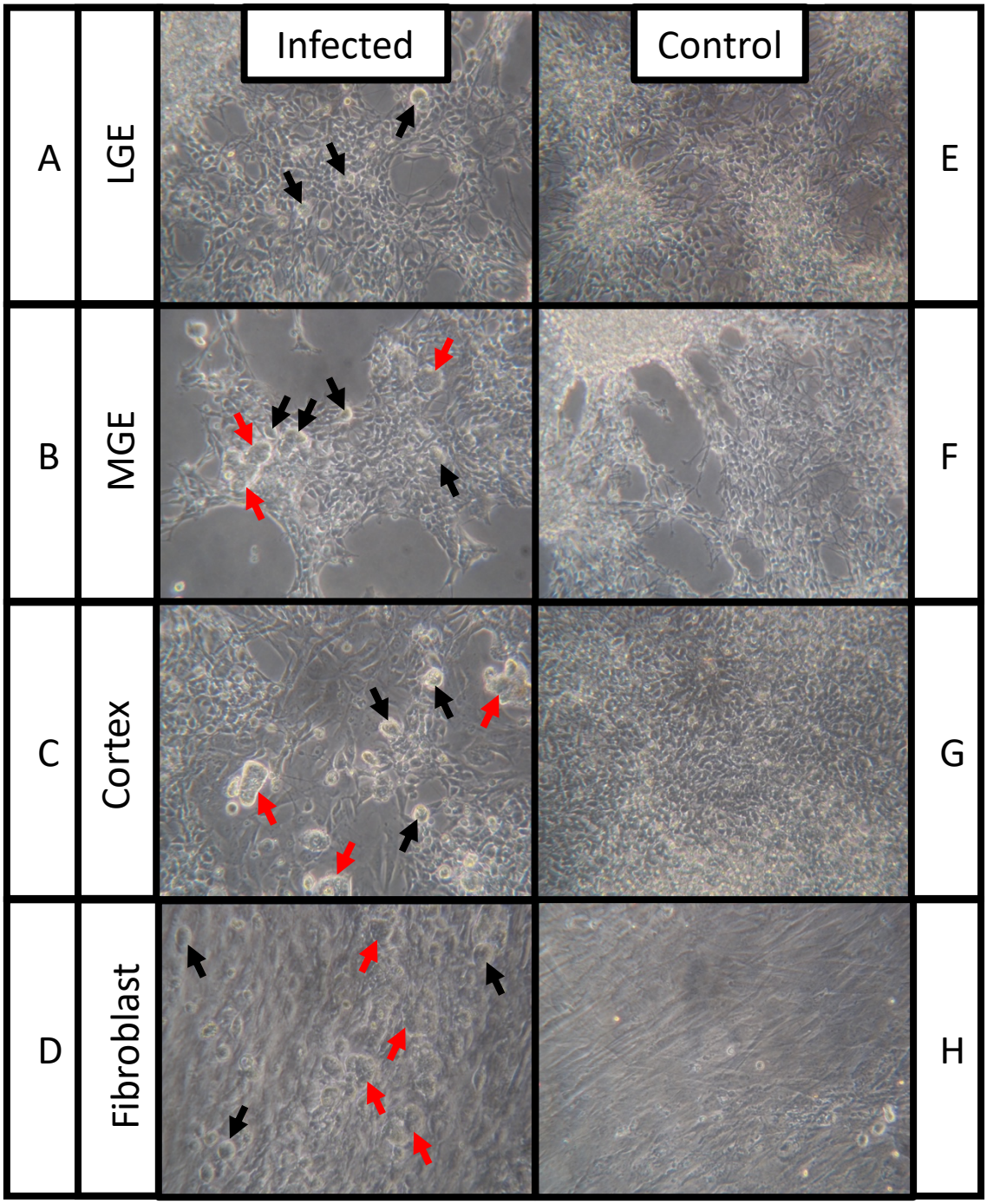


Figure 3.2: The Generation of hiPSCs from human fetal LMCF tissues.

A-D. Brightfield images of LMCF fetal tissues six days post infection with OSKM expressing SeV vectors. Infected cells are seen exhibiting morphological traits typical of cells undergoing reprogramming. Notably the early swelling cells (indicated with black arrows) that later divide into emerging hPSC-like colonies with individual hPSC-like cells (indicated with red arrows).
E-H. Brightfield images of control uninfected LMCF fetal tissues cultured under the same conditions, photos taken at the same time point.

Most of the derived lines did not pass all aspects of this selection criteria, but following this observational period there were at least three lines from each tissue type reprogrammed that exhibited traditional hPSC morphology. Of these lines, 3 LGE-derived hiPSC lines were taken forwards for further characterisation (L1, L2 and L11), additionally 1 MGE-derived hiPSC line (M3), 1 Cortex-derived hiPSC line (C6), and one Fibroblast-derived hiPSC line (F1) were also taken forwards as genomic and tissue origin controls.

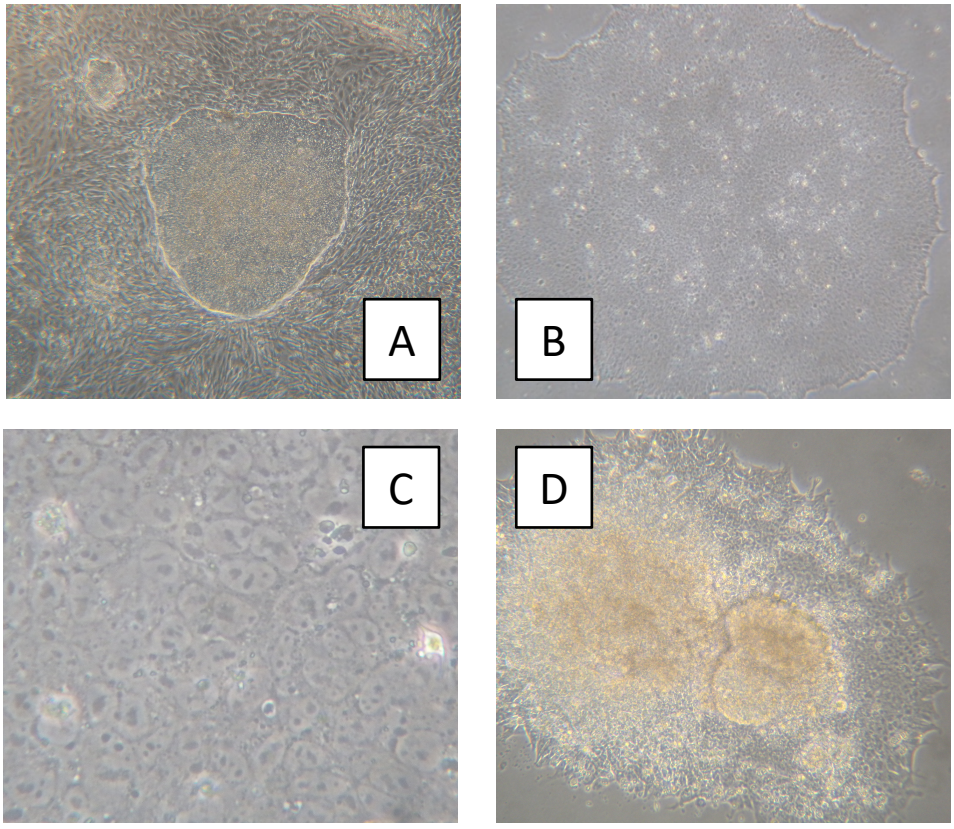


Figure 3.3: Morphological Characterisation of hiPSCs derived from fetal LMCF tissues.

A. Brightfield image of infected fibroblast culture prior to picking sample colonies, hiPSC colonies exhibit a clearly defined border between colonies and surrounding unaffected cells.
B. Brightfield image of LGE derived hiPSC colony (L2) containing tightly packed cells exhibiting a scant rounded cytoplasm with a high nucleus to cytoplasm ratio and prominent nucleoli.
C. Brightfield image of MGE derived hiPSC colony (M3) exhibiting a flat and rounded shape consisting of tightly packed but individually definable cells with low levels of intracellular space.
D. Brightfield image of fibroblast derived hiPSCs (F10) that was later removed from the experiment for its atypical phenotype. Examples observable here include a tendency to dome at the center of the colony, variable cell size and morphology across the colony, borders appearing within the colony center indicating spontaneous differentiation and high levels of apoptosis across the colony body.

3.2.4 hiPSCs derived from human fetal tissues express hallmark pluripotency markers

Next, pluripotency was confirmed in each of these novel lines by examining their expression of genes associated with pluripotency, and comparing results to a human ESC control line (H9). L1, L2, L11, M3, C6, F1, and H9 cells were cultured as hPSCs (Methods: 2.1.5). Samples were processed for immunocytochemistry, RT-PCR, and qPCR analysis (Methods 2.2.1-2.2.2).

hPSC cultures between passage 9-13 (hiPSC) and a passage 18 (hESC) were fixed and underwent immunocytochemical staining. Fluorescent microscopy revealed that all derived hiPSC lines expressed NOGGIN, OCT4, SOX2, TRA-1-60, and SSEA3/4 in a similar manner to the hESC control (Figure 3.4A-E, data shown for hiPSC lines L1, M3, C6, and F1 only). SOX2 was co-stained with the ectodermal marker NESTIN to confirm that any observed SOX2 expression was indicative of pluripotency rather than ectodermal commitment. Although low level expression of NESTIN was observed in the periphery of all hiPSC colonies, it did not exceed that observed in the hESC control, and the locus of expression was distinct from that observable in committed neuroectodermal cells (see Figure 3.6E).

Following this, RT-PCR was used to confirm the immunocytochemical results. RNA was extracted from hPSC cultures at passage 9-12 from hiPSC lines and passage 16 of a H9 hESC control line. Gel electrophoresis of PCR products confirmed the expression of hallmark pluripotency genes OCT4, and SOX2 in all samples (Figure 3.5A).

Finally, qPCR was used to quantify any differences in expression of these hallmark pluripotency genes between the hiPSCs and control hESCs. Cells from three separate cultures from each of the lines, L1, L2, L11, M3, C6, F1, and H9, were harvested between passages 6-16 (hiPSCs) or 16-20 (hESCs H9). QPCR was performed and relative expression of each gene analysed was calculated according to H9 expression levels using the ΔΔCt method (Methods 2.2.2.3). A one way ANOVA was conducted, and no differences were observed in OCT4 expression across all lines regardless of cell origin ($F_{6,20}$=1.89, p=0.154; Figure 3.5B). However, whilst SOX2 was positively expressed in all lines, there was a significant difference in relative SOX2 expression between the lines ($F_{6,20}$=3.03, p=0.041), and a post hoc Dunnett's T test comparing hiPSCs to the H9 line revealed that the MGE derived line M3 expressed significantly lower levels of SOX2 than the hESC control line (relative fold difference of 0.44 ± 0.03, p=0.026; Figure 3.5C).

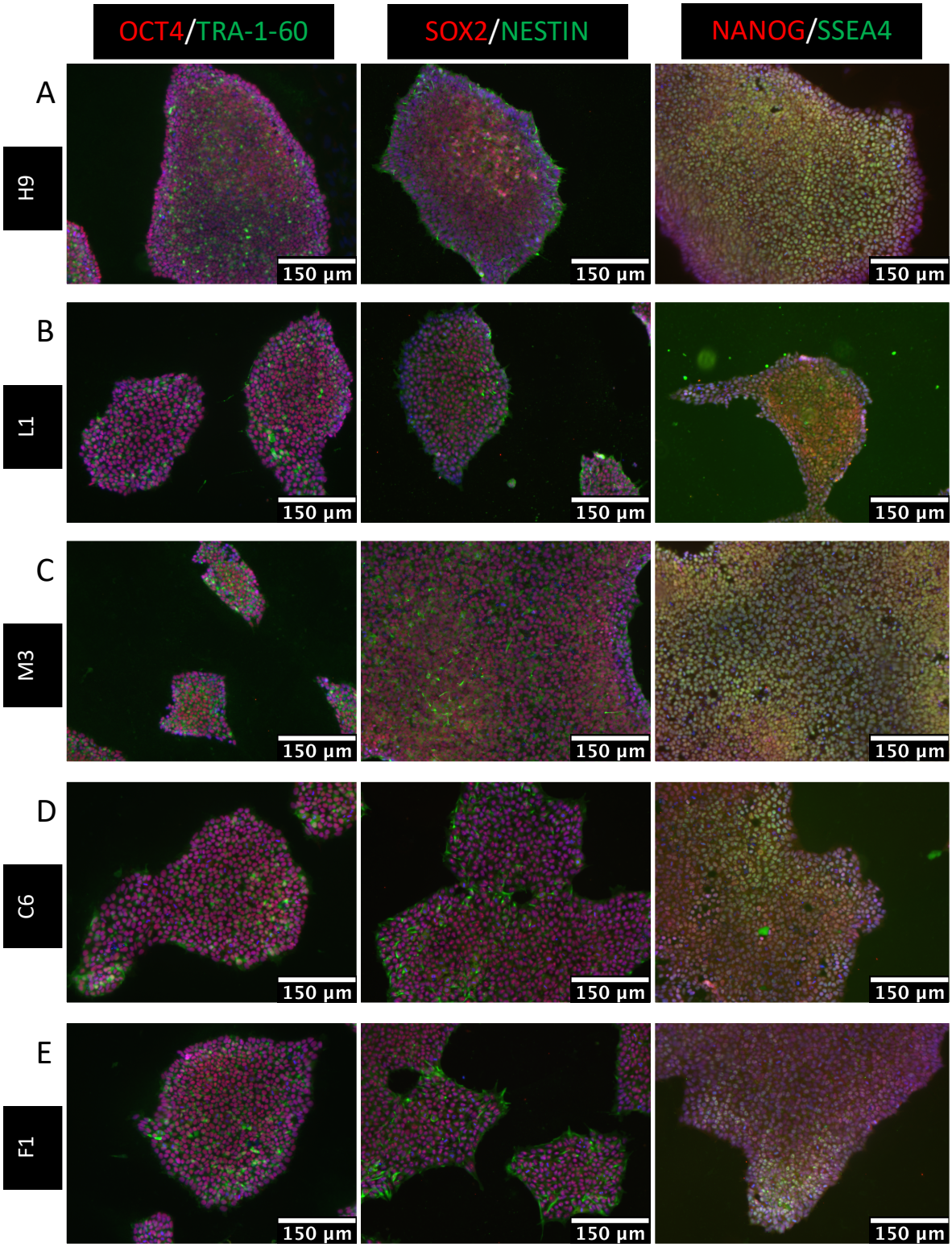


Figure 3.4: hiPSCs derived from fetal LMCF tissues exhibit hallmark pluripotency markers in a similar pattern to ESCs.

A-E. Immunocytochemical fluorescence staining for OCT4 (left, red), TRA-1-60 (left, green), SOX2 (centre, red), NANOG (right, red), and SSEA4 (right, green) is present in ESC line H9 (A), and LMCF fetal tissue derived hiPSC lines L1 (B), M3 (C), C6 (D), and F1 (E). SOX2 is co-stained with NESTIN (centre, green).

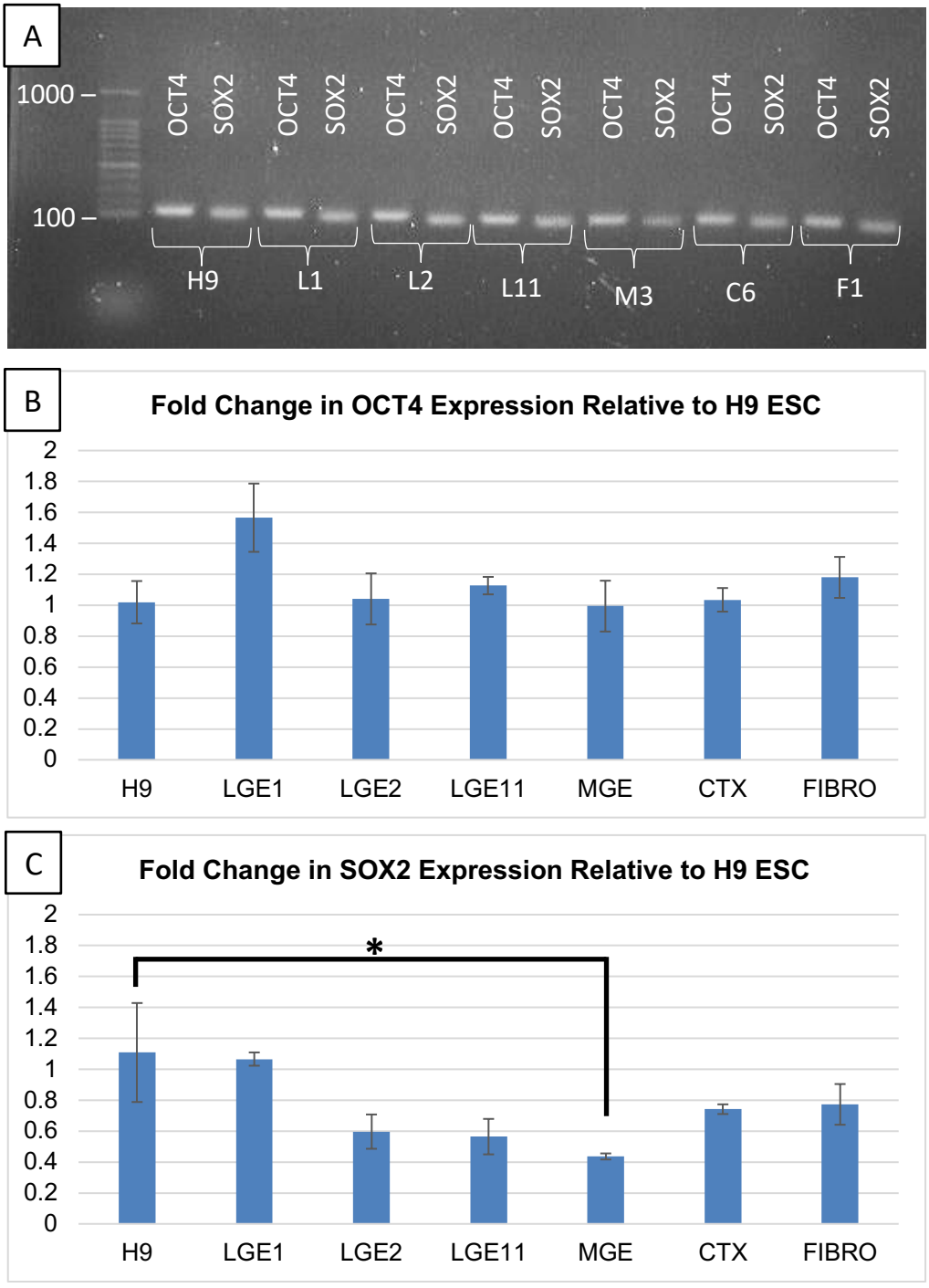


Figure 3.5: PCR analysis of pluripotency markers OCT4 and SOX2 in hiPSC derived lines compared to an ESC control.

A. RT-PCR product visualized following 30 amplification cycles using gel electrophoresis indicates both OCT4 and SOX2 are expressed by all lines.
B-C. Q-PCR analysis of OCT4 (B) and SOX2 (C) expression relative to H9 hESC control, data shown (n=3 for each group) taken at passages ranging between passage 6 and 16 (hiPSCs), or 16-20 (hESC). *p<0.05, post hoc Dunnett's t test.

3.2.5 hiPSCs derived from human fetal tissues are functionally pluripotent

A fundamental aspect of pluripotency is the capacity to differentiate towards all three germ layers of the developing embryo. Once the expression of pluripotency markers was confirmed in these lines, their functional capacity for pluripotency was characterised. hiPSC cultures of L1, L2, L11, M3, C6, and F1 (all <12 passages) were allowed to spontaneously differentiate for 5 days in culture (for mesoderm and endoderm fate commitment, Methods 2.1.6), or underwent the initial 9 days of a MSN differentiation protocol (for ectoderm fate commitment, Methods 2.1.7, Protocol 2), before being processed for immunocytochemical analysis (Methods 2.2.1).

Immunocytochemical staining was used to identify cells committed to each of the three germ layers. The expression of DESMIN (Figure 3.6A, green) and Alpha-Smooth Muscle Actin (α-SMA Figure 3.6B, red) was used to identify cells that had begun to differentiate towards a mesodermal lineage. Expression of Alpha-Fetoprotein (α-FP, Figure 3.6C, red) and VIMENTIN (Figure 3.6D, green) was used to identify cells that had committed to an endodermal lineage. Finally, NESTIN (Figure 3.6E, green) and βIII-tubulin (Figure 3.6F, green) expression was used to identify cells differentiating towards an ectodermal lineage. These six markers were identifiable in all hiPSC derived cultures (Figure 3.6A-F, data shown for L1 only).

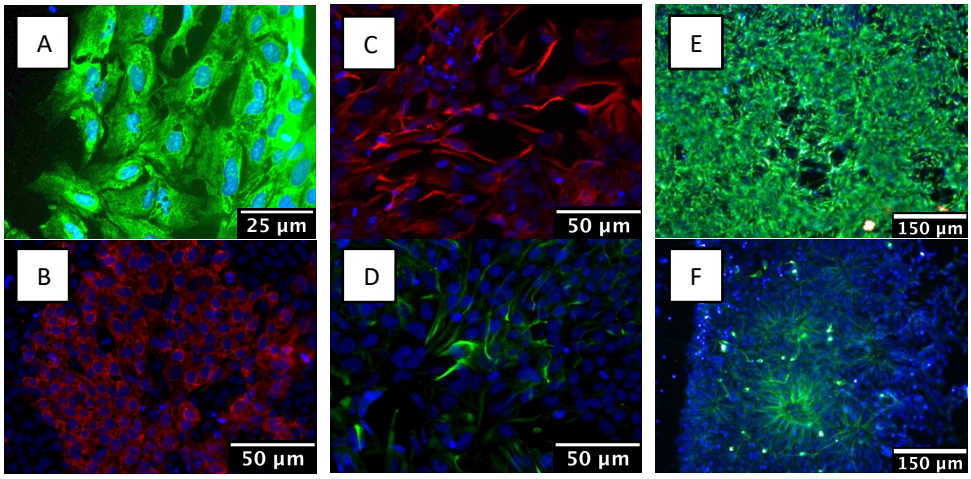


Figure 3.6: hiPSCs derived from fetal LMCF tissues can differentiate towards all three germ layers.

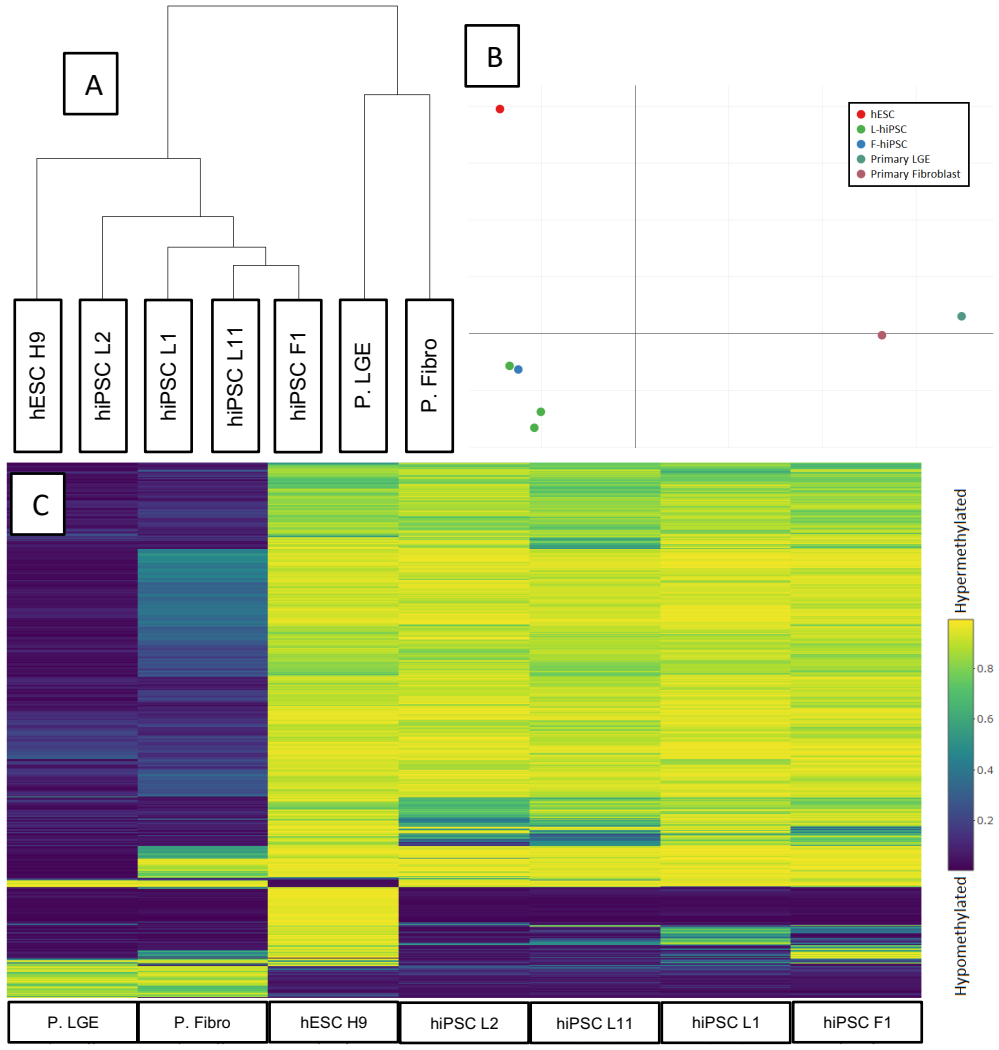
A-F. Immunocytochemical fluorescence staining of hiPSC line L1 following 5 days of spontaneous differentiation (A-D) or 11 days of an MSN differentiation protocol (E-F). The hiPSCs exhibit a capacity to differentiate towards all three germ layers, as demonstrated by positive expression of mesodermal markers DESMIN (A, green) and α-SMA (B, red); endodermal markers α-FP (C, red) and VIMENTIN (D, green); and ectodermal markers NESTIN (E, green) and βIII-tubulin (F, green). In all cases Hoechst staining is conducted in blue. Cells in E are co-stained with OCT4 (red). Note, different magnifications have been used between images, to emphasise locus of expression.

3.2.6 Profiling the epigenetic methylation landscape of hiPSCs derived from human fetal tissues

Finally, having generated and validated LMCF derived hiPSCs, it remained necessary to explore whether early passage hiPSCs exhibited different methylation profiles, potentially reflective of their tissue of origin, and to observe whether this separated them from hESCs. Seven samples were processed for genome wide methylation profiling: hPSCs from the hiPSC lines L1, L2, L11, and F1; a sample of the primary LGE and fibroblast tissues that were used to derive these lines; and a sample of H9 hESCs. For resource reasons, only a single sample could be included for each line, so statistical analysis was not appropriate. These samples were processed and analysed using Illumina's Infinium Human Methylation 450K system, and the resulting data was then prepared and analysed using the ChAMP package available from Bioconductor (Methods 2.5).

Two datasets were generated during the analysis. The first contained all seven samples listed above, which after QC (Methods 2.5.2.1, including removal of SWIFT 2285 specific methylation as discussed in Chapter 5) included 411,852 probes in the analysis. The data set was used to generate a dendrogram describing unsupervised hierarchical clustering of the samples based on the methylation status of all probes, and revealed that the samples primarily cluster by cell type (hiPSC, hESC, and primary tissue/cells), but that the hiPSCs were more similar to the hESCs than the primary tissue samples (Figure 3.7A). This distinction is further consolidated by comparing a heatmap of the 1000 most variable probes between all seven samples (Figure 3.7C), which demonstrates that much of this variance is due to higher methylation in the hESC sample than in all hiPSCs samples, and conversely lower methylation in primary tissues samples. An MDS plot calculated using those same 1000 most variable probes indicated again that cell type was the most important feature when clustering samples (Figure 3.7B).

To examine the methylation differences between the hPSCs without the noise of primary tissues, the second dataset generated only included the five pluripotent samples, and following QC 412,331 probes were included in the analysis. The second dataset was used to generate a new MDS plot based on the 1000 most variable probes within the pluripotent samples, which clearly clusters the LGE derived hiPSC lines L1, L2, and L11 together, and away from both the fibroblast derived hiPSC line F1 and the hESC line H9. This indicates that whilst hiPSCs appear generally similar compared to hESCs and primary tissue, there are epigenetic differences between these samples that result in separate clustering patterns that group hiPSC lines by tissue of origin (Figure 3.7D).



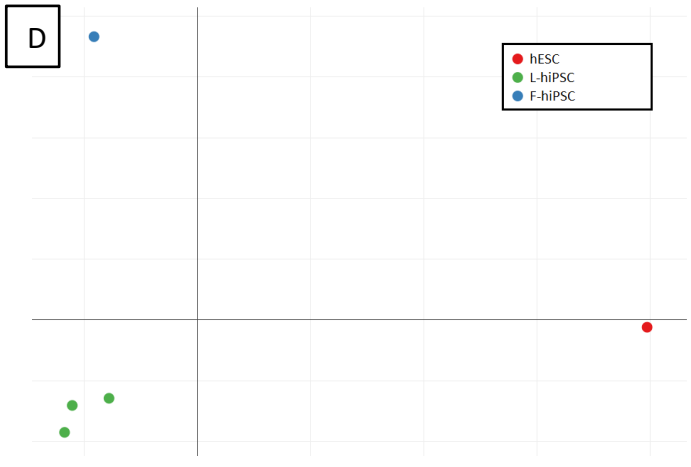


Figure 3.7: DNA Methylation Cluster Analysis

Single samples of SWIFT 2285 Primary LGE tissue (P.LGE); SWIFT 2285 Primary Fibroblast tissue (P.Fibro); hiPSC lines L1, L2, L11, and F1; and hESC line H9.

A. Unbiased hierarchical clustering of samples using entire methylome, indicating samples group by cell type (hiPSC, hESC, or Primary).

B. MDS plot of the 1000 most variable CpGs across samples. Samples cluster by cell type (hESC, hiPSC, Primary).

C. Methylation heatmap of the 1000 most variable CpGs across samples, showing low methylation in primary tissues, high methylation in hESCs, and that hiPSCs share features of both.

D. MDS plot of the 1000 most variable CpG sites between hPSCs only, indicating that hPSCs cluster according to their source tissue.

3.3 Discussion

3.3.1 The Sendai virus can successfully deliver OSKM factors to fetal LMCF tissues to induce pluripotency

The SeV is a well validated method of transducing pluripotency in a variety of tissue types (Ban *et al.*, 2011; MacArthur *et al.*, 2012), although to date there is no published work describing hiPSC generation from fetal LGE, MGE or cortical tissues using this method. Here, I have demonstrated the capacity of the RNA based SeV to deliver OSKM factors to these cell types, thereby avoiding the issue of genomic integration. During the induction process, colony formation was observed as early as 5 days post infection, a notably shorter generation time than the advised time of 14 days given for adult fibroblast tissues. It is possible this is at least in part due to the young age of the tissues (i.e. fetal as opposed to adult cells) as it is understood cell age and DNA damage can inhibit reprogramming (Marión *et al.*, 2009) and that the majority of cells infected would have been progenitors, not yet terminally differentiated towards their final stage. It is also unclear how variable generation times are between tissue types and across ages, as due to the wide spread establishment of iPSC generation, such details are rarely reported. However, as the speed of generation was not a focus of this work, the mechanisms underpinning this finding could not be investigated for reasons of time.

Following colony formation, 96 individual pluripotent-like colonies were isolated for potential use (24 from each tissue type), and examined for desirable morphological traits indicative of authentic hiPSCs to remove partially reprogrammed samples. Authentic hPSC morphology includes a scant rounded cytoplasm containing a large pale nucleus and prominent nucleoli, resulting in a high ratio of nucleus to cytoplasm, and the cells should be individually distinguishable within a tightly packed but flat and uniform colony (Thompson *et al.*, 1998). Using this criteria, three LGE tissue derived lines and one line derived from each of the MGE, cortex and fibroblast tissues were selected for further characterisation.

Morphological characterisation is not considered robust enough to establish whether an iPSC line has completely reprogrammed to a pluripotent state, and must be followed up with additional gene expression and functional assays (Chan *et al.*, 2009). Immunocytochemistry confirmed that all 6 hiPSC lines exhibited positive expression of key pluripotency regulators OCT4, SOX2 and NANOG in a similar fashion to the hESC control line, and of reliable pluripotency surface markers SSEA4, and TRA-1-60 that are essential to determine complete hiPSC reprogramming (Chan *et al.*, 2009). Whilst there was evidence of the intermediate filament protein NESTIN in the periphery of all

hiPSC cultures, the amount and locus of expression was not observed to differ from the hESC control line, indicating that this was within acceptable levels of authentic pluripotent populations. This was further corroborated by comparing hiPSC cultures to cells found in hESC cultures that expressed both NESTIN and SOX2 in a morphologically distinct manner to authentic hESCs. OCT4 and SOX2 expression was then validated using RT-PCR for all lines, and then quantified using RT-qPCR relative to the hESC control. OCT4 expression was highly consistent between hiPSC lines and the hESC control line, with only hiPSC line L1 seeming to express slightly higher levels, although this was not significantly different. However, SOX2 tended to be expressed at a lower level in hiPSC lines (observable in L2, L11, M3, C6 and F1) and was significantly lower in the MGE derived hiPSC line M3 compared to the hESC control line.

Along with OCT4 and NANOG, SOX2 expression is required to maintain pluripotency in hPSCs (Niwa, 2007). Its functional purpose is to regulate and maintain consistent OCT4 expression levels (Masui *et al.*, 2007, Fong *et al.*, 2008), which in turn regulates SOX2 expression in hPSCs in tandem with NANOG as part of a regulatory circuit of pluripotency (Boyer *et al.*, 2005). Deviations in this triad of gene expression equilibrium in both directions are known triggers for differentiation (Boer *et al.*, 2007, Kopp *et al.*, 2008), which suggests that lower levels of SOX2 expression in hPSCs might be indicative of imperfect reprogramming. However, in 2012 Wakao *et al*. characterised ≈800 colonies generated across two experiments and demonstrated that a key distinction between completely and incompletely reprogrammed hiPSCs is the positive expression of SOX2, rather than variant levels of expression. Additionally, it is not yet established how much variation in expression of pluripotency genes can be expected in hPSCs before they lose pluripotency, and there is evidence that expression of pluripotent markers can vary drastically between authentic pluripotent cells. For instance, in 2003 Carpenter *et al.* reported varying levels of SSEA4 and CD9 expression across four ESC lines (H1, H7, H9 and H14) without loss of pluripotency.

For cell line M3, SOX2 expression was detected on average only 1.2 (± 0.1) PCR cycles later than in the H9 control, and for all hiPSC lines SOX2 expression was clearly identifiable in all conditions, indicating that it was still present and consistently expressed in all hiPSC lines tested. Furthermore, there were no observable differences in OCT4 expression between hiPSC lines that expressed lower levels of SOX2 (L2, L11, M3, C6, F1) compared to the hESC control, and the only cell line to deviate from this general trend was L1, which tended to express higher (though not significantly so) levels of OCT4 and very similar levels of SOX2 when compared to the hESC control.

Perhaps the most defining feature of pluripotency is the ability for PSCs to differentiate towards any somatic cell of an organism. This is routinely validated in

mouse-derived iPSCs by forming a chimeric blastocyst with the iPSCs to observe whether they are capable of forming a chimeric animal. As this is not possible for human derived hiPSCs, the gold standard is to instead observe if the cells are capable of differentiating towards all three germ layers that give rise to all somatic cells of the body, using *in vivo* teratoma formation assays and *in vitro* spontaneous and directed differentiation. All lines including M3 readily expressed markers indicative of all three germ layers, demonstrating a functional capacity for pluripotency. Therefore, whilst there were variable levels of pluripotent gene expression observable between hiPSC lines and the hESC control line, these do not necessarily reflect a lack of pluripotent potential.

3.3.2 Genomically identical hiPSCs exhibit DNA methylation profiles potentially indicative of tissue of origin

Previous works have reported variation in hiPSC differentiation capacity linked to the original tissue source they were derived from, and that this variation is somewhat mediated by epigenetic differences between these lines (Introduction 1.5.4). The methylome of single samples from pluripotent hiPSC lines L1, L2, L11 and F1 were compared with pluripotent ESC line H9, and the original primary LGE and fibroblast tissues from which the hiPSCs were derived. Three forms of cluster analysis were used to explore the methylomes of these samples. From this analysis, it is apparent that the most distinct difference between these samples is the type of cell (primary vs hiPSC vs hESC). As seen in the first MDS plot and in the hierarchical clustering dendrogram (Figure 3.7), the samples most similar are the hiPSC lines, with the primary tissues clustering furthest from the other samples, although the primary LGE and fibroblast tissues still remain quite distinct from one another. The hiPSCs have a methylation profile more similar to that of hESCs than to that of the primary tissues from which they were derived, which is driven by a genome wide increase in methylation as demonstrated by the heatmap of the 1000 most variable probes between these samples. Additionally, from this heatmap it is evident that the hiPSCs share methylation characteristics with both the primary tissues of origin from which they were derived and with the pluripotent hESCs that they have become functionally similar to. This suggests that although their methylome has shifted towards that of the hESCs, they clearly still retain some residual epigenetic characteristics of their tissues of origin. Lastly, the final MDS plot, generated by comparing only the pluripotent samples, indicates that a more subtle clustering pattern exists within the samples, that separates the hiPSCs by their tissue of origin (fibroblast and LGE derived). However, due to the small sample sizes and lack of statistical analysis, this result remains provisional.

3.3.3 Conclusions

In this chapter I have demonstrated the successful derivation of hiPSCs from human fetal LMCF tissues from the same donor using a SeV to deliver OSKM transgenes, and have begun testing the observable differences between these lines compared to a true pluripotent hESC control. Collectively, this analysis indicates that the hiPSCs generated in this study have been successfully reprogrammed to a pluripotent state, and are functionally able to differentiate towards any of the three germ layers. Furthermore, there is some evidence suggesting that these hiPSCs have retained some DNA methylation distinctions indicating an epigenetic memory of tissue of origin.

Chapter 4: Characterisation of hPSC-MSNs

4.1 Introduction

CRT for HD is understood to require a transplantable population of striatal MSN progenitors in order to restore the damaged circuitry and thereby restore function (Peschanski, Cesaro and Hantraye, 1995; Celland, Barker and Watts, 2008). There is proof of principle that such interventions can somewhat ameliorate the symptoms of HD, if the transplanted population are authentic MSN progenitors at the correct stage of development (Rosser and Bachoud-Levi, 2012). To date this has been most consistently achieved with fetal WGE, which in typical development is the origin of the adult striatum and is therefore a source of authentic MSN progenitors (Olsson, Björklund and Campbell, 1998; Brazel *et al.*, 2003). There have now been several papers published detailing the directed differentiation of hPSCs towards an MSN fate using morphogenic factors (Introduction 1.4.4). However, to date these cells have not been able to reliably restore motor function in an animal model of HD (Introduction 1.4.5). Precisely why these cells do not restore function remains unclear. One possible reason is that the protocols used for deriving MSN progenitors are not yet sufficient to produce an authentic MSN phenotype, and therefore these cells cannot function as required (Introduction 1.4.6).

Previously within our lab, we have generated hiPSCs from fetal WGE, and demonstrated that they are capable of differentiating towards an MSN fate. Furthermore, there is provisional evidence that these cells produce more DARPP32 positive cells when transplanted *in vivo* compared to similarly cultured hESCs (Choompoo, 2015). It is possible this is due to the established effect of epigenetic memory of tissue of origin that persists in hiPSCs, which has been observed to bias hiPSC differentiation towards cell types similar to the tissues from which they were derived (Kim *et al.*, 2010; Polo *et al.*, 2010; Introduction 1.5.4). Of note, this epigenetic memory has since been explored by several labs in an attempt to enhance differentiation, particularly with regards to ‘hard to culture’ cell types (Bar-Nur *et al.*, 2011; Hiler *et al.*, 2015). We therefore have theorised that hiPSCs derived from fetal LGE tissues (the origin of MSNs; Olsson *et al.*, 1995) may harbour an authentic epigenetic memory of their tissue of origin, that could potentially enhance their subsequent MSN differentiation. However, our previous work did not control for genetic variation between PSC lines, nor did it include a comparison with hiPSCs derived from tissues other than LGE to confirm that the results were due to a retained epigenetic memory of hiPSC tissue of origin (Choompoo, 2015). Therefore this provisional

evidence needs to be validated through a systematic study which controls for these possible variables.

In Chapter 3, I derived hiPSCs from a variety of fetal tissues including LGE (L1, L2, L11), MGE (M3), cortex (C6) and fibroblasts (F1), to attempt to generate hiPSC lines with an epigenetic memory of their tissue of origin. Critically, these were derived from the same fetal donor using a non-integrating method of induction, to ensure the hiPSCs were genomically identical, and that any differences observed between the lines in their differentiation capacity was due to epigenetic differences. In that Chapter, I also provided some evidence that these cells can be distinguished by their DNA methylation profile, which seemed to be related to their tissue of origin. Here, I examine the differentiation potential of these hiPSCs in order to determine if there are differences between these lines that cannot be attributed to genetic variation or cell type (hiPSC vs hESC), using several widely studied measures of characterising hPSC derived MSNs.

Importantly, there are two established general methods of differentiating hPSCs towards an MSN phenotype. The first and most widely used relies on the addition of SHH, which is based on the knowledge that SHH is expressed at increased levels on the ventral side of the developing forebrain, and is known to have a potent, dose-dependent, ventralising effect on the fate commitment of forebrain progenitor cells (Aubry *et al.*, 2008; Ma *et al.*, 2012; Nicoleau *et al.*, 2013; Noakes *et al.*, 2019). To date, the most convincing functional recovery using hPSC derived MSNs has been reported using an SHH based protocol (Ma *et al.*, 2012), however, functional recovery using similar methods has not been found reliably between labs. Furthermore, there have been several cases of graft overgrowth in SHH based protocols (Aubry *et al.*, 2008; Ma *et al.*, 2012; Delli Carri *et al.*, 2013), and high volumes of SHH have been repeatedly demonstrated to induce a fate commitment more associated with the striatal interneurons of the developing MGE (Ma *et al.*, 2012; Nicoleau *et al.*, 2013; Arber *et al.*, 2015 Noakes *et al.*, 2019). More recently, Activin A (Activin), a TGFβ family protein, has been shown to specifically upregulate genes associated with downstream MSN fate patterning, including FOXP1 and CTIP2 (Arber *et al.*, 2015). The exact mechanisms underpinning this interaction is not yet clear, however TGFβ activity can act to promote phosphorylated Smad2 expression, which has been detected in the developing fetal ganglionic eminence (GE) co-localising with the pan-GE marker DLX2, acting to enhance DLX2 expression (Feijen, Goumans and van den Eijnden-van Raaij, 1994; Maira *et al.*, 2010). This method has not yet produced functional PSC derived grafts, but, there is presently only one paper published that uses Activin to direct cells towards an MSN fate. It is therefore difficult to judge if this method is more or less likely to produce functional MSNs than SHH based protocols. However, there is compelling

evidence that Activin acts on different pathways than SHH, and a direct comparison of their ability to specify an MSN fate, indicated that Activin, unlike SHH, could rapidly induce MSN precursor markers (Arber *et al*., 2015). This is consistent with other published work which details that much longer periods of SHH exposure are required to induce an MSN phenotype and prevent graft overgrowth (Aubry *et al.*, 2008), which suggests that SHH instead acts indirectly to facilitate an environment for spontaneous MSN differentiation. As such, I elected to use an Activin based protocol for the characterisation of the differentiation potential of these hiPSCs.

4.1.1 Aims

In this chapter, I aimed to characterise the effect of the epigenetic memory of tissue of origin in hiPSCs derived from primary LGE tissues on their capacity to undergo an MSN differentiation. To achieve this, I differentiated the hiPSC lines generated in Chapter 3 and a hESC control line towards an MSN fate, using a standardised protocol adapted for all hPSCs tested. I compared the capacity of these cells to produce striatal MSN progenitors and neurons *in vivo* and *in vitro*, using a multitude of different measures to explore the cells both during and at the end of the differentiation process.

4.2 Results

4.2.1 Standardising a hPSC-MSN differentiation protocol

In order to systematically investigate the MSN differentiation capacity of the cell lines derived in Chapter 3, it was first necessary to confirm their ability to undergo an MSN differentiation protocol, and if necessary, adapt it to be suitable for all the cell lines being examined to allow for a uniform comparison. Initial attempts were conducted on hiPSC lines L1, M3, F1, and an hESC control line: H9. The H9 line was chosen as the control line because previously it has been established that this hESC line is highly efficient at undergoing a similar differentiation (Arber *et al.*, 2015). First, the cells underwent a monolayer MSN protocol which has been optimised through the Repair HD European Consortium (Repair-HD, 2013). This protocol is based on Arber *et al.*, 2015, and heavily informed by Nicoleau *et al.*, 2013. It also took account of findings published in Chambers *et al.*, 2009 and Boergermann *et al.*, 2010, and is currently being prepared for publication (Protocol 1; Figure 4.1A).

It was quickly determined across three rounds of differentiation that Protocol 1 was ill-suited for hiPSC lines L1 and M3, as both lines persistently spontaneously detached from their culture plates coated with reduced growth factor (RGF) Matrigel by day 8 of the protocol, before passage 1 could be conducted. Trial and error revealed that passage of these cells on day 7 of the protocol, rather than day 9, and plating onto fibronectin coated PDL treated plates was sufficient to maintain these cells for further culturing (Figure 4.1B, P1). Furthermore, both hiPSC line F1 and hESC line H9 tolerated this earlier passage time well. Administration of morphogens was maintained for the originally intended duration according to Protocol 1, and by day 11 of the protocol PSC cultures exhibited widespread expression of the neural progenitor marker NESTIN and LGE progenitor marker GSX2 (Figure 4.1D, data only shown for L1, no differences observed in the other lines).

Around the time of the second intended passage (day 20) additional problems arose. Firstly, there were notably fewer cells in L1 and M3 cultures, which was presumed to be caused by a combination of notably higher levels of cell death and possibly less proliferation. Morphological differences were also apparent between lines at this stage, with L1 in particular exhibiting a mature neuronal phenotype as evidenced by long neurite outgrowth and defined cytoskeletal expression of βIII-tubulin at day 20, compared to both H9 and F1 which were still recruiting βIII-tubulin to the developing cell cytoplasm (Figure 4.1E). Furthermore, during passage 2, the viability of L1 was notably lower than the other lines (L1: ≈55%, F1≈90%; H9 ≈90%). Based on the live cell counts determined using a Trypan Blue exclusion assay, cells were plated at the

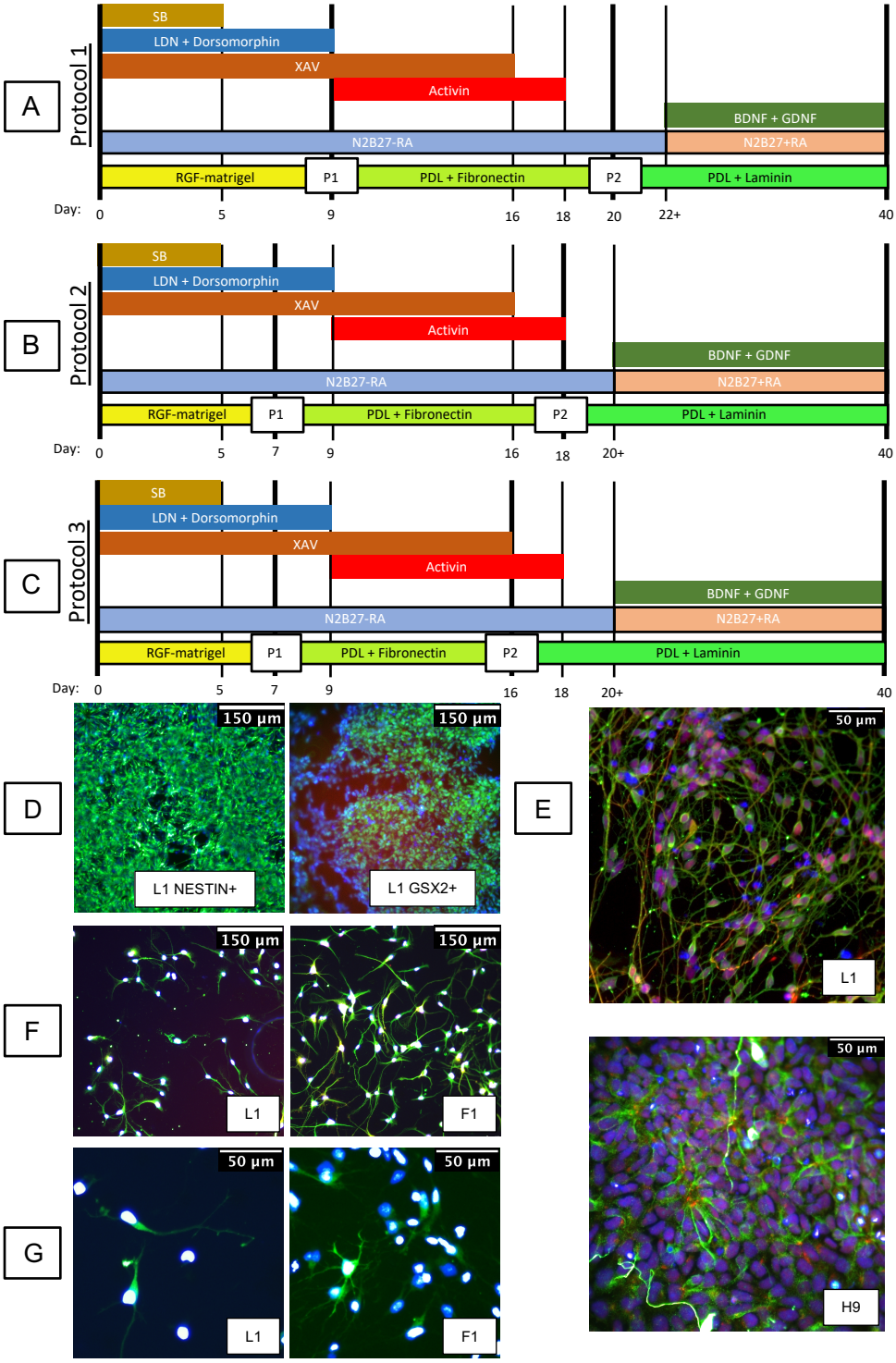


Figure 4.1: Standardisation of an MSN differentiation protocol
A-C. Schematics of the different versions of the MSN differentiation protocol.
D. hiPSCs express early neuronal (NESTIN, left) and WGE/LGE (GSX2, right) markers following Protocol 2. Data shown for L1 only. Green is NESTIN/GSX2, blue is Hoechst. Cells at day 11.
E. L1 (top) exhibited signs of advanced neuronal maturation and less cell proliferation by the end of Activin treatment compared to H9 (bottom). Green is βIII-tubulin, red is FOXP1. L1 at day 20, H9 at day 18. No differences observed between H9 and F1 at this time.
F. L1 demonstrated poor survival compared to F1 (shown) and H9 (not shown) following P2 at day 20 (images taken at day 33). Green is MAP2a/b. Red is DARPP32. Blue is Hoechst. Note: DARPP32 expression at this time is easier to visualise in G.
G. hiPSCs are capable of differentiation into DARPP32 (green) expressing neurons (images taken at day 33). Blue is Hoechst.

same density during this passage, yet, post plating, L1 cultures were observed to exhibit additional cell death and thus produced much sparser populations than F1 and H9 cultures passaged at the same time (Figure 4.1F). In spite of this, all three lines produced surviving cells that went on to express MAP2a/b and DARPP32 by day 33 of the differentiation (Figure 4.1F-G) and were able to be maintained as viable cells up to at least day 60 (data not shown). Collectively, this seemed to indicate that line L1 and potentially line M3 were maturing at a faster rate than H9 and F1, although additional study is required to confirm this. In an attempt to improve viability following passage 2, it was moved to 2 days earlier than in Protocol 1 (day 18; Figure 4.1B, P2) which seemed sufficient to improve viability to levels comparable to F1 and H9. This optimisation period resulted in an adapted protocol (Protocol 2, depicted in Figure 4.1B), that produced more consistently viable cells that expressed DARPP32 by day 30 (Figure 4.2A). This protocol was used to differentiate the cells characterised *in vivo* that are described in sections 4.2.2, 4.2.3, and 4.2.4, and also to produce samples for later DNA methylation analysis (Chapter 5).

However, additional difficulties came to light during cell preparation for the transplantation study. Specifically, the MSN progenitors derived from lines L1, L2, L11, M3, and C6 were all found to have a severe reduction in viability following a freeze/thaw cycle prior to transplantation. Due to time constraints following animal lesions, these cells had to be used for the transplantation experiment (with the necessary adjustments made to ensure transplantation of adequate numbers of viable cells). However, it was decided to make one further protocol adjustment to improve the cell viability for subsequent experiments. It seemed likely that the drop in viability post freeze/thaw was related to the presumed accelerated neuronal maturity exhibited by hiPSCs derived from neural tissues, as described above (L1 and to an extent M3; Figure 4.1E). Therefore, additional tests were conducted to determine if conducting the second passage earlier (day 16), with activin treatment continuing the day following a freeze/thaw, would improve viability. All lines tested (L2, L11, C6, F1, and H9) were found to tolerate this final modification and still produced DARPP32 positive neurons (Figures 4.9-11). Therefore, a third protocol (Protocol 3, depicted in Figure 4.1C), was used to differentiate the cells characterised *in vitro* as described in sections 4.2.5, 4.2.6 and 4.2.7.

4.2.2 The generation of MSN progenitors from hPSCs using Protocol 2

Following the development of Protocol 2 (Figure 4.1B), hPSCs were differentiated to build a bank of MSN progenitor cells for future *in vivo* engraftment and *in vitro* characterisation. Specifically, hiPSC lines L1, L2, L11, M3, C6, and F1, and hESC line H9 underwent an MSN differentiation using Protocol 2 (Methods 2.1.7). Each line was differentiated in three separate rounds of differentiation, with each round from each hPSC line separated by 1-2 hPSC passages. This was achieved by splitting pluripotent cells into two populations at each passage during expansion, with the majority going forwards into Protocol 2 for MSN differentiation, and a subset being maintained under pluripotent culture conditions to continue expansion as a pluripotent population for subsequent differentiations. Differentiations were initiated between passages 12-18 for hiPSCs and 18-22 for hESCs. At passage 2 of Protocol 2, hPSCs were frozen as described (Methods 2.1.8), but a small subset was maintained to observe the differentiation up to day 30. These plated cells were then harvested on day 30 for later DNA methylation analysis (described in Chapter 5). Furthermore, to prepare for *in vivo* engraftment, single wells from one passage of lines L1, L2, L11, F1, and H9 were fixed at day 30 for pilot immunocytochemical analysis (Methods 2.2.1). These lines were chosen to undergo this additional step specifically because the original *in vivo* engraftment experiment was intended to examine MSN progenitor grafts derived from two LGE derived hiPSCs, and compare these to MSN progenitor grafts derived from lines F1 and H9. Therefore, this served two purposes, firstly providing evidence that these MSN progenitors were capable of producing MSNs *in vitro*, and secondly to help select which LGE derived hiPSC lines to engraft.

On day 30, MSN progenitors from all examined lines (L1, L2, L11, F1, and H9) exhibited cells immuno-positive for DARPP32 and MAP2a/b (Figure 4.2A). Provisional counts were conducted (biological n=1, cell counts conducted using 5 random fields across 3 coverslips) to help determine which LGE derived hiPSC lines produced the most DARPP32 *in vitro* to guide line selection for *in vivo* engraftment. This indicated that H9 produced the highest percentage of MAP2a/b expressing cells (≈93% of Hoechst positive cells), and L1 produced the least (≈83%), with the other lines producing similar numbers (L2 ≈87%; L11 ≈89%, F1 ≈89%; Figure 4.2B). Examination of the proportion of MAP2a/b expressing cells that co-localised with DARPP32 indicated a trend for LGE derived iPSC lines to express increased levels of DARPP32 compared to F1 (≈16% of MAP2a/b positive cells) and H9 (≈23%), with L11 expressing the highest proportion (≈44%), L2 the next highest (≈38%), and L1 the least (≈28%; Figure 4.2C).

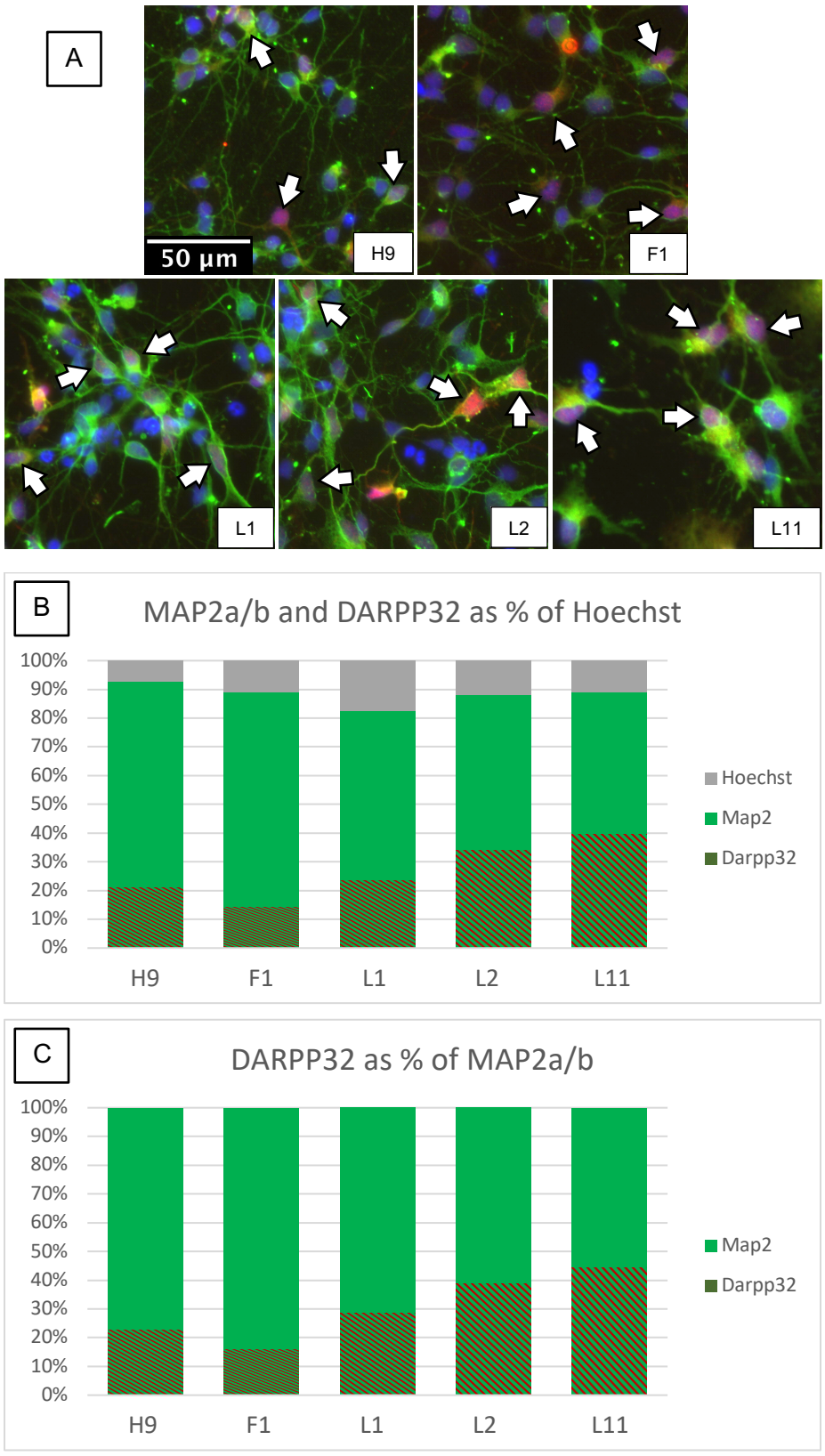


Figure 4.2: hPSCs can differentiate to a DARPP32 expressing neuronal phenotype
A. hPSCs undergoing Protocol 2 in preparation for *in vivo* engraftment. At day 30 all lines are largely positive for the neuronal marker MAP2a/b (green) and some neurons express DARPP32 (red), as indicated by arrows.
B. Bar chart representing the proportion of Hoechst positive cells (grey) that are positive for MAP2a/b (green) and DARPP32 (red stripes).
C. Bar chart representing the proportion of MAP2a/b positive neurons (green) that co-express DARPP32 (red stripes).

4.2.3 *In vivo* examination of grafted hPSC derived MSN progenitors

Following the complete generation of a bank of MSN progenitors from multiple cell lines, and the provisional results described in section 4.2.2, the process of transplanting these cells into a HD rat model began, to systematically study how these cells performed following transplantation into the brain. Originally, it was planned that MSN progenitors from lines L2, L11, F1, and H9 would be transplanted into the QA lesioned rat striatum (L1 was not planned for use as it produced the lowest number of DARPP32 positive cells of the three LGE derived hiPSC lines; Figure 4.2C). As such, 56 rats underwent QA lesion surgery (Methods 2.3.3),and underwent baseline apomorphine rotation analysis (Methods 2.3.5) to assess the lesion success and subsequently group the rats into balanced behavioural groups (n=12 for L2, L11, F1, H9; n=8 lesion only, n=6 control). PSC derived MSN progenitors were defrosted two days prior to planned surgery, with the intention of culturing them for an additional 2 days to allow them to recover from thawing and bring them to the same time point previously used to transplant hESCs cultured under similar techniques (Arber *et al.*, 2015, transplanted on day 20). During the first round of thawing it became immediately apparent that MSN progenitors from lines L2 and L11 were of a severely reduced viability than expected (L2 ≈10%, L11 ≈25% viable cells as per a Trypan Blue count) which was surprising considering no adverse effects had been observed during passage 2 of Protocol 2, or in the culture period of these cells following passage 2 (Section 4.2.2), therefore this seemed to be specifically related to reduced viability following a freeze/thaw process. This was also found to be the case for MSN progenitors derived from lines L1 (≈15%), M3 (<5%), and C6 (≈30%), but not from lines F1 (≈90%), nor H9 (≈90%). Therefore, QA lesioned rats were allocated into new groups to receive grafts as follows: L1 *n*=2; L2 *n*=1; L11 *n*=11; C6 *n*=6, F1 *n*=10, H9 *n*=10 (Methods 2.3.4). These new groups were still balanced between groups for their baseline apomorphine induced rotations. Following 6 months of incubation, animals were perfused (Methods 2.3.6) and the grafts were processed for immunohistochemical and stereological analysis (Methods 2.4). Graft survival at 6 months was confirmed by HuNu positive staining, and survival was low as only 27.5% of grafts survived across this period (Figure 4.3A; graft survival by group: L1 *n*=0/2; L2 *n*=1/1; L11 *n*=0/11; C6 *n*=3/6; F1 *n*=3/10; H9 *n*=4/10).

There were 4 surviving grafts derived from the control hESC line H9, none of which exhibited overgrowth (Figure 4.3A: H9). These grafts had a dense graft core, but some cells at the edge of these grafts migrated a little beyond this boundary into the surrounding tissues (Figure 4.3B: H9). Routinely this was into the neighbouring tail of the ipsilateral external capsule, from where the cells then appeared to migrate into the

corpus callosum (Figure 4.3C: H9). In two of the grafts there was also evidence that a few of the cells had migrated through the external capsule into the neighbouring ipsilateral cortex (Figure 4.3D: H9). There were 0 surviving grafts derived from hiPSC lines L1 or L11, but the 1 L2 derived graft did survive (Figure 4.3A: L2). This graft was confined to the ventral striatum, and did not migrate beyond the graft site (Figure 4.3B: L2), no cells being observed in the corpus callosum or neighbouring ipsilateral cortex (Figure 4.3 C-D: L2). There were 3 C6 derived grafts that survived, all of which exhibited signs of overgrowth to varying degrees (Figure 4.3A: C6). Two of these grafts did not appear to negatively affect the hosts' native anatomy, but one (Table 4.1, animal 487) exhibited localised growth in the cortex that likely occurred through unwanted cell placement in the needle tract during transplantation, and this cortical portion of the graft did distort the native anatomy. The C6 derived grafts were consistently large with a less defined graft core, instead cells seemed to proliferate and migrate to fill the majority of the striatum (Figure 4.3B: C6). Each C6-derived graft contained cells that had migrated beyond the graft into the ipsilateral external capsule, yet, in spite of large cell numbers, very few cells were observed to have migrated into the corpus callosum, instead cells appeared to be more inclined to migrate into the ipsilateral cortex (Figure 4.3C-D: C6). There were 3 surviving grafts derived from hiPSC line F1, and of all hiPSC derived grafts these were most similar to the H9 derived grafts. They exhibited no obvious overgrowth but consistently produced large grafts (Figure 4.3A: F1) with a dense core, but cells migrated beyond this boundary more often than H9 (Figure 4.3B: F1). These cells were consistently found in the external capsule, and migrating through the corpus callosum and far into the contralateral hemisphere (Figure 4.3C: F1). These cells were also found to migrate into the ipsilateral cortex, although to a lesser extent than the C6 derived grafts (Figure 4.3D: F1). All grafts except one C6 derived graft were positive in areas for human DARPP32 (Figure 4.3E), representing expression of at least one MSN marker. Table 4.1 contains key information about all grafts.

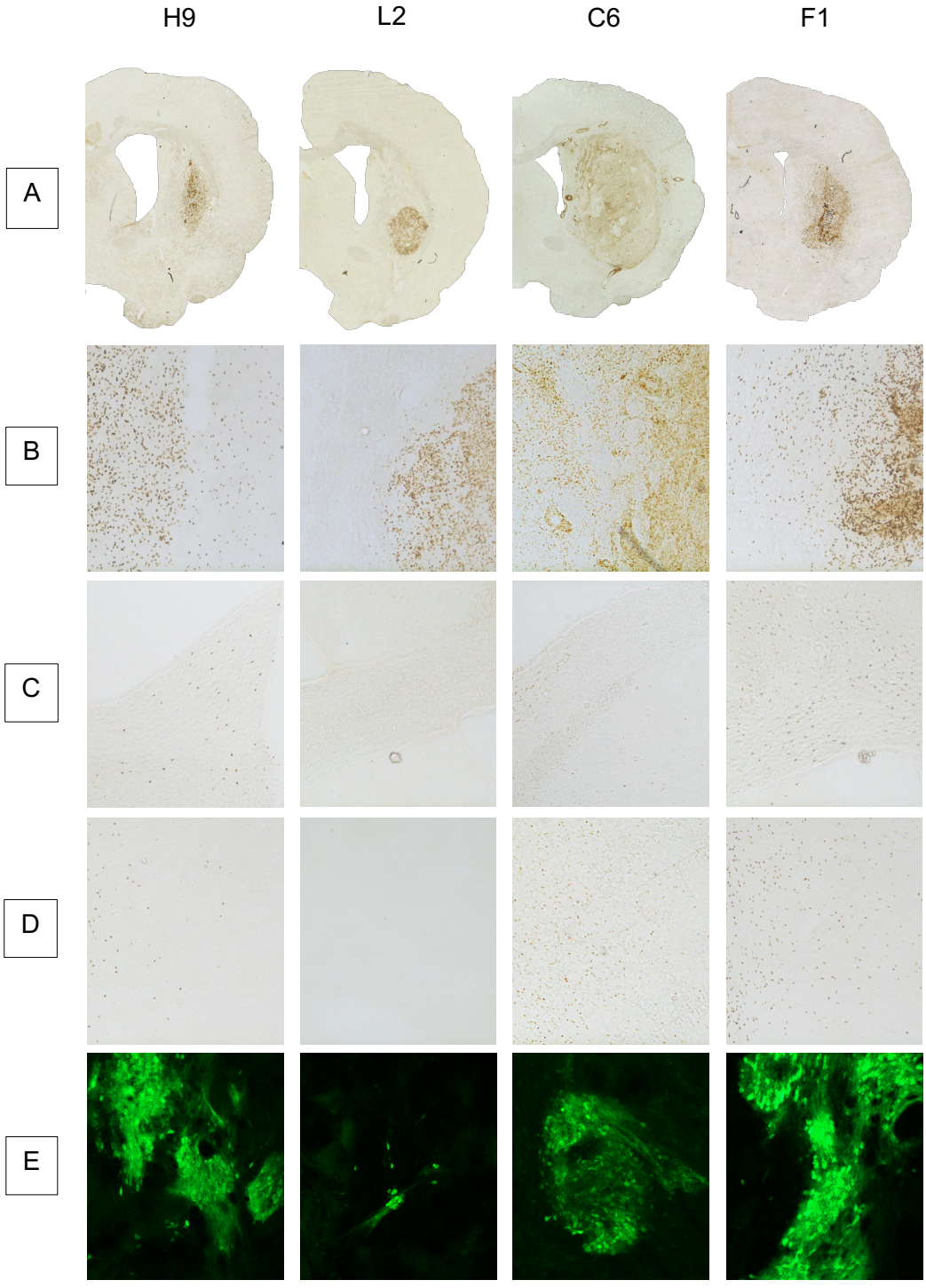


Figure 4.3: Engraftment of hPSC derived MSNs
A-D. Representative HuNu stain by group. Animals shown are 479 (H9), 485 (L2), 476 (C6), and 459 (F1).
B. Higher magnification of graft core edge, and example of migratory cells. Here, H9 exhibits cells migrating from the graft core (left) into the external capsule, and ipsilateral cortex (right). L2 does not exhibit migratory cells beyond the graft core. C6 cells have expanded to fill much of the striatum, but still exhibit signs of further migration (left) beyond the denser graft core (right). F1 exhibit clear migration from the graft core (right) more commonly into the medial striatum (left).
C. Higher magnification of corpus callosum migratory cells (none in L2, few in C6).
D. Higher magnification of cortex migratory cells (few in H9, none in L2).
E. Examples of positive human DARPP32 staining within the grafts (green).

Table 4.1 – Summary of graft features

Animal	Cell Line	Graft core?	Graft periphery?	Migration into external capsule?	Migration into cortex?	Migration into corpus callosum?	Positive DARPP32?
428	H9	Yes	Little	Yes	No	Yes	Yes
444	H9	Yes	Little	Yes	No	Yes	Yes
468	H9	Yes	Little	Yes	Little	Yes	Yes
479	H9	Yes	Little	Yes	Little	No	Yes
485	L2	Yes	No	No	No	No	Yes
476	C6	Yes*	Yes	Yes	Yes	Little	Yes
487	C6	Yes*	Yes	Yes	Yes	Little	Yes
489	C6	Yes*	Yes	No	No	No	No
457	F1	Yes	Yes	Yes	Little	Yes	Yes
459	F1	Yes	Yes	Yes	Yes	Yes	Yes
480	F1	No	Yes	Yes	Little	Yes	Yes

* This was less defined than in other grafts
(note: table is shaded to more easily indicate hPSC graft group)

Stereological analysis was then conducted on the grafts (Methods 2.4.4), the results of which are detailed in Figure 4.4 below. Stereology was conducted initially on the whole graft, including the less dense, migratory part of the graft (whole graft, left), and separately on the dense graft core (graft core, right), to estimate total cell count and graft volume (Figure 4.4A-D). These values were then used to estimate graft density for each region (cells per mm^3; Figure 4.4E-F). Statistical analysis was then conducted on these estimates using one way ANOVAs, L2 was not included as there was only one rat with a surviving graft (n=1). No significant differences were observed between the groups for graft volume, cell number, or graft density in either total graft or core (p>0.05; Figure 4.4E-F).

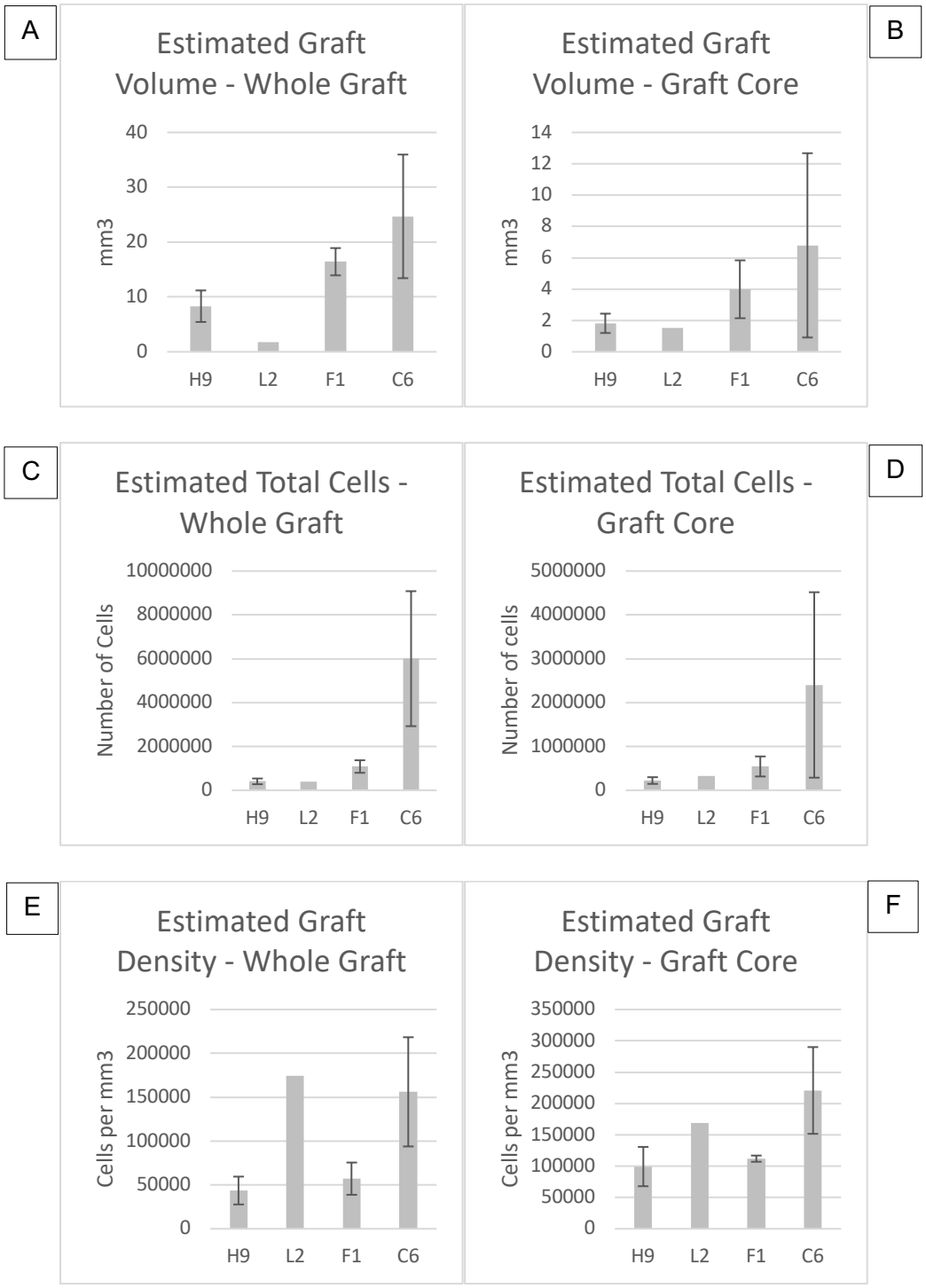


Figure 4.4: Graft Stereology
A-F. Bar chart representing the estimated whole graft volume (**A**); estimated graft core volume (**B**); estimated total number of cells across the whole graft (**C**); estimated total number of cells in the graft core (**D**); estimated whole graft density as cells per mm^3 (**E**); and estimated graft core density as cells per mm^3 (**F**). Error bars are Standard Error from the Mean. There are no statistical differences between the groups (p>0.05), L2 was not included in the statistical tests as there was only one surviving graft (n of 1).

4.2.4 Behavioural analysis of hPSC derived grafts

Grafted animals also underwent behavioural analysis to observe if the grafts had any effect on functional recovery. Apomorphine rotations are one of the most robust measurements of human tissue graft mediated functional recovery in a unilateral QA striatal lesion model (Jerrusi and Glick, 1975; Pundt *et al.*, 1996b; Sanberg *et al.*, 1997; Lelos *et al.*, 2016), and as such this was the behavioural assessment used. Animals were tested before grafting (baseline), and at weeks 12, 16, 20, and 24 following transplantation surgery (Methods 2.3.5). Animals were removed from the final analysis if they were not found to have a surviving graft, so final groups contained only the animals outlined above (Table 4.1). A two way repeated measures ANOVA was conducted on the rotation data with Week and Group as factors, L2 was not included as there was only one rat with a surviving graft (n=1). A significant main effect of Week ($F_{4,72}$=12.53, p<0.001) and Group ($F_{4,18}$=17.27, p<0.001) was found, and post hoc analysis indicated that Control rats rotated significantly less than all other groups (range of mean difference of experimental groups to control rats at all time points: 4.9-6.4 rotations per minute, p=0.001), no other group differences were observed. Additionally, rats rotated fewer times at baseline than at later time points which is in keeping with the expected behavioural trajectory for this surgical procedure (p=0.005). Collectively this indicated that no graft group exhibited functional recovery across the experiment (Figure 4.5).

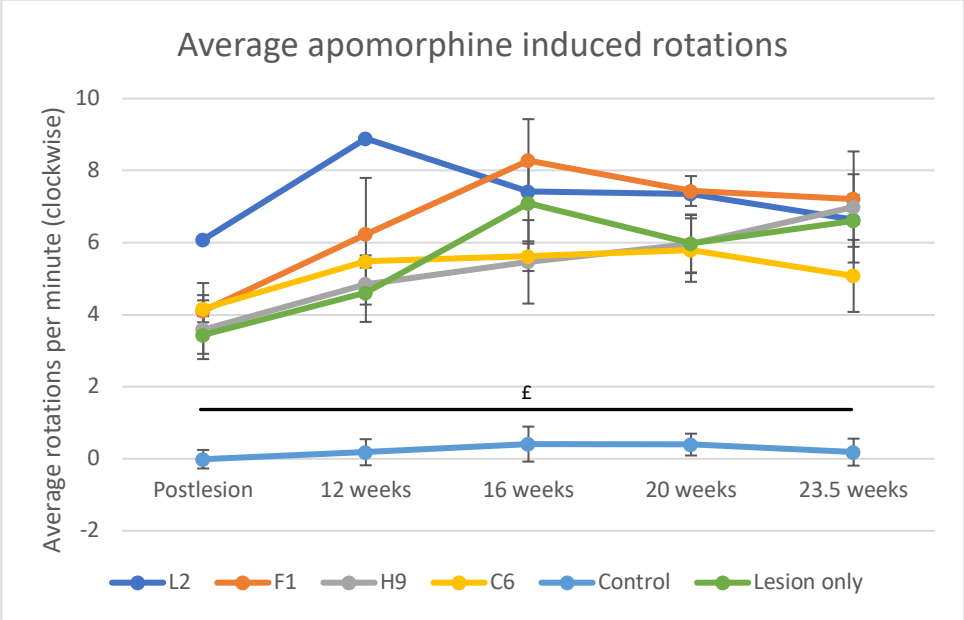


Figure 4.5: Behavioural Analysis
Average number of apomorphine induced rotations per minute by weeks post-transplant, error bars are SEM. Only animals with surviving grafts are included in the grafted groups (L2 n=1; F1 n=4; H9 n=4; C6 n=3; Control n=6; Lesion only n= 8). No graft groups exhibited a significant change in number of rotations (p>0.05), and consistently rotated significantly more than the control rats (p=0.001). £ indicates main effect of group, control rats consistently rotated less than all other experimental groups.

4.2.5 qPCR analysis of hPSC derived MSN populations across *in vitro* maturation

Following the discovery that neural derived hiPSCs did not survive a freeze/thaw cycle after a passage conducted at day 18, the MSN differentiation protocol was revised to Protocol 3 (Fig 4.1C), as described above. This protocol was applied to hiPSC lines L2, L11, C6, F1 and hESC line H9 in a similar manner as described in section 4.2.2, resulting in 3 separate MSN progenitor differentiations from each line (Methods 2.1.7). These cells were then thawed on the same day and plated for *in vitro* maturation, so as to minimise variation caused by the maturation environment and experimenter induced variability. Samples of these cultures were fixed for immunocytochemical analysis or underwent RNA extraction at day 20, 30, and 40. These three time points were chosen for the following reasons: day 20 is when PSC cultures should have completed localised patterning for an LGE progenitor phenotype and this is equivalent to the day similar cells were transplanted in the previous described *in vivo* experiment (4.2.3); day 30 is when the key MSN marker DARPP32 is first routinely expressed by hESC lines undergoing a similar differentiation protocol (Protocol 1; Figure 4.1A); and day 40 is when DARPP32 expression in hESC lines following Protocol 1 plateaus and total MSN yield can be quantified.

qPCR was used to examine expression of critical subpallium, LGE, and MSN markers across these time points (Methods 2.2.2.3). All qPCR outputs were calculated as relative fold change to the expression of the same gene from H9-MSNs at day 30. All outputs were analysed for each gene independently, using a two way repeated measures ANOVA between cell line (Line) and across each time point (Day). Where appropriate, Sidak's t test was used for post hoc analysis. Fold changes are described with ± SEM, except when indicated.

Early fate specification was examined using DLX2 for pan-WGE fate specification and response to Activin (Feijen, Goumans and van den Eijnden-van Raaij, 1994; Maira *et al.*, 2010), and NKX2.1 for ventral/MGE fate specification (Noakes *et al.*, 2019). There was a significant interaction between the effects of Day and Line on DLX2 expression ($F_{8,20}$=3.31, p=0.014; Figure 4.6A). As expected, DLX2 expression was significantly higher at Day 20 than at Day 30 and Day 40 in H9 cultures (H9 Day 20 to Day 30: 2.5 ± 0.42 fold difference, p=0.001). This was also observed in L2 (1.9 ± 0.42 fold difference, p=0.003), L11 (1.3 ± 0.42 fold difference, p=0.041), and F1 (2.1 ± 0.42 fold difference, p=0.002) cultures, but not in C6 cultures which expressed statistically similar levels of DLX2 across all timepoints, and significantly lower levels of DLX2 than

H9 cultures at Day 20 (mean=1.7 ± SEM=0.4 fold difference between H9 and C6 at day 20, p=0.016), but not compared to the other hiPSC lines. Furthermore, there was a significant main effect of Day ($F_{2,20}$=9.2, p=0.001) and of Line ($F_{4,10}$=4.12, p=0.032) on NKX2.1 expression, but no interaction between the two (Figure 4.6B). Post hoc analysis revealed that the significant main effect of Day was driven by an overall drop in NKX2.1 expression between Day 20 and the later time points (8.1 ± 2.8 fold reduction by Day 30, p=0.049; and 8.9 ± 2.7 fold reduction by Day 40, p=0.025). Post hoc analysis of the main effect of Line was not significant (p>0.05). However, there was a clear trend for C6 cultures to express more NKX2.1 than H9 (30.8 ± 9.1 fold difference to C6, p=0.066), F1 (30.0 ± 9.1 fold difference, p=0.075) and L11 (28.9 ± 9.1 fold difference, p=0.093) cultures. Expression of NKX2.1 varied greatly between the three C6 derived cultures, but was somewhat consistent within each independent differentiation (mean fold difference compared to average expression in H9 at day 30: C6 MSN differentiation 1, total mean across day 20, day 30 and day 40 = 55.11, st.dev = 24.99; C6 MSN differentiation 2, total mean: 35.32 st.dev: 5.52; C6 MSN differentiation 3, mean: 6.87 st.dev: 1.47).

Early LGE specific patterning was examined using the progenitor LGE marker NOLZ1, which exhibits reduced expression in terminally mature MSNs (Urbán *et al.*, 2010). Additionally, maturing MSN fate was examined using CTIP2 (Arlotta *et al.*, 2008), FOXP1 (Precious *et al.*, 2016), and FOXP2 (Devanna, Middelbeek and Vernes, 2014) which are all highly expressed in developing striatal MSN populations. There was a significant main effect of Line on NOLZ1 expression ($F_{4,10}$=3.54, p=0.048; Figure 4.6C), and post hoc analysis revealed that F1 expressed significantly more NOLZ1 than L11 (1.1 ± 0.3 fold difference, p=0.044). No other differences were observed including across day (p>0.05), indicating the MSNs here were not terminally mature. Similarly, there was a significant main effect of Line on CTIP2 expression ($F_{4,10}$=12.15, p=0.001; Figure 4.6D). Post hoc analysis revealed that F1 expressed significantly more CTIP2 than L2 (1 ± 0.17 fold difference, p=0.002), L11 (1.1 ± 0.17 fold difference, p=0.001), and C6 (0.8 ± 0.17 fold difference, p=0.008). There was a significant interaction between the effects of Day and Line on FOXP1 expression ($F_{8,20}$=5.82, p=0.001; Figure 4.6E). Post hoc analysis revealed that FOXP1 expression increased over time in all lines (all p<0.05). At day 20, H9 expressed significantly more FOXP1 than L11 (0.28 ± 0.07 fold difference, p=0.023), but not at later time points. Whereas at day 40, F1 expressed significantly more FOXP1 than all other lines (H9: 2.2 ± 0.42 fold difference to F1, p=0.004; L2: 2.3 ± 0.42 fold difference, p=0.003; L11: 2.3 ± 0.42 fold difference, p=0.003; C6: 2.4 ± 0.42 fold difference, p=0.002) but only at this time point. There was a significant interaction between the effects of Day and Line on FOXP2 expression ($F_{8,20}$=6.85, p<0.001; Figure 4.6F). Post hoc analysis revealed only lines H9 and L11

exhibited an increase in FOXP2 expression over time, and this increase was only significant for these lines between Day 20 and Day 30/Day 40 (H9 Day 20 to Day 40: 1.1 ± 0.12 fold difference, $p<0.001$; L11 Day 20 to Day 40: 0.6 ± 0.12 fold difference, $p=0.001$). Furthermore, at Day 20, H9 expressed significantly more FOXP2 than C6 (0.2 ± 0.05 fold difference, $p=0.019$). By Day 30 H9 expressed significantly more FOXP2 than all lines (L2: 0.8 ± 0.13 fold difference to H9 at day 30, $p=0.001$; L11: 0.5 ± 0.13 fold difference, $p=0.016$; C6: 0.7 ± 0.13 fold difference, $p=0.002$; F1: 0.6 ± 0.13 fold difference, $p=0.005$), and this remained unchanged by day 40 (L2: 1.1 ± 0.17 fold difference to H9 at day 40, $p=0.001$; L11: 0.6 ± 0.17 fold difference, $p=0.047$; C6: 1.1 ± 0.17 fold difference, $p=0.001$; F1: 1 ± 0.17 fold difference, $p=0.002$).

Additionally, mature striatal MSN phenotypes were examined using classical MSN markers DARPP32, CALB1, DRD1, and DRD2 (Gerfen, 1985; Anderson and Reiner, 1991; Ferré *et al.*, 1997; Ouimet, Langley-Gullion and Greengard, 1998; Quiroz *et al.*, 2009; Straccia *et al.*, 2015). There was a significant interaction between the effects of Day and Line on DARPP32 expression ($F_{8,20}=11.48$, $p<0.001$; Figure 4.6G). Post hoc analysis revealed that all lines exhibited a significant increase in expression between Day 20 and Day 30 ($p<0.01$, not depicted on Figure 4.6G), which was expected as DARPP32 was not expressed in cultures until Day 30. Only H9 and F1 exhibited a further significant increase between Day 30 and 40 (H9: 1.8 ± 0.39 fold increase by day 40, $p=0.003$; F1: 2.9 ± 0.39 fold increase, $p<0.001$). At day 30 there were no differences between lines, however by day 40, H9 expressed significantly more DARPP32 than L2 (1.8 ± 0.49 fold difference to H9, $p<0.044$), and F1 expressed significantly more DARPP32 than all other lines (H9: 1.9 ± 0.49 fold difference to F1, $p=0.026$; L2: 3.7 ± 0.49 fold difference, $p<0.001$; L11: 3.4 ± 0.49 fold difference, $p<0.001$; C6: 2.7 ± 0.49 fold difference, $p=0.003$). There was a significant interaction between the effects of Day and Line on CALB1 expression ($F_{8,20}=7.40$, $p<0.001$; Figure 4.6H). Post hoc analysis revealed all lines increased expression of CALB1 over time, although when this occurred varied by line: H9 and F1 expressed significantly more CALB1 by Day 30 than they had at Day 20, and exhibited no further changes by Day 40, whereas L2, L11, and C6 expressed significantly more CALB1 at day 40 than they had at day 20, but not before then. Due to this, F1 expressed significantly more CALB1 at day 30 than both L2 and L11 (L2: 0.9 ± 0.24 fold difference, $p<0.037$; L11: 1 ± 0.24 fold difference, $p<0.024$). However, this was no longer the case by day 40, and instead L2 expressed significantly more CALB1 than L11 (1.3 ± 0.29 fold difference, $p<0.013$). There was a significant interaction between the effects of Day and Line on DRD1 expression ($F_{8,20}=2.97$, $p=0.023$; Figure 4.6I). Post hoc analysis revealed that H9 expressed significantly more DRD1 at Day 40 than at Day 30 (0.7 ± 0.2 fold difference, $p<0.013$), and by Day 40, H9 expressed significantly more DRD1 than L11, but not any

other line (1.2 ± 0.32 fold difference, $p<0.029$). There was a significant interaction between the effects of Day and Line on DRD2 expression ($F_{8,20}=7.17$, $p<0.001$; Figure 4.6J). Post hoc analysis revealed that L11 expressed significantly more DRD2 by day 40 compared to Day 20 (0.7 ± 0.22 fold difference, $p<0.024$), and both H9 and F1 expressed significantly more DRD2 by Day 30 than they did at Day 20 (H9: 1 ± 0.33 fold difference, $p=0.033$; F1: 2.9 ± 0.33 fold difference, $p<0.001$), although no further increase occurred after this point. This increase meant that H9 expressed more DRD2 than C6 by day 40 (1.1 ± 0.29 fold difference, $p=0.033$). Furthermore, F1 expressed significantly more DRD2 than all other lines, both at Day 30 (H9: 1.9 ± 0.49 fold difference, $p=0.025$; L2: 2.6 ± 0.49 fold difference, $p=0.003$; L11: 2.4 ± 0.49 fold difference, $p=0.006$; C6: 2.8 ± 0.49 fold difference, $p=0.002$) and Day 40 (H9: 1.2 ± 0.29 fold difference, $p=0.024$; L2: 1.8 ± 0.29 fold difference, $p=0.001$; L11: 1.7 ± 0.29 fold difference, $p=0.002$; C6: 2.3 ± 0.29 fold difference, $p<0.001$).

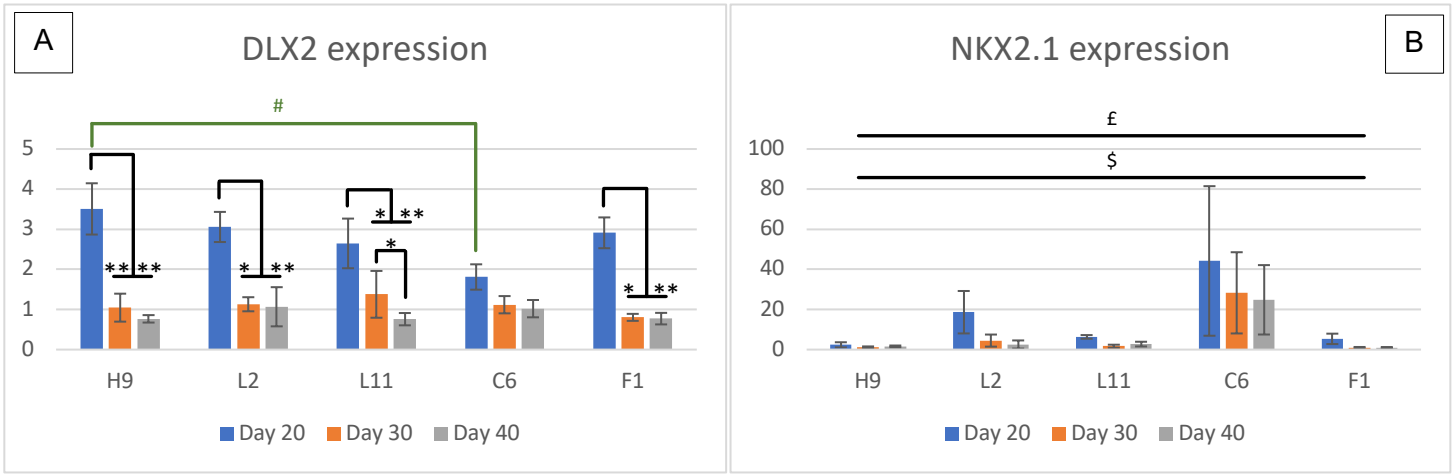
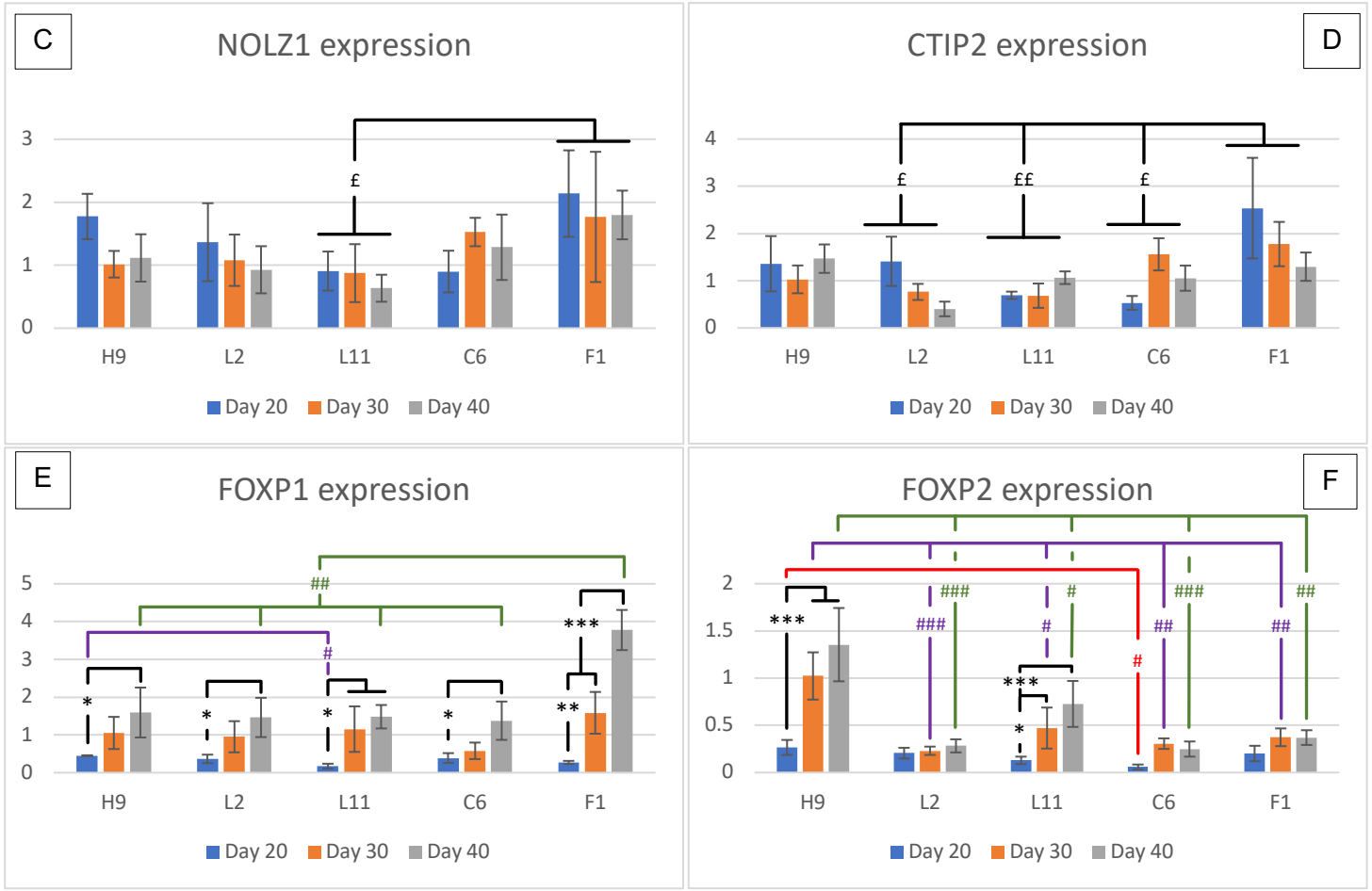
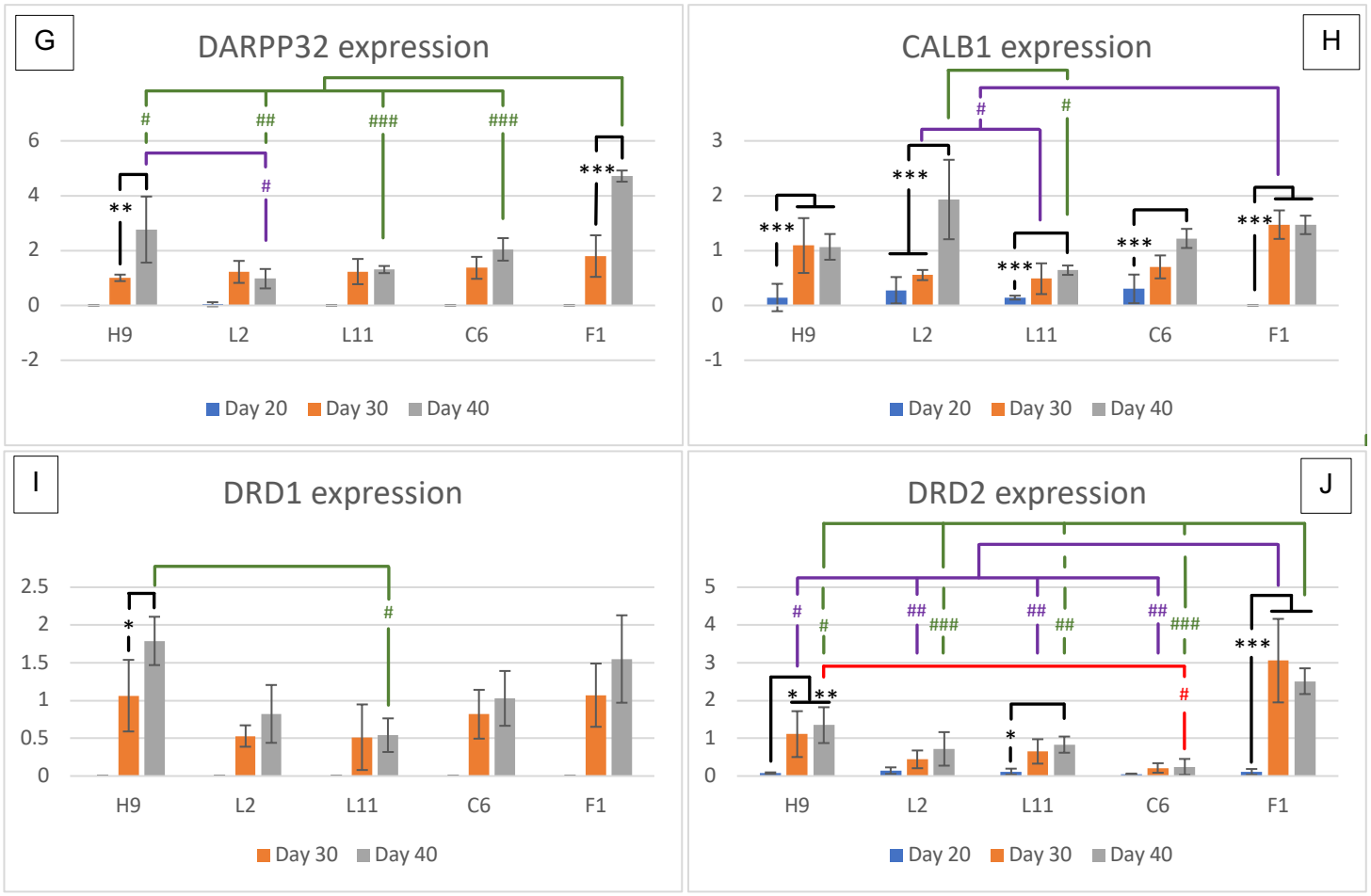


Figure 4.6 (continued overleaf): qPCR analysis of key striatal genes
A-J. Bar chart representations of qPCR gene expression data collected from hPSC-MSNs at day 20, 30 and 40 of 3 independent MSN differentiations. All data is normalised to H9 cultures at day 30. A two way repeated measures ANOVA was conducted on all genes examined with Line and Day as IVs. Significant results are depicted on bar charts where £ symbolises main effects of line, $ symbolises main effects of Day, # symbolises interactions of Line/Day and * symbolises interactions of Day/Line. The number of symbols indicates *p* value, where 1 symbol = $p<0.05$, 2 = $p<0.01$, and 3 = $p<0.001$. **A** = DLX2; **B** = NKX2.1; **C** = NOLZ1; **D** = CTIP2; **E** = FOXP1; **F** = FOXP2. **G** = DARPP32; **H** = CALB1; **I** = DRD1; **J** = DRD2.


C
NOLZ1 expression
3
2
1
0
£
H9
L2
L11
C6
F1
Day 20
Day 30
Day 40
D
CTIP2 expression
4
3
2
1
0
£
£
£
H9
L2
L11
C6
F1
Day 20
Day 30
Day 40
E
FOXP1 expression
5
4
3
2
1
0
##
#
*
*
*
*
***
**
H9
L2
L11
C6
F1
Day 20
Day 30
Day 40
F
FOXP2 expression
2
1.5
1
0.5
0
***
###
###
#
#
#
***
*
##
##
##
H9
L2
L11
C6
F1
Day 20
Day 30
Day 40


G
DARPP32 expression
6
4
2
0
-2
H9
L2
L11
C6
F1
Day 20
Day 30
Day 40
H
CALB1 expression
3
2
1
0
-1
I
DRD1 expression
2.5
2
1.5
1
0.5
0
J
DRD2 expression
5
4
3
2
1
0

4.2.6 Immunocytochemical analysis of hPSC derived MSN populations during *in vitro* maturation

Next, immunocytochemical analysis (Methods 2.2.1) was conducted on the previously described hPSC MSN populations (4.2.5). Samples of hPSC derived MSNs were fixed at Day 30 and Day 40, then stained with either βIII-tubulin and CTIP2, to determine total neuronal specification and the proportion of CTIP2 positive neurons, or CTIP2 and DARPP32 to determine MSN fate commitment. Immuno-positive cells were counted and calculated as a percentage of total Hoechst positive cells. Cell counts were then analysed for each marker independently, using a two way repeated measures ANOVA between cell line (Line) and across both time points (Day). Where appropriate, Sidak's t test was used for post hoc analysis.

As expected from earlier cultures using Protocol 1 and 2, hPSC-MSNs derived using Protocol 3 produced a highly neuronal population. βIII-tubulin was widely expressed by the majority of cells across all lines at both time points examined, and the hPSC derived neurons consistently exhibited widespread axons indicative of maturing cultures (Figure 4.7, data shown for Day 30 only). However, the lines varied by the proportion of neurons they produced, and statistical analysis revealed there was a significant main effect of Line on proportion of βIII-tubulin expressing cells ($F_{1,10}$=7.16, p=0.005; Figure 4.8A). Post hoc analysis revealed that hiPSC line C6 had a significantly higher neuronal population than L11 (C6: mean Hoechst cells co-labelled for βIII-tubulin = 96.5% ± 2.4; L11: 78.9% ± 2.4; mean difference of 17.6% ± 3.4, p=0.004). No other statistical differences were observed.

Next, the progenitor MSN marker CTIP2 was examined, and was identifiable in all cultures at both time points (Figures 4.7, 4.9, 4.10). However, statistical analysis revealed there was a significant interaction between the effects of Day and Line on the number of CTIP2 expressing cells ($F_{4,10}$=3.50, p=0.049; Figure 4.8B). Post hoc analysis revealed that at Day 30, lines L2, C6, and F1 had a significantly higher proportion of CTIP2 positive cells than lines H9 and L11 (H9: mean Hoechst cells co-labelled for CTIP2 at Day 30 = 43.2% ± 3; L2: 58.2% ± 3; L11: 42% ± 3; C6: 64.7% ± 3; F1: 62.7% ± 3; range of significant mean differences: 15% - 22.6% ± 4.2; L2 p<0.05; C6 p<0.01; F1 p<0.01). However, between Day 30 and Day 40 lines H9, L11, and C6 exhibited a significant increase in the number of CTIP2 expressing cells (H9: mean increase of 15.7% ± 3.5, p=0.001; L11: 19.6% ± 3.5, p<0.001; C6: 10% ± 3.5, p=0.017). This resulted in no final differences in CTIP2 populations by day 40.

The final marker examined by immunocytochemistry was the mature MSN marker DARPP32. DARPP32 was identifiable in all lines by Day 30 (Figures 4.9).

Additionally, each line yielded a significant increase in the proportion of positive cells by Day 40 (Figure 4.10) as demonstrated by a significant main effect of Day ($F_{1,10}$=179.5, p<0.001; Figure 4.8C). Post hoc analysis revealed that this increase was lowest in L2 (mean increase of 23.8% ± 4.7) and highest in F1 (mean increase of 33.9% ± 4.7) but all were highly significant (all p<0.001). Additionally, there was a significant main effect of Line ($F_{4,10}$=5.8, p=0.011; Figure 4.8C), and post hoc analysis indicated that F1 cultures had a higher proportion of DARPP32 positive expressing cells than both L2 (mean difference of 11% ± 2.8, p=0.031) and L11 (mean difference of 12.4% ± 2.8, p=0.014).

Of note, at Day 30 the majority of DARPP32 positive cells observed in culture were only expressing the marker weakly, however cells were identifiable in L2 and L11 derived cultures that were expressing it to a higher degree than in other lines (Figure 4.9). By day 40 all lines contained cells expressing high levels of DARPP32 (Figure 4.10). As with all other stains, counts were conducted using all positive DARPP32 staining compared to a control well that was not exposed to the primary antibody.

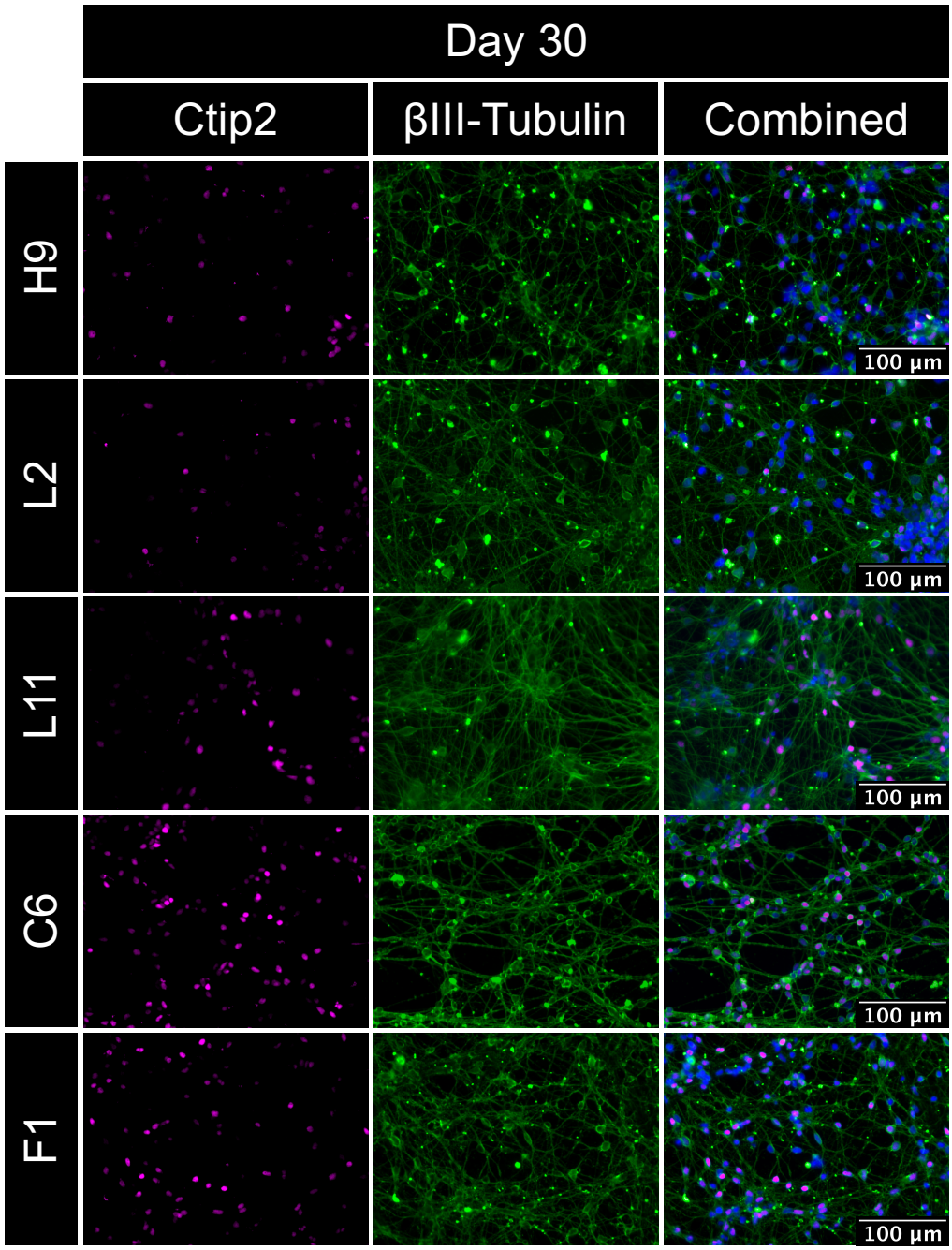


Figure 4.7: Immunocytochemical stains for neuronal marker βIII-tubulin and CTIP2 at day 30 of Protocol 3
Immunocytochemical stains of hPSC-MSN differentiations generated using Protocol 3. CTIP2 is in pink, βIII-tubulin is in green, merged with Hoechst stain in blue. All cell lines are capable of differentiation towards a neuronal phenotype as indicated by high βIII-tubulin expression across cultures.

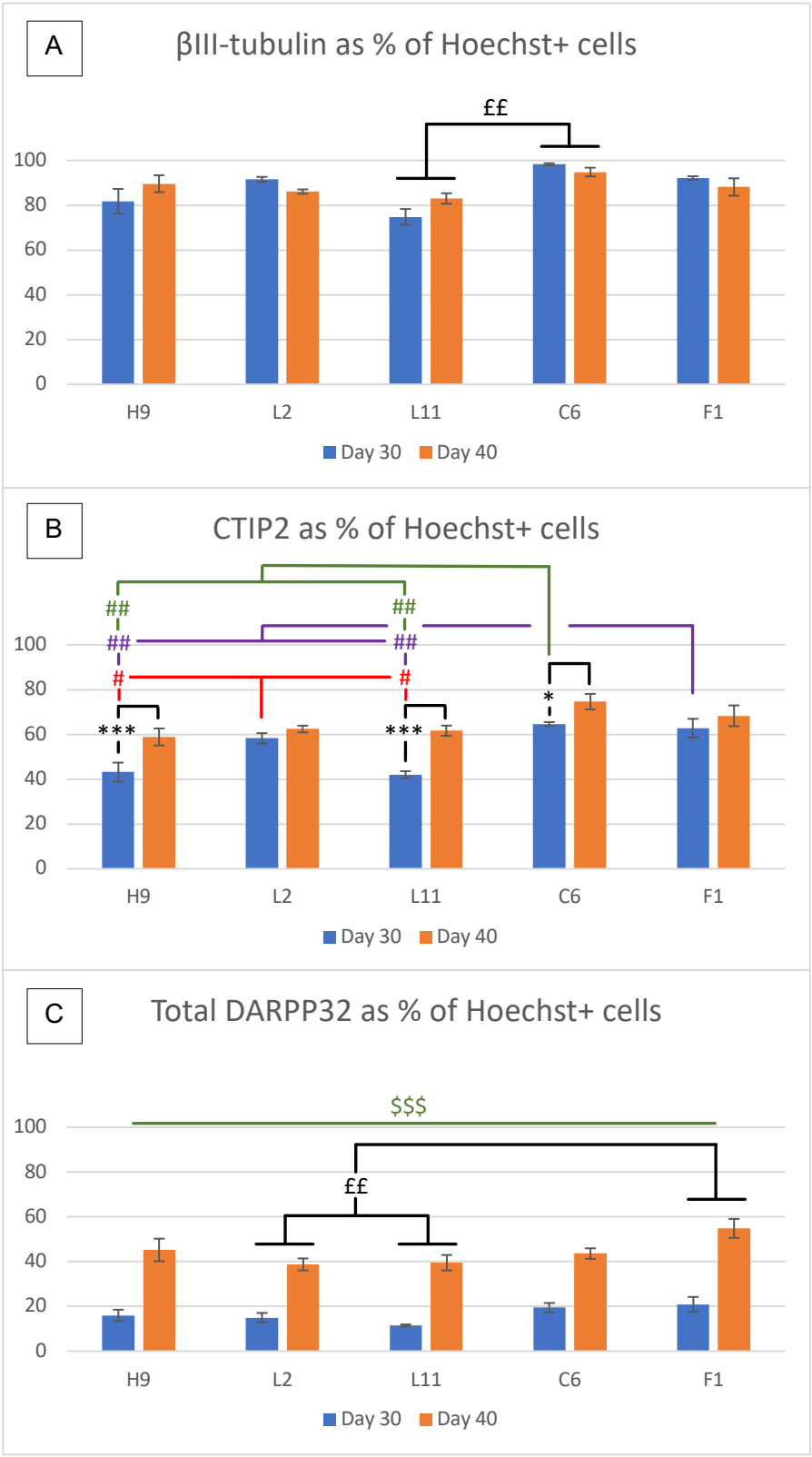


Figure 4.8: Cell counts following immunocytochemistry
A-C. Bar chart representations of percentage of Hoechst positive cells positive for **A.** βIII-tubulin, **B.** CTIP2, and **C.** DARPP32. Data collected from PSC-MSNs at day 30 and 40 of 3 independent MSN differentiations. A two way repeated measures ANOVA was conducted on all genes examined with Line and Day as factors. Significant results are depicted on bar charts where £ symbolises main effects of line, $ symbolises main effects of Day, # symbolises interactions of Line/Day and * symbolises interactions of Day/Line. The number of symbols indicates *p* value, where 1 symbol = $p<0.05$, 2 = $p<0.01$, and 3 = $p<0.001$.

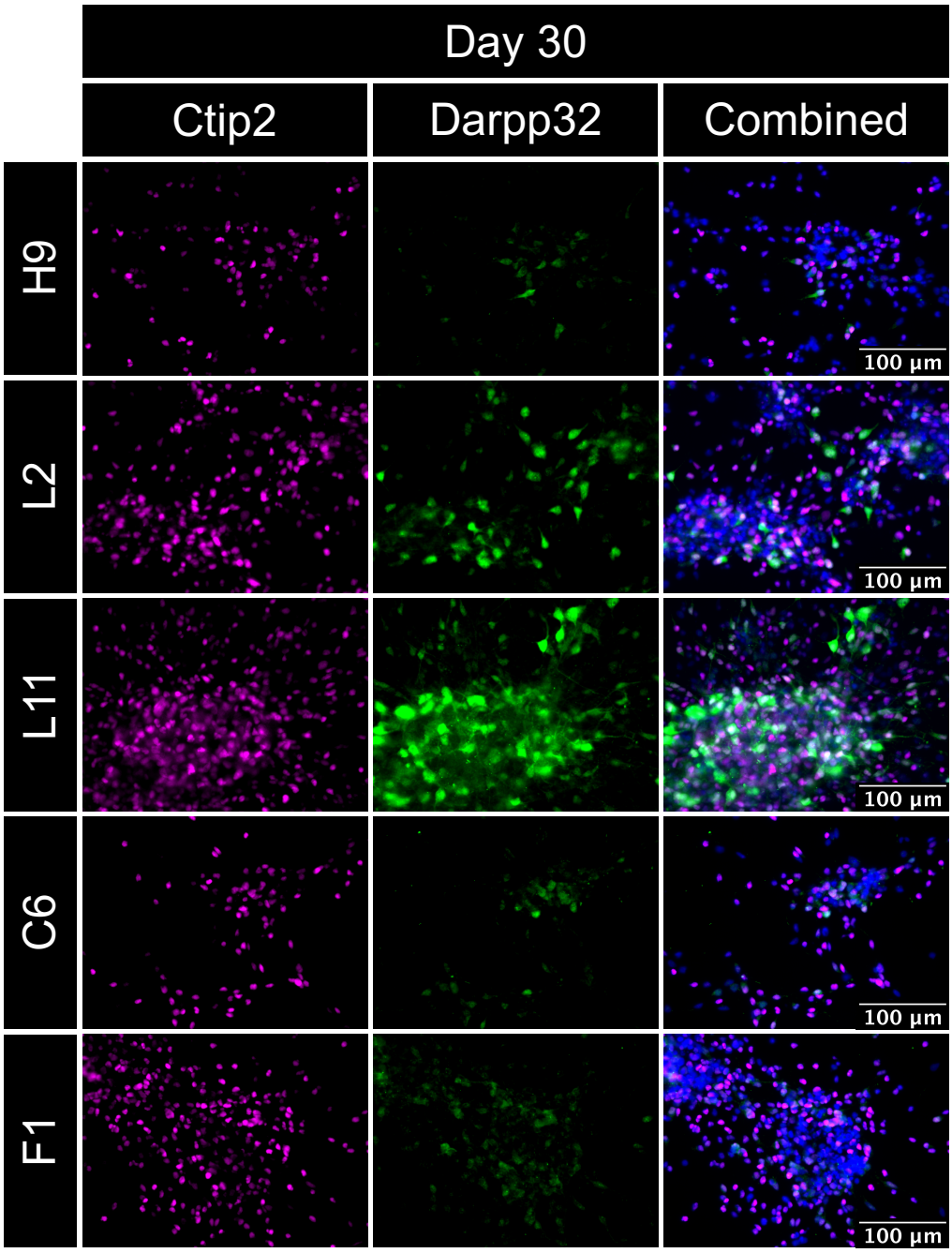


Figure 4.9: Immunocytochemical stains for MSN markers CTIP2 and DARPP32 at day 30 of Protocol 3
Immunocytochemical stains of hPSC-MSN differentiations generated using Protocol 3. CTIP2 is in pink, DARPP32 is in green, merged with Hoechst stain in blue. Whilst DARPP32 is expressed in all cultures, L2 and L11 both exhibited cells strongly immuno-positive for DARPP32. These were not observed in H9, C6, or F1 cultures.

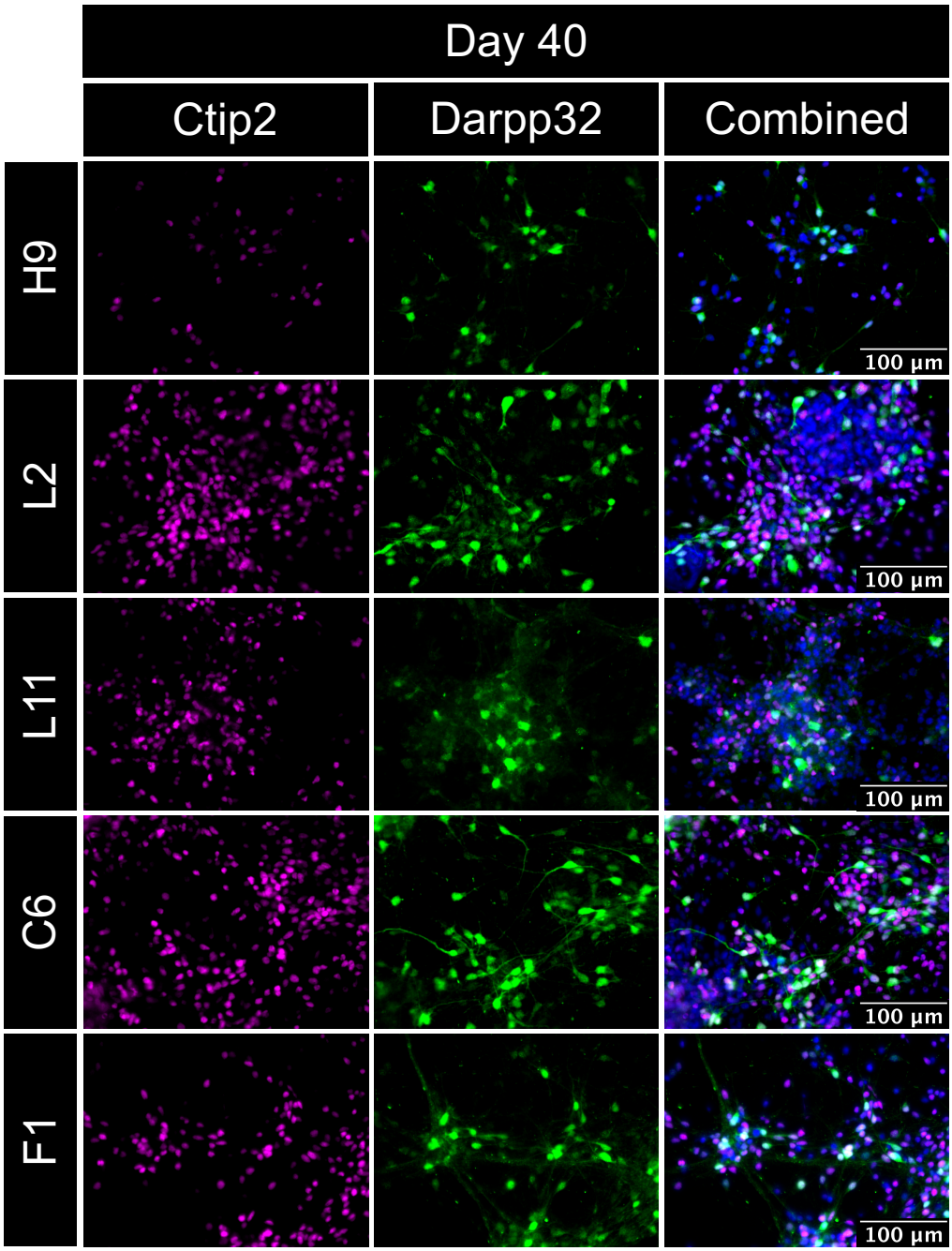


Figure 4.10: Immunocytochemical stains for MSN markers CTIP2 and DARPP32 at day 40 of Protocol 3
Immunocytochemical stains of hPSC-MSN differentiations generated using Protocol 3. CTIP2 is in pink, DARPP32 is in green, merged with Hoechst stain in blue. By this time point, all lines contain cells strongly immuno-positive for DARPP32.

4.2.7 Multi-electrode analysis of spontaneous electrical activity of hPSC derived MSN populations

The spontaneous electrical activity of the previously described hPSC MSN populations (4.2.5) was examined using multi-electrode array (MEA) analysis (Methods 2.2.3). MEA analysis allows the examination of the electrical activity of a population of cells simultaneously. Data presented was collected across a 2 minute window at day 40, which was during the peak of electrical activity observed for all lines.

Spontaneous electrical activity was observed across all lines at this time point, however this activity was generally infrequent, and the spikes were of low amplitude (Figure 4.11; 4.12 activity is indicated by red marks along electrode read out). One well of L11 derived MSNs exhibited frequent activity at this time point (Figure 4.12E), which was observed in the same well at other time points. However the other 5 wells containing L11 derived cells did not exhibit similar levels of activity (Figure 4.12F). Examination of the peak to peak amplitude of single electrical bursts by line, revealed that lines L11 and C6 exhibited the largest peak to peak amplitudes (Figure 4.11; 4.12 indicated by purple boxes/arrow). Additionally, L2 exhibited rapid burst activity at one electrode which may be indicative of neuronal excitability (Figure 4.14A indicated by green boxes) however, this was not observed in other L2 derived wells. There was too little consistency observed within the lines to conduct further meaningful analysis.

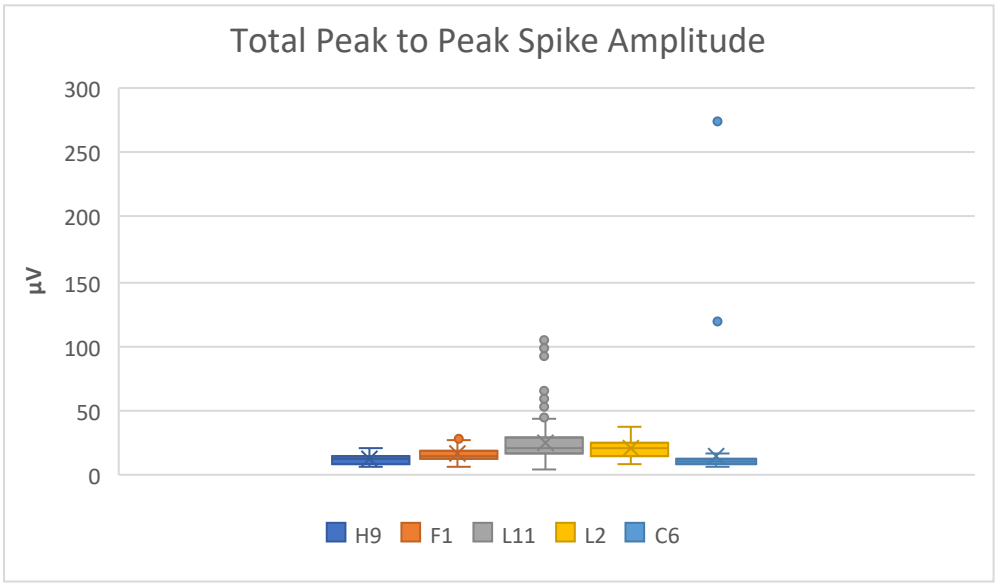
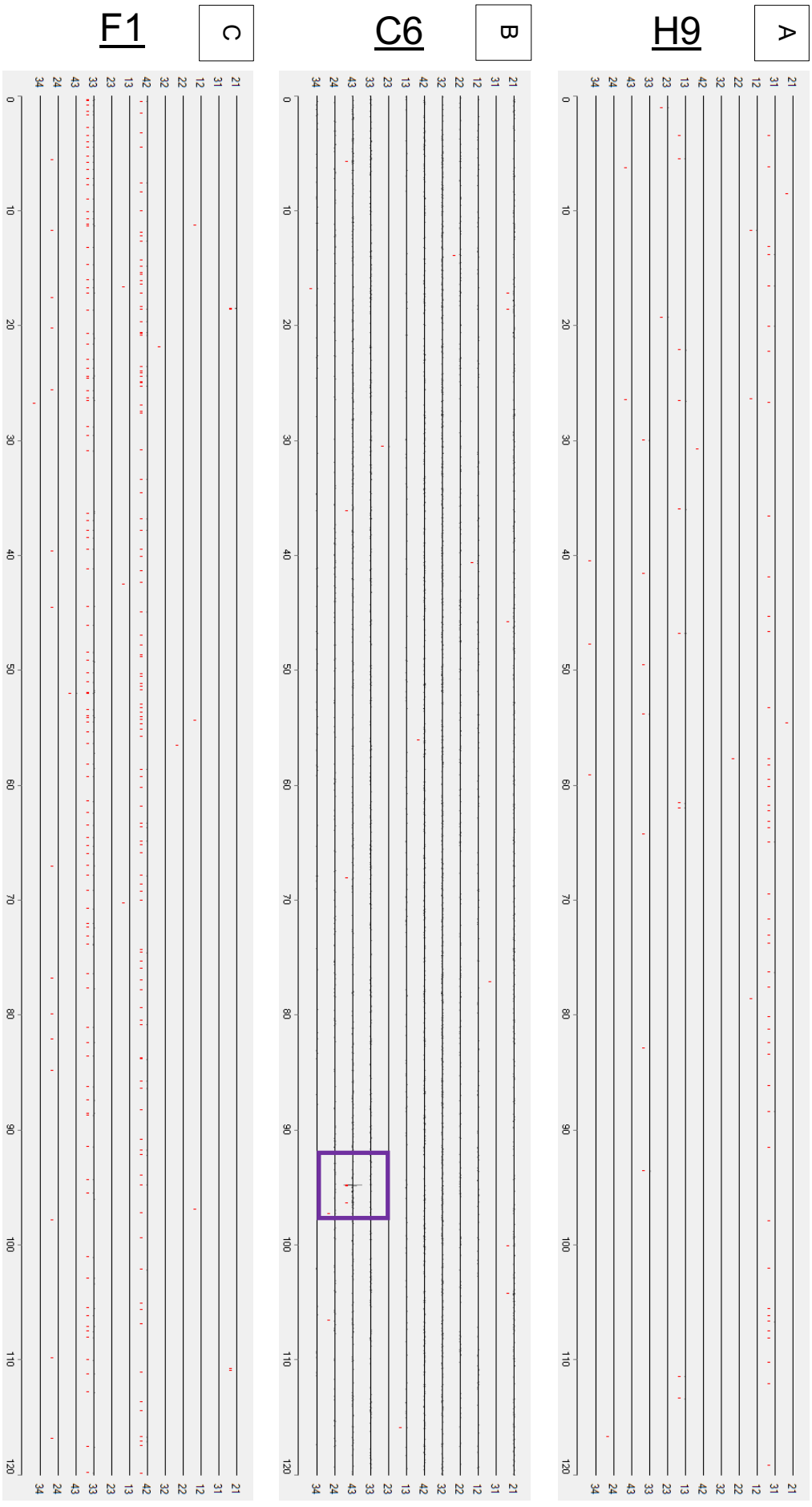
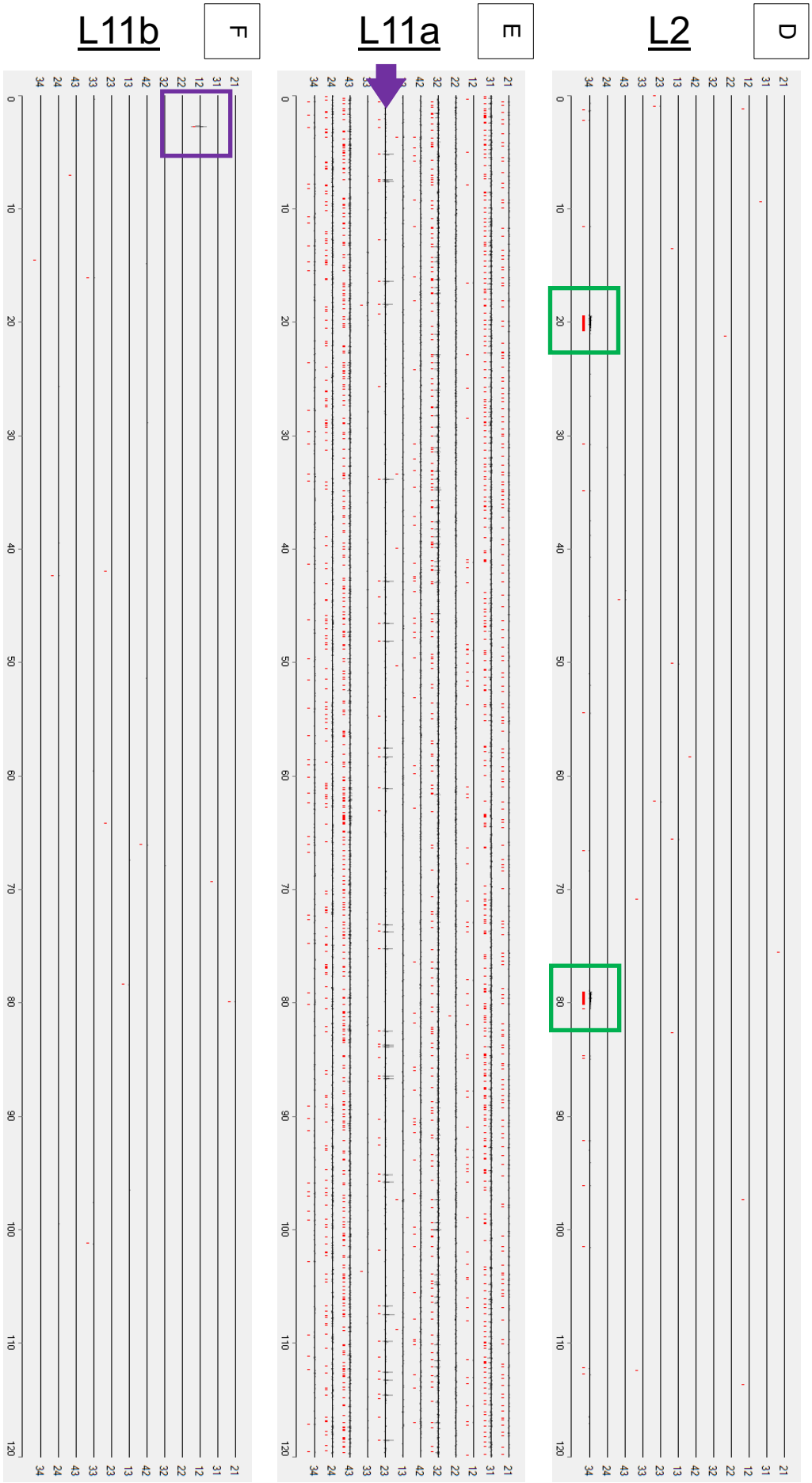


Figure 4.11: Peak to peak spike amplitudes measured at day 40
Boxplot of all peak to peak spike amplitudes measured in hPSC-MSNs at day 40. L11 and C6 exhibit occasional amplitudes far beyond the ranges observed in other lines.

Figure 4.12 (overleaf): Overview of all electrical activity measured at day 40
MEA output graph of electrical activity detected at each well electrode (numbered on Y axis) over time (X axis is seconds). Red marks indicate when activity is measured above the background level, deemed spontaneous electrical activity/spikes. The purple boxes/arrow indicate spikes with a large amplitude. The green boxes indicate rapid burst activity. **A** = H9; **B** = C6; **C** = F1; **D** = L2; **E** = L11 (high activity); **F** = L11 (typical).


A
H9
B
C6
C
F1


L2
D
L11a
E
L11b
F

4.3 Discussion

In this chapter I have characterised the differentiation potential of the isogenic hiPSC lines generated in Chapter 3, and found that each are able to follow a standardised MSN differentiation protocol to produce cells expressing critical MSN markers frequently used to establish MSN fate commitment in hPSC cultures. However, differences have emerged between these lines which could be indicative of their epigenetic memory of tissue of origin.

4.3.1 Fibroblast derived hiPSCs were not hindered by their tissue of origin in their capacity to undergo an MSN differentiation

The fibroblast derived hiPSC line F1 was generated to act as a control for hiPSC tissue of origin, under the assumption that if these cells retained an epigenetic memory of their tissue of origin, it would not be beneficial for the generation of MSNs and could potentially hinder MSN differentiation in this line (Chapter 3, Introduction). Yet, across the various comparisons made in experiments across this chapter, the F1 line typically performed as well as, or arguably better than, the hESC control line, in terms of the measures used here. This was initially apparent in the standardisation of the differentiation protocol (4.2.1), as no difficulties were encountered with these cells undergoing Protocol 1, which was also the case for H9. Initial characterisation via immunocytochemistry seemed to indicate a slightly reduced efficiency of neural induction and reduced DARPP32 expression (4.2.2). However, upon transplantation the F1 derived cells developed into grafts similar to H9 grafts (and comparatively less like those derived from C6 or L2), that did not overgrow and expressed human DARPP32 (4.2.3). These grafts exhibited similar migratory patterns to H9 derived grafts, and in both groups cells were observed to have migrated into the corpus callosum, and occasionally into the contralateral hemisphere (Table 4.1). The main *in vitro* characterisation study further validated this finding. First, the qPCR analysis indicated that F1 derived cultures expressed key striatal genes in a similar pattern to H9 (4.2.5). Specifically, the mature MSN marker DARPP32 was highest in F1 cultures at day 40 compared to all other lines, followed by H9. Additionally, DRD2, was expressed at a higher rate in F1 than all other lines, but again H9 expression was second highest. This coincided with significantly higher expression of FOXP1, NOLZ1 and CTIP2 compared to various other lines across the differentiation time points; importantly, all three of these genes act upstream of DARPP32 expression (Arlotta *et al.*, 2008; Urbán *et al.*, 2010; Precious *et al.*, 2016). Immunocytochemical analysis (4.2.6) indicated that F1 cultures were highly efficient at differentiating into a neuronal

phenotype, and expressed similar numbers of cells expressing CTIP2 to other lines, surpassing both H9 and L11 at day 30, and equal to all other lines at day 40. Furthermore, they produced a significantly higher proportion of cells expressing DARPP32 than lines L2 and L11. Collectively, this indicates that the F1 line was not negatively affected in its ability to respond to the various differentiation protocols used by producing MSN-like cells (according to the small number of markers used), and performed in this as well as, or better than, an hESC control, or genomically identical hiPSCs derived from different tissues.

4.3.2 Cortex derived hiPSCs exhibit reduced efficiency of differentiation towards an MSN phenotype

Similar to the fibroblast derived line F1, the cortex derived line C6 was generated to examine how precise the effects of epigenetic memory were on hiPSC lines, as it was unknown how a tissue background that was neural, but of the “wrong” regional specificity, would impact an MSN differentiation. Throughout this chapter the C6 line has exhibited various differences compared to other lines. First, following neural transplantation (4.3.3) C6 produced large grafts that exhibited overgrowth, unlike the grafts derived from H9, F1, or L2 cultures. This was most apparent in animals 476 and 487, which received different batches of C6 MSN progenitors, but were transplanted on the same day. Both of these grafts exhibited widespread migration into the neighbouring cortex and external capsule, but in spite of producing the largest grafts with the most cells, there was little evidence of their inclusion in the corpus callosum spanning the two hemispheres. This was again different to both H9 and F1 derived grafts, which instead tended to exhibit migration into the corpus callosum, with a few cells identifiable in the neighbouring hemisphere. It is possible this tendency to overgrow following transplantation is an indication that this line had not yet committed to a maturing phenotype, which in turn could suggest conflicting endogenous gene expression or epigenetic apparatus that interfered with the MSN differentiation of these cells. However further research is required to determine if this is the case or not. Additionally, it must be highlighted that the total number of surviving grafts for this study was relatively low, and a more expansive study might find these differences are reduced.

In the main *in vitro* study using Protocol 3, qPCR analysis (4.3.5) revealed that at day 20, C6 derived cultures tended to express the lowest amount of the pan-GE marker DLX2, and the highest amounts of NKX2.1, although the expression of NKX2.1 was highly variable between C6 differentiations. DLX2 is a gene associated with early subpallium patterning and is understood to activate downstream of TGFβ activity, which

in this protocol is mediated by the addition of Activin (Feijen, Goumans and van den Eijnden-van Raaij, 1994; Maira *et al.*, 2010). Whilst NKX2.1 is expressed in both the LGE and MGE, higher concentrations of this gene are associated with more ventral regions like the MGE, and specifically interneurons (Noakes *et al.*, 2019). Therefore, this pattern of low DLX2, and high NKX2.1 expression could indicate that the C6 hiPSC line was less responsive to the addition of exogenous Activin, resulting in a proclivity to differentiate towards a wider variety of neuronal phenotypes. When examining key striatal genes more indicative of a mature MSN phenotype, C6 typically performed as well as the LGE derived lines L2 and L11, yet, C6 was found to produce a "purer" neuronal population compared to lines L2 and L11 (4.2.6). This also seems to indicate that these reduced levels of key striatal genes are not due to a reduced neuronal population, but rather a reduced efficiency within the present protocol to differentiate towards an MSN phenotype.

4.3.3 LGE derived hiPSCs perform similarly throughout the analysis

A general trend was observed for LGE derived lines L1, L2, and L11 to exhibit similarities across these characterisation experiments, which suggests that they are being influenced by similar underlying mechanisms. Provisional cell counts from the MSN differentiation conducted using Protocol 2, indicated that lines L1, L2 and L11 exhibited more DARPP32 than H9 or F1 (4.2.2). However, no comparison with M3 or C6 was made at the time to determine if this was specific to neuronal derived hiPSCs or LGE derived hiPSCs, and this dataset was not large enough for statistical analysis so it remains unknown how meaningful this distinction was. Later, evidence from the immunocytochemical staining conducted as part of the second *in vitro* experiments using Protocol 3 (4.2.6), suggested that cells strongly immuno-positive for DARPP32 are only present in L2 and L11 derived cultures at day 30, but were observable in C6, F1, and H9 derived cultures by day 40, again indicating a trend shared by LGE derived hiPSC lines that makes them distinct from other lines. Aside from these indications, it is also apparent from the second *in vitro* study using Protocol 3 that the lines L2 and L11 appeared to consistently underperform compared to the other lines with regards to gene expression of expected MSN fate markers. For example, the qPCR data (4.2.5) indicated that both L2 and L11 (along with C6) expressed less CTIP2 than F1 cultures. In addition, the immunocytochemistry counts (4.2.6) revealed that both had significantly fewer DARPP32 positive cells than F1 cultures, which was not the case for C6 cultures.

Differences were also found between LGE derived hiPSC lines. For example, the qPCR analysis indicated that L2 expressed significantly more Calbindin 1 at day 40,

and a trend to express more NKX2.1 at day 20, whereas L11 showed a trend to express higher levels of FOXP2 at all time points. Immunocytochemical analysis also indicated that L2 cultures contained significantly more CTIP2 expressing cells than L11 at day 30 (p<0.05). Collectively, this suggests that there are line specific differences between LGE derived hiPSCs, but that they have a general tendency to perform similarly under the present conditions for MSN differentiation, when compared to other genetically identical hiPSCs derived from different tissues.

4.3.4 Evidence that hiPSCs derived from neural tissues more rapidly commit to a neuronal phenotype.

It is possible, although not demonstrated conclusively here, that hiPSCs derived from neural tissues more rapidly commit to a neuronal phenotype. Early in this characterisation, when standardising a differentiation protocol that was suitable for all lines, there were differences observed that separated hiPSC lines L1 and M3 from F1 and H9. The first of these was the observation that, whilst undergoing dual SMAD inhibition to induce a forebrain progenitor phenotype, both L1 and M3 PSCs detached from the RGF Matrigel typically two days before they were due to be passaged, according to the initial protocol implemented. This was subsequently overcome by making an alteration to the standard hESC dual SMAD protocol. RGF Matrigel is a mixture of extracellular matrix proteins derived from Engelbreth-Holm-Swarm mouse sarcoma cells, comprising primarily laminin, collagen IV, and enactin (Hughes, Postovit and Lajoie, 2010). Collagen IV is known to be a poor substrate for the culture of neural stem cells, however this is not the case for laminin which is widely used as a matrix for the culture of developing and maturing neurons. Tate *et al.*, 2004 conducted a series of experiments that demonstrated that neurospheres adhered to fibronectin and laminin equally, and consistently with 6 fold the efficiency of collagen IV. However, they also established that this occurred through different integrins, with neurospheres binding to laminin through $\alpha_6\beta_1$ and fibronectin through $\alpha_5\beta_1$. Precisely why L1 and M3 detached early was not examined in the present study and no direct attempt was made to transfer these cells to a laminin matrix at this time. However, the observation that they repeatedly spontaneously detached from RGF Matrigel seems to be indicative of a more rapid change in cellular construct in these lines when exposed to conditions designed to induce a neuronal fate. Similarly, at a later time point L1 exhibited a more mature neuronal phenotype than was expected compared to cells derived from the hESC line H9, exhibiting widespread axonal growth and βIII-tubulin expression. It is important to note that this was not systematically studied in this thesis, and the morphological differences were less apparent in following differentiations using adapted

protocols, however this is again suggestive that L1 committed to a neuronal fate faster than the H9 line. This is further supported by the difficulties encountered passaging L1 at the second time point, which coincides with the time point when hESC line H9 was previously suitable for transplantation (Arber *et al.*, 2015). When passaged at this time, L1 exhibited reduced viability, and either reduced viability or proliferation following plating, compared to H9 and F1 lines cultured at the same time, again suggesting that these cells had matured beyond the level observed in H9 and F1.

Later experiments also provide some indication that lend support to this idea. First, initial examination of DARPP32 expression in MSNs undergoing Protocol 2 revealed more widespread expression of the MSN marker in lines L1, L2, and L11 compared to H9 and F1 (Figure 4.2). However, this data set was small and inappropriate for statistical analysis (n=1), as such, a later experiment was conducted on a larger dataset, and the findings generated here seem to conflict with this earlier finding (Figure 4.8). There are several differences between these two datasets, the most apparent are the differences in culture technique. The first were maintained in culture throughout the differentiation, and there is now significant evidence that these cells suffered huge viability losses when undergoing a freeze/thaw cycle at day 18. In contrast, the second data set was generated using Protocol 3, which included a freeze/thaw cycle as part of this experiment in an attempt to reduce bias at the point of measurement. This occurred on day 16 during the later stages of LGE specification via Activin treatment. The inclusion of a freeze/thaw cycle at this time point was not deemed to be detrimental to the cell viability of any line tested at the time, but without a comparison with cultures differentiated in parallel that did not undergo a freeze/thaw cycle, it is not possible to be certain this had no effect on the final fate specification observed here. An interesting observation in this second larger data set was the cells that highly expressed DARPP32 at day 30, which were only observed in L2 and L11 cultures at this time (Figure 4.9), but were observed at day 40 in all cultures (Figure 4.10). This again seems to be an indication of advanced maturity in the L2 and L11 cultures. Finally, the spontaneous electrical activity of these same cells was examined at day 40 using MEAs. Though the data here was highly inconsistent within lines, it is worth noting that lines L2, L11, and C6 all exhibited some signs of electrical activity more indicative of authentic neurons (Figure 4.12: L2 burst activity indicative of localised excitability; L11 and C6 exhibited large peak to peak amplitudes in spiking activity, indicative of attempted action potentials), and no such activity was observed in F1 or H9 derived cultures.

There is some evidence in the literature that suggests this tendency to advance towards a neuronal phenotype faster than other PSCs could be a consistent feature of

hiPSCs derived from neural tissues. For example, Roost *et al.*, 2017 generated hiPSCs from human fetal whole brain and fibroblasts from the same fetal donor. They found hiPSCs derived from fetal whole brain cells expressed βIII-tubulin earlier and at a higher level than hiPSCs derived from fibroblasts. Within this thesis there is no systematic study of this effect, as it was secondary to the main goal of determining differences in final MSN populations. Nonetheless, within the context of the evidence outlined above, there is sufficient reason to consider this collection of observations an indication of variability between these lines that may be related to tissue of origin.

4.3.5 Evidence that hiPSCs derived from neural tissues are less efficient at producing MSNs than hiPSCs derived from non-neural tissues

In direct contrast to this capacity for early neuronal specificity, there are several findings throughout this chapter that seem to indicate that hiPSCs derived from neural tissues are less efficient at producing DARPP32 positive MSN populations than a genomically identical hiPSC line derived from non-neural tissues, and a hESC control. Foremost of these is the previously discussed qPCR and immunocytochemistry data (4.2.5-4.2.6) collected during the main *in vitro* characterisation. Here, there was consistent evidence across multiple differentiations and cell lines that the fibroblast derived hiPSC line F1 produced the richest population of DARPP32 expressing neurons. This was also true of other key striatal genes that act upstream of DARPP32 and have been previously shown to be critical for MSN specification including NOLZ1, CTIP2, FOXP1 (Arlotta *et al.*, 2008; Urbán *et al.*, 2010; Precious *et al.*, 2016). This appears to be in direct conflict with the majority of the literature attempting to observe the effect of epigenetic memory of iPSCs (Introduction 1.5.4; Table 1.1). However, it is possible that this reduced capacity is a product of the methodology used here. Specifically, the differentiation protocol on which this work was based was tailored to the developmental time frame of hESCs, and it is therefore possible that critical windows in development are being missed if they do not occur at the same time across all lines examined, preventing the efficient differentiation of these cells towards an MSN phenotype. As previously discussed, there is evidence throughout this chapter that the neural derived hiPSCs do appear to mature faster than the hESC line and hiPSC line derived from fibroblast tissues. Therefore, it seems plausible that the differentiation capacity of these hiPSCs still needs to be further researched using optimised differentiation protocols that are tailored to each cell line to fully determine if the residual epigenetic memory of these cells can enhance an MSN differentiation.

4.3.6 Conclusions

In this chapter I aimed to characterise the differences in genomically identical hiPSCs derived from different tissues, to determine the effect that any residual epigenetic memory of tissue of origin had on an MSN differentiation. Across the experiments included in this chapter, it has become apparent that there are differences between hiPSC lines that cannot be attributed to genetic differences or variability in experimental conditions, although it remains undetermined if these differences are due to tissue specific epigenetic memory. Yet, there are trends that seem to indicate this: first, there is evidence from across these experiments that hiPSCs derived from non-neural tissues act more like hESCs whilst undergoing an MSN differentiation than they do genetically identical hiPSCs derived from neural tissues. There was also an observed trend for neural tissue derived hiPSCs to commit and mature into neuronal phenotypes faster than an other non-neural tissue derived hiPSC line, and a hESC control. It therefore seems somewhat surprising that hiPSCs derived from neural tissues do not efficiently differentiate towards an MSN phenotype, but given that this study uses a protocol which had originally been designed with developmental timings worked around a hESC line, it is perhaps possible that critical windows are being missed to properly direct these cells towards the correct phenotype. Alternatively, it is also possible that the retained epigenetic mechanisms are instead conflicting with the exogenous cues being used in culture, which is again plausible considering that regional specificity is largely dictated by precise spatial and temporal expression of morphogenic factors, and regulated by genes that act antagonistically in a gradient to downregulate other developmental pathways.

Further research will be required to determine if these effects are due to tissue specific epigenetic mechanisms that have been preserved in these hiPSCs, and to establish whether whatever is driving these differences can be a help, or only ever a hindrance, to the differentiation of these lines towards an MSN fate.

Chapter 5: Genome-wide Methylation Analysis

5.1 Introduction

DNA methylation is a heritable and flexible epigenetic mechanism that is understood to play a fundamental role in the regulation of gene transcription (Bird, 2002). Due to this functional relationship, DNA methylation has been shown to reflect tissue and cellular phenotype in both adult and fetal tissues (Nagae, *et al.*, 2011; Fernandez *et al.*, 2012, Lister *et al.*, 2013; Lokk *et al.*, 2014). Furthermore, during early embryogenesis the entire genome undergoes several hypo- and hyper- methylation cycles, and it is known to be highly conserved and regulated during development (Cedar and Bergman, 2012). During the generation of iPSCs, it is established that the entire genome undergoes a similar methylation shift, which expunges the previous methylome, by instead acquiring high levels of methylation across the genome to more closely resemble hESCs (Papp and Plath, 2013). This change in global methylation is however not immediate or absolute, and early passage iPSCs are known to retain a small degree of their previous epigenome. This includes some of the DNA methylation observed in the parent cells from which these iPSCs are derived, within iPSCs, this is referred to as their epigenetic memory of their tissue of origin (Kim *et al.*, 2010; Polo *et al.*, 2010). There is evidence that this retained methylation includes some tissue specific methylation that could potentially influence the differentiation of these iPSCs towards cellular fates similar to their tissue of origin (Introduction 1.5.4).

Previously across both Chapters 3 and 4, I have generated hiPSCs from fetal LGE tissues (L1, L2, and L11) and explored their capacity to differentiate towards an MSN fate. Through the work conducted in Chapter 4, I have established that these hiPSCs do indeed vary in their differentiation potential compared to genomically identical hiPSCs derived from control tissues (C6, F1), and genomically distinct hESCs (H9). Additionally, in Chapter 3, I presented some evidence that a critical difference between these lines is their variable epigenome that is indicative of their tissue of origin. However, to establish if these differences are indeed due to a retained epigenetic memory specific to the hiPSC tissue of origin, it is important to next analyse the epigenome of these cells. Such an analysis will also establish how similar the methylome of these hPSC-MSNs are to the fetal LGE tissues which we are attempting to emulate in culture.

The currently published data on DNA methylation in human fetal brains collectively establishes that DNA methylation is more highly conserved and regulated in human fetal brains compared to human child and human adult brains. Furthermore,

it is predictive of neuronal and glial cell phenotypes, and appears to serve a role mediating developmental gene expression (Introduction 1.5.3; Numata *et al.*, 2012; Lister *et al.*, 2013; Spiers *et al.*, 2015; Roost *et al.*, 2017). However, to date there are no published studies that describe the global DNA methylation differences between the various structures of the human fetal brain including the WGE or its substructures (LGE, MGE, CGE), although examination of post-natal rat brains has indicated DNA methylation can vary according to the developmental structures of the neonatal brain (Simmons *et al.*, 2013). Therefore, it is also important to establish how the methylome of primary LGE differs from other neural tissues, and to establish the common methylation profile of primary fetal LGE. This will allow for a comparison between the common methylome of primary fetal LGE tissues and the aforementioned hPSC-MSNs.

Due to the scarcity of fetal tissues, and the time and material costs of generating MSNs from hPSCs, it was preferable to use technology that offers precise and sensitive measures to maximise the quality of data generated, whilst minimising the need for a high number of samples per condition. Furthermore, as there is no present data on the methylation profiles of primary fetal LGE tissues, it was important to use technology that offered genome-wide profiling of samples. For these reasons, I elected to use the Infinium Human Methylation 450 bead chips (450K) from Illumina, which targets CpGs across the whole genome, which are associated with ≈99% of RefSeq genes and include both CpG rich and reduced regions. The resulting data has been shown to be highly reproducible and sensitive to subtle differences between sample methylation (R^2=0.992 between technical replicates; Bibikova *et al.*, 2011). Furthermore, since its release, this technology has been used broadly to characterise DNA methylation across a diverse range of tissues, and the wider research community have consequently further refined this technology (Nordland *et al.*, 2013; Zhou *et al.*, 2017), and improved the methods of data analysis (Morris & Beck, 2015).

1.6.1 Summary of Aims

In this chapter I aimed to explore if and how the methylation profile of *in vitro* cultured hPSC-MSNs differed, and how this varied by hPSC line. I also sought to explore the methylation profiles of primary human fetal LGE tissues and compare them to similar and dissimilar fetal tissues to establish their methylation profile. Finally, I aimed to compare the methylome of primary fetal LGE tissues to hPSC-MSNs to establish if any retained DNA methylation was advantageous to the LGE derived hiPSCs as they underwent an MSNs differentiation that encouraged a more authentic phenotype.

5.2 Results

Thirty two samples were processed for genome wide methylation analysis (Methods 2.2.3, see Table 5.1 for sample details). In summary these were: single samples of the primary LGE, MGE, cortex, and fibroblast tissues (all from SWIFT 2285, CRL=42.1mm, 61 days p.c.) that had been used to generate the hiPSCs studied throughout this thesis; single samples of two additional primary LGEs (from different SWIFT donors: SWIFT 2415, CRL=53.4 mm, 69 days p.c.; and SWIFT 2451, CRL=37.8 mm, 70 days p.c.); single samples of three Primary WGEs (from different SWIFT donors: SWIFT 2230, CRL≈55 mm, 65 days p.c.; SWIFT 2240, CRL=40 mm, 62 days p.c.; and SWIFT 3005, CRL=39.5 mm, 59 days p.c.), single samples of three WGEs (termed 'Cultured WGE') that had been taken from the same WGE SWIFT donors as above (SWIFT 2230, 2240, 3005) and allowed to spontaneously differentiate for four weeks *in vitro* (Methods 2.1.3; WGE maturation culture conducted independently by Dr. Sophie Precious); and the remaining twenty samples comprised a set of four samples from each of the hPSC lines L1, L2, L11, F1, and H9. Each set from these lines included a single pluripotent sample (described previously in Chapter 3; passage 9-13 for hiPSCs, passage 23 for hESC), and three further samples harvested at day 30 from independent rounds of hPSC-MSN differentiations (harvested from cells plated during the first round of differentiation under Protocol 2, described in section 4.2.2 of Chapter 4. Differentiations were initiated between passages 12-18 for hiPSCs and 18-22 for hESCs). Of these samples, one (hESC H9 MSN ii) was removed from further analysis because 62% of the sample probes failed, whereas all other samples were found to have <1% failed probes (Table 5.1). During the course of the analysis the data was compiled into multiple data sets, these are described in Table 5.2.

Table 5.1 – Sample information

Sample Name	Genetic Identity*	Experimental Group	Cell Phenotype	% of failed probes	Datasets:
Primary LGE 2415	SWIFT 2415	Primary LGE/GE	Fetal LGE	0.1%	1, 2, 3, 5, 6
Primary LGE 2451	SWIFT 2451	Primary LGE/GE	Fetal LGE	0.01%	1, 2, 3, 5, 6
Primary LGE 2285	SWIFT 2285	Primary LGE/GE	Fetal LGE	0.06%	1, 2, 3, 5, 6
Primary MGE 2285	SWIFT 2285	Primary MGE/GE	Fetal MGE	0.05%	1, 2, 3, 6
Primary CTX 2285	SWIFT 2285	Primary Cortex	Fetal cortex	0.1%	1, 2, 3
Primary FIB 2285	SWIFT 2285	Primary Fibroblast	Fetal fibroblast	0.05%	1, 2, 3
Primary WGE 2230	SWIFT 2230	Primary WGE/GE	Fetal WGE	0.05%	1, 2, 3, 5, 6
Primary WGE 2240	SWIFT 2240	Primary WGE/GE	Fetal WGE	0.07%	1, 2, 3, 5, 6
Primary WGE 3005	SWIFT 3005	Primary WGE/GE	Fetal WGE	0.06%	1, 2, 3, 5, 6
Cultured WGE 2230	SWIFT 2230	Cultured WGE	Fetal WGE	0.08%	1, 2, 3
Cultured WGE 2240	SWIFT 2240	Cultured WGE	Fetal WGE	0.06%	1, 2, 3
Cultured WGE 3005	SWIFT 3005	Cultured WGE	Fetal WGE	0.06%	1, 2, 3
hiPSC L1 Pluri	SWIFT 2285	Pluri-L1	hiPSC	0.06%	1, 3
hiPSC L2 Pluri	SWIFT 2285	Pluri-L2	hiPSC	0.17%	1, 3
hiPSC L11 Pluri	SWIFT 2285	Pluri-L11	hiPSC	0.06%	1, 3
hiPSC F1 Pluri	SWIFT 2285	Pluri-F1	hiPSC	0.08%	1, 3
hESC H9 Pluri	ESC H9	Pluri-H9	hESC	0.05%	1, 3
hiPSC L1 MSN i	SWIFT 2285	L1-MSN/L-MSN	hiPSC-MSN	0.06%	1, 3, 4, 5, 6
hiPSC L1 MSN ii	SWIFT 2285	L1-MSN/L-MSN	hiPSC-MSN	0.05%	1, 3, 4, 5, 6
hiPSC L1 MSN iii	SWIFT 2285	L1-MSN/L-MSN	hiPSC-MSN	0.11%	1, 3, 4, 5, 6
hiPSC L2 MSN i	SWIFT 2285	L2-MSN/L-MSN	hiPSC-MSN	0.72%	1, 3, 4, 5, 6
hiPSC L2 MSN ii	SWIFT 2285	L2-MSN/L-MSN	hiPSC-MSN	0.06%	1, 3, 4, 5, 6
hiPSC L2 MSN iii	SWIFT 2285	L2-MSN/L-MSN	hiPSC-MSN	0.50%	1, 3, 4, 5, 6
hiPSC L11 MSN i	SWIFT 2285	L11-MSN/L-MSN	hiPSC-MSN	0.36%	1, 3, 4, 5, 6
hiPSC L11 MSN ii	SWIFT 2285	L11-MSN/L-MSN	hiPSC-MSN	0.05%	1, 3, 4, 5, 6
hiPSC L11 MSN iii	SWIFT 2285	L11-MSN/L-MSN	hiPSC-MSN	0.06%	1, 3, 4, 5, 6
hiPSC F1 MSN i	SWIFT 2285	F1-MSN	hiPSC-MSN	0.06%	1, 3, 4, 5, 6
hiPSC F1 MSN ii	SWIFT 2285	F1-MSN	hiPSC-MSN	0.31%	1, 3, 4, 5, 6
hiPSC F1 MSN iii	SWIFT 2285	F1-MSN	hiPSC-MSN	0.06%	1, 3, 4, 5, 6
hESC H9 MSN i	ESC H9	H9-MSN	hESC-MSN	0.05%	1, 3, 4, 5, 6
hESC H9 MSN ii	ESC H9	H9-MSN	hESC-MSN	62.56%	---
hESC H9 MSN iii	ESC H9	H9-MSN	hESC-MSN	0.05%	1, 3, 4, 5, 6

**Genetic identity is described here as SWIFT donor number or hESC line, to clarify which samples are isogenic.*
Key: Primary = tissue dissected and prepared for analysis without any in vitro culture. Cultured = tissue that has been allowed to spontaneously differentiate in culture (methods 2.1.3). CTX = cortex. FIB = fibroblast. Pluri = undifferentiated/pluripotent sample. MSN = hPSCs at day 30 of an MSN directed differentiation (Methods 2.1.7). i/ii/iii = *independent round of MSN differentiations.*

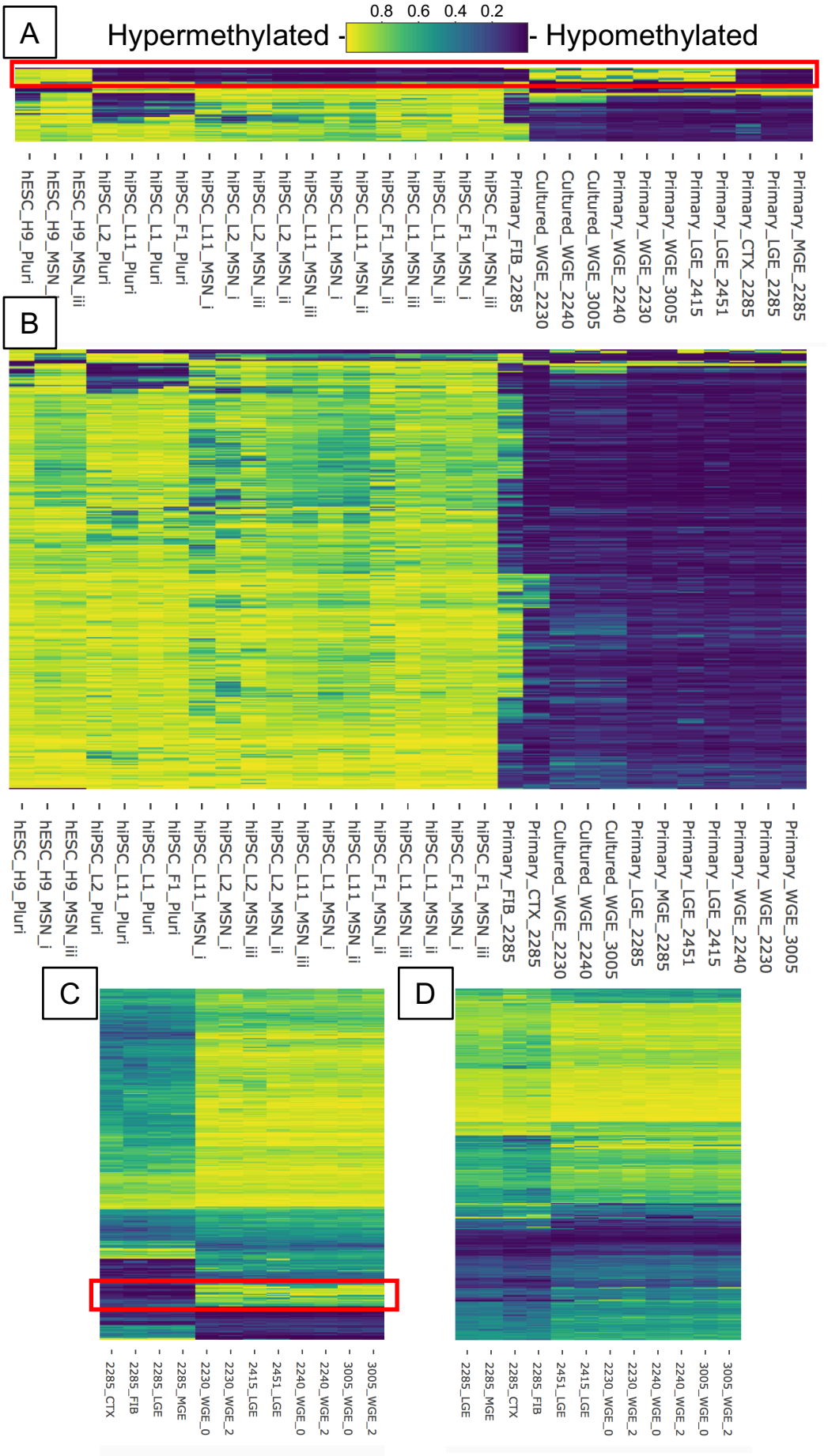
Table 5.2 – Dataset information

No.	Chapter Section(s):	No. of probes:	Groups included:	Purpose:
1	5.2.1	408,233	All	Exploratory analysis, initial identification of SWIFT 2285 specific CpG methylation
2	5.2.1	410,018	All Primary tissues (including cultured WGE)	To locate SWIFT 2285 specific CpG methylation, gene enrichment analysis
3	5.2.1, 5.2.2, 5.2.3	407,406	All	Exploratory analysis
4	5.2.3	408,391	L1-MSNs; L2-MSNs; L11-MSNs; F1-MSNs; H9-MSNs	Exploratory analysis
5	5.2.4	407,739	Primary LGE; Primary WGE, Cultured WGE; L1-MSNs; L2-MSNs; L11-MSNs; F1-MSNs; H9-MSNs	DMP analysis
6	5.2.4, 5.2.5	407,801	Primary GE (includes all LGE, MGE and WGE samples); L-MSN (includes all L1-, L2-, and L11-MSNs); F1-MSNs; and H9-MSNs	DMP analysis, gene enrichment analysis

5.2.1 Identification and exclusion of SWIFT 2285 specific methylation

During exploratory analysis (dataset 1) of the 31 samples, it became apparent that there was a risk of genetic bias in the analysis due to the high number of samples derived from SWIFT 2285. Specifically, there is a distinct methylation signature unique to samples derived from SWIFT 2285, which appears in all SWIFT 2285 derived primary tissues, and remains observable through induction to pluripotency, and *in vitro* differentiation towards an MSN phenotype (Figure 5.1A, red box). It was noted that maintaining these specific CpGs in the analysis could potentially bias results, and prevent analysis of any methylation specific to tissue of origin retained by the hiPSCs, and therefore needed to be removed before data exploration.

To do so, a second dataset (dataset 2) was generated to locate which probes were significantly differentially methylated in SWIFT 2285 samples compared to other fetal tissues. The 12 fetal tissue samples were divided into two groups based on genetic identity rather than phenotype: LGE, MGE, cortex, and fibroblast tissues taken from SWIFT 2285 were included in the first group; whereas the second group comprised LGE tissues taken from SWIFTs 2415 and 2451, and both primary and cultured WGE samples taken from SWIFTs 2230, 2240, and 3005. hESCs, hiPSCs, and hPSC-MSNs were excluded from this comparison to avoid the unwanted removal of hESC, hiPSC and hPSC-MSN phenotype differences. Differentially methylated probe (DMP) analysis (Methods 2.5.2.3) was conducted on dataset 2, and a total of 827 significant DMPs/CpGs were observed between these two groups (p<0.05, Figure 5.1C shows all 827 significant DMPs, the red box indicates the same CpGs identifiable in 5.1A).


A
Hypermethylated -
0.8 0.6 0.4 0.2
- Hypomethylated
Primary_MGE_2285
Primary_LGE_2285
Primary_CTX_2285
Primary_LGE_2451
Primary_LGE_2415
Primary_WGE_3005
Primary_WGE_2230
Primary_WGE_2240
Cultured_WGE_3005
Cultured_WGE_2240
Cultured_WGE_2230
Primary_FIB_2285
hiPSC_F1_MSN_iii
hiPSC_F1_MSN_i
hiPSC_L1_MSN_ii
hiPSC_L1_MSN_iii
hiPSC_F1_MSN_ii
hiPSC_L11_MSN_ii
hiPSC_L1_MSN_i
hiPSC_L11_MSN_iii
hiPSC_L2_MSN_ii
hiPSC_L2_MSN_iii
hiPSC_L2_MSN_i
hiPSC_L11_MSN_i
hiPSC_F1_Pluri
hiPSC_L1_Pluri
hiPSC_L11_Pluri
hiPSC_L2_Pluri
hESC_H9_MSN_iii
hESC_H9_MSN_
hESC_H9_Pluri
B
Primary_WGE_3005
Primary_WGE_2230
Primary_WGE_2240
Primary_LGE_2451
Primary_LGE_2415
Primary_MGE_2285
Primary_LGE_2285
Cultured_WGE_3005
Cultured_WGE_2240
Cultured_WGE_2230
Primary_CTX_2285
Primary_FIB_2285
hiPSC_F1_MSN_iii
hiPSC_F1_MSN_j
hiPSC_L1_MSN_ii
hiPSC_L1_MSN_iii
hiPSC_F1_MSN_ii
hiPSC_L11_MSN_ii
hiPSC_L1_MSN_i
hiPSC_L11_MSN_iii
hiPSC_L2_MSN_ii
hiPSC_L2_MSN_iii
hiPSC_L2_MSN_i
hiPSC_L11_MSN_i
hiPSC_F1_Pluri
hiPSC_L1_Pluri
hiPSC_L11_Pluri
hiPSC_L2_Pluri
hESC_H9_MSN_iii
hESC_H9_MSN_i
hESC_H9_Pluri
C
3005_WGE_2
3005_WGE_0
2240_WGE_2
2240_WGE_0
2451_LGE
2415_LGE
2230_WGE_2
2230_WGE_0
2285_MGE
2285_LGE
2285_FIB
2285_CTX
D
3005_WGE_2
3005_WGE_0
2240_WGE_2
2240_WGE_0
2415_LGE
2230_WGE_0
2230_WGE_2
2451_LGE
2285_FIB
2285_CTX
2285_MGE
2285_LGE

Previously, it has been determined that there are autosomal sex differences in fetal brain methylation (Spiers *et al.*, 2015). Therefore, the 827 CpGs were next compared to these previously identified sex specific CpGs (n=521 CpGs), but there was not a single CpG present in both lists, indicating this is not due to sex differences between samples, and instead are a product of genetic variation. Genes associated with these DMPs (n=57) were analysed using gene enrichment analysis (Methods 2.5.2.4) to identify if there was a functional association between these differentially methylated genes, three genes (SIK3, RORA, and PDK1) were significantly associated with the regulation of cellular carbohydrate metabolic process (p=0.002) and regulation of glucose metabolic process (p=0.006) in the GO Biological Process (2018) database (which aims to group genes by general biological function), but no tissue specific enrichment was observed.

When these significant DMPs were removed from the dataset (dataset 3) the SWIFT 2285 specific methylation signature was no longer observed (Figure 5.1B, D). Subsequently, these 827 CpGs were filtered from all future datasets to reduce the bias effect of SWIFT 2285s genome specific methylation.

Figure 5.1: (Overleaf) Heatmaps of CpG Methylation before and after removal of SWIFT 2285 specific methylation.

A. Exert of the top 1000 most variable CpGs between samples in dataset 1. Red box indicate a series of CpGs that appear hypomethylated in all SWIFT 2285 samples, but hypermethylated in all other samples.
B. The top 1000 most variable CpGs between samples in dataset 3, following the removal of the 827 CpGs identified using dataset 2.
C. Summary of the 827 significantly differentially methylated CpGs between groups in dataset 2. Red box indicates the same differentially methylated CpGs as seen in A.
D. Summary of the most differentially methylated CpGs between groups in dataset 2 (none are significant), after the removal of the 827 CpGs.

5.2.2 Unsupervised hierarchical clustering analysis

Following the exclusion of SWIFT 2285 specific methylation, exploratory analysis was undertaken (dataset 3). Unsupervised hierarchical clustering (Methods 2.5.2.2) was conducted on the global methylation data collected from all CpG probes, to explore whether DNA methylation could be used to distinguish and group the samples. The results are expressed as a dendrogram and are discussed in reverse clustering order (Figure 5.2).

The final clustering, and therefore greatest distinction, was made between the primary fibroblast sample from SWIFT 2285 and the remaining thirty samples (Figure 5.2 (1)), indicating the unique, tissue specific, nature of its methylome. The three H9-ESC derived samples were the next to separate from the main cluster (2). This group included the pluripotent sample and the two differentiated H9-MSN samples, suggesting that line specific methylation patterns of either genomic or hESC origin were overall more defining than the cellular phenotype (hPSC vs hPSC-MSN), although this distinction between cell phenotype was later identifiable (7). The next grouping (3) included all hiPSC pluripotent samples, regardless of tissue of origin. Three hiPSC derived MSN samples then clustered away from the main group (4) containing two L2-MSN samples and one L11-MSN sample. The next division (5) separated all primary neural samples (LGE, MGE, WGE, and cortex) from the remaining samples. Then the main group clustered into two distinct groups (6); the first contained the three WGE samples that had been allowed to spontaneously differentiate for 4 weeks *in vitro* (Cultured WGE); and the second contained the remaining nine hiPSC derived MSN samples. The next division (7) was within the H9 cluster, separating the pluripotent sample from the differentiated MSN samples, indicating that the difference in H9 cell phenotype was less pronounced than all previously mentioned distinctions. The final noteworthy distinction (8) was then made between the primary cortex sample and the various primary striatal tissues. All further groupings were within samples of similar phenotypes or from similar experimental conditions. Of note, the hiPSC derived MSN samples did not separate in this analysis based on tissue of origin, which suggests any effects caused by tissue specific residual epigenetic memory are largely overshadowed by other defining features of these populations, such as the variability created during the differentiation process. This is further supported by the early separation of three hiPSC-MSN samples (4), which clearly indicates that the differentiation process did not have a uniform effect on the methylome of these samples in all instances.

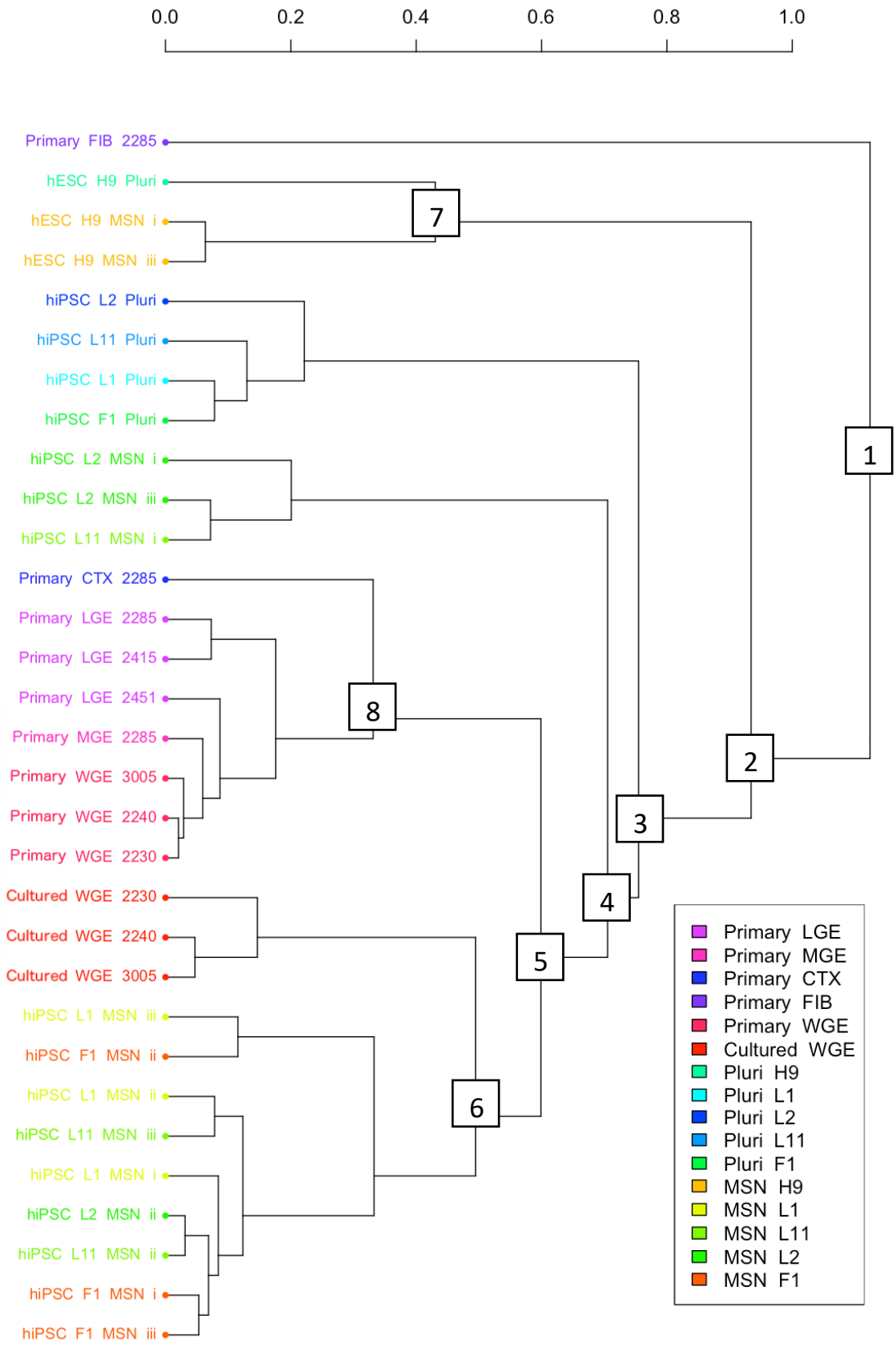


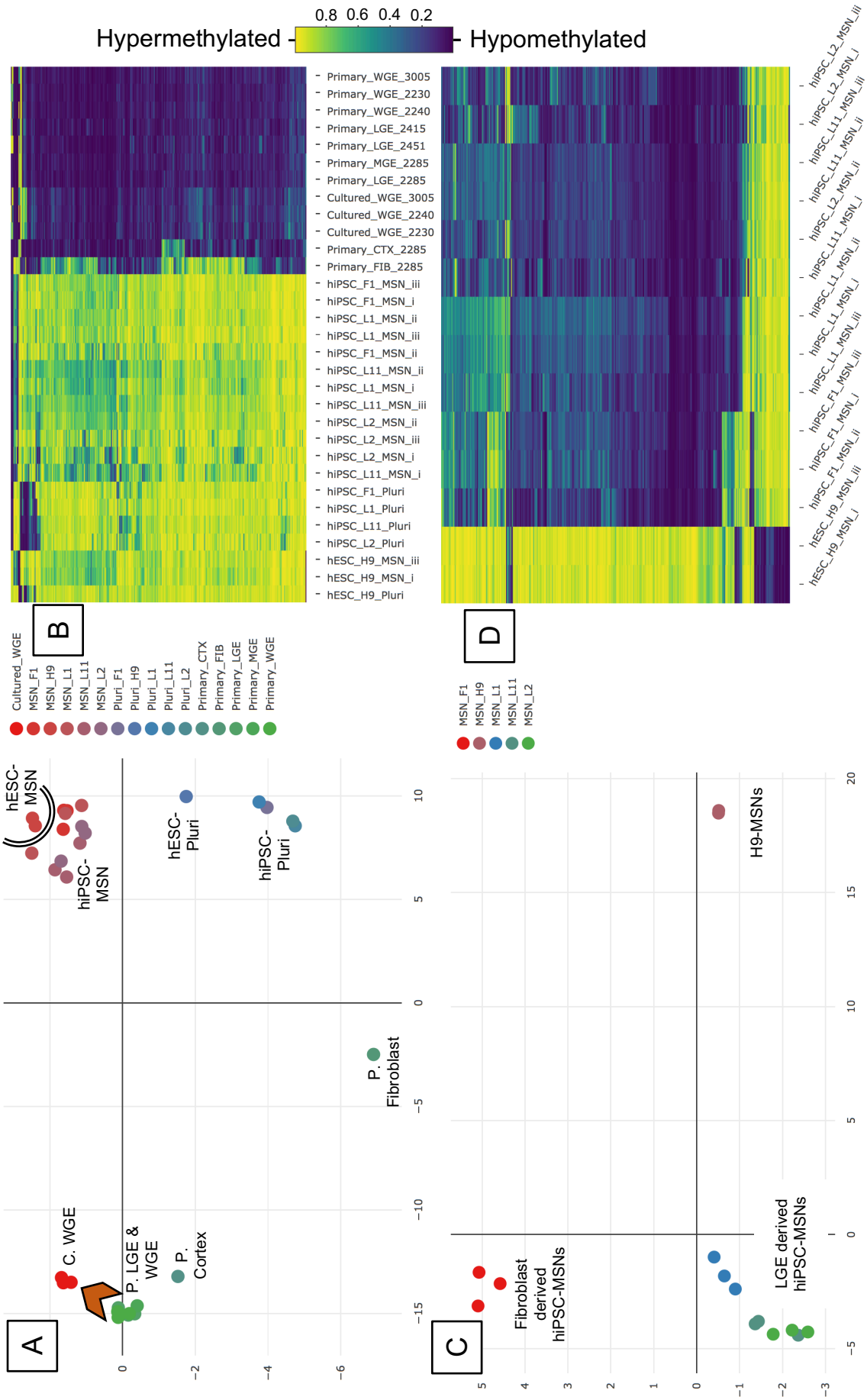
Figure 5.2: Dendrogram of unsupervised hierarchical clustering analysis undertaken across all samples using dataset 3.
The axis describes the calculated Euclidean distance between clusters, with 0 representing samples that are indistinguishable, and gradual increases representing increased difference between clusters. The final eight clusters to form are annotated at the point of clustering in reverse order (1-8). Samples are coloured by experimental group.

5.2.3 Examination of the 1000 most variable CpGs

Next samples were compared using only the 1000 CpG probes that exhibited the largest range of β-values across all samples, indicating which CpGs varied most in methylation status. These data are expressed using a heatmap of the methylation status of these CpGs (Figure 5.3A), and through an MDS cluster plot (Figure 5.3B). From the heatmap, it is apparent that there is a higher degree of methylation across all hESC and hiPSC samples compared to primary and cultured primary tissues, which is not obviously ameliorated by the *in vitro* differentiation process. Distinct trends were still noticeable between samples based on phenotype, reflective of the unsupervised hierarchical clustering patterns based on all CpGs outlined previously. For example, the primary fibroblast sample exhibited a methylation pattern unlike all other samples, and WGE samples appear to become more methylated following 4 weeks of *in vitro* culture compared to their primary tissue counterparts (Figure 5.3.A). These similarities and differences are also visible on the MDS plot (Figure 5.3B), as the primary fibroblast sample clusters independently from all other samples, and there is a subtle division between the primary WGE and cultured WGE indicative of this increase in methylation. However, further distinctions are also visible, the most notable being that *in vitro* MSNs cultured from hiPSCs and hESCs clustered in a group loosely together, but separately from pluripotent samples (Figure 5.3B).

To determine if tissue of origin still affected hiPSCs after an *in vitro* MSN differentiation, a new dataset (dataset 4) was generated that included only the hPSC-MSNs samples. hESC-MSNs were notably more methylated than hiPSC-MSNs (Figure 5.3D) and cluster separately from hiPSC-MSNs (Figure 5.3C). Furthermore, whilst the differences between LGE derived hiPSC-MSNs and fibroblast derived hiPSC-MSNs are not as pronounced compared to hESC-MSNs (Figure 5.3D), their methylomes still cluster according to source tissue (Figure 5.3C), suggesting that they have retained a

Figure 5.3: (Overleaf) Heatmaps and MDS plots exploring the 1000 most variable CpGs across samples in dataset 3 and 4.
A. MDS plot generated using the 1000 most variable CpGs across samples in dataset 3. Clusters are labelled, and individual samples are coloured by group. Cultured WGE clusters separately from Primary WGE and LGE, indicating a methylation shift is acquired *in vivo* (highlighted by an orange chevron). hPSC-MSNs cluster loosely together, a black line indicates which samples are hESC-MSNs and which are hiPSC-MSNs.
B. A heatmap generated using the 1000 most variable CpGs across samples in dataset 3. hPSC derived samples are more widely methylated than primary tissues, regardless of cellular phenotypes and experimental groups. Trends are identifiable between samples that cluster together in A.
C. MDS plot generated using the 1000 most variable CpGs across samples in dataset 4. Clusters are labelled, and individual samples are coloured by group. hPSC-MSNs cluster according to their tissue of origin rather than only by cell line, as exhibited by separate hESC, LGE-hiPSC and fibroblast-hiPSC clusters, and the strong similarity between all 9 LGE-hiPSC derived samples.
D. A heatmap generated using the 1000 most variable CpGs across samples in dataset 4. H9-MSNs are typically more methylated than all hiPSC-MSNs, and methylation trends can be observed within lines and groups that led to the sample clustering observed in C.


A
C. WGE
P. LGE & WGE
P. Cortex
P. Fibroblast
hESC-MSN
hiPSC-MSN
hESC-Pluri
hiPSC-Pluri
B
Hypermethylated
Hypomethylated
0.8 0.6 0.4 0.2
Primary_WGE_3005
Primary_WGE_2230
Primary_WGE_2240
Primary_LGE_2415
Primary_LGE_2451
Primary_MGE_2285
Primary_LGE_2285
Cultured_WGE_3005
Cultured_WGE_2240
Cultured_WGE_2230
Primary_CTX_2285
Primary_FIB_2285
hiPSC_F1_MSN_iii
hiPSC_F1_MSN_i
hiPSC_L1_MSN_ii
hiPSC_L1_MSN_iii
hiPSC_F1_MSN_ii
hiPSC_L11_MSN_ii
hiPSC_L1_MSN_i
hiPSC_L11_MSN_iii
hiPSC_L2_MSN_ii
hiPSC_L2_MSN_iii
hiPSC_L2_MSN_i
hiPSC_L11_MSN_i
hiPSC_F1_Pluri
hiPSC_L1_Pluri
hiPSC_L11_Pluri
hiPSC_L2_Pluri
hESC_H9_MSN_iii
hESC_H9_MSN_i
hESC_H9_Pluri
Cultured_WGE
MSN_F1
MSN_H9
MSN_L1
MSN_L11
MSN_L2
Pluri_F1
Pluri_H9
Pluri_L1
Pluri_L11
Pluri_L2
Primary_CTX
Primary_FIB
Primary_LGE
Primary_MGE
Primary_WGE
C
Fibroblast derived hiPSC-MSNs
H9-MSNs
LGE derived hiPSC-MSNs
D
hiPSC_L2_MSN_iii
hiPSC_L2_MSN_i
hiPSC_L11_MSN_iii
hiPSC_L11_MSN_ii
hiPSC_L2_MSN_ii
hiPSC_L11_MSN_i
hiPSC_L1_MSN_ii
hiPSC_L1_MSN_iii
hiPSC_F1_MSN_iii
hiPSC_F1_MSN_i
hiPSC_F1_MSN_ii
hESC_H9_MSN_iii
hESC_H9_MSN_i
MSN_F1
MSN_H9
MSN_L1
MSN_L11
MSN_L2

methylation signature indicative of their tissue of origin. Finally, there is some observed separation between LGE derived hiPSC-MSNs here, which suggests that there are also subtle line specific patterns that can persist between hiPSCs derived from the same tissues (Figure 5.3C).

5.2.4 Global DMP analysis

A DMP analysis (Methods 2.5.2.3) was conducted on all experimental groups that contained a minimum of two samples for analysis (Table 5.1; Primary LGE, Primary WGE, Cultured WGE, L1-MSNs, L2-MSNs, L11-MSNs, F1-MSNs, H9-MSNs). The results of this comparison are detailed in Table 5.3. Groups that were experimentally similar routinely exhibited the least number of significantly different DMPs. For instance, Primary LGE and Primary WGE were observed to have 0 DMPs between them, and between the groups of LGE-hiPSC derived MSNs there were only 2 DMPs between MSNs derived from lines L2 and L11 which had been noted previously to have less consistent methylation profiles. However, when compared to other samples with different phenotypes, differences between these otherwise similar groups became more apparent, for example, between the Primary LGE and L1-MSN groups there were 54,407 DMPs, whereas between the Primary WGE and L1-MSN groups there were 82,643 DMPs. The difference of 28,236 DMPs between these comparisons could potentially be a consequence of small sample sizes used in these analyses. It is therefore possible that a larger study would have sufficient power to identify individual DMPs between similar tissues and experimental groups. In light of this, all subsequent analyses were therefore restricted to comparisons of the most prominent differences between groups.

Table 5.3 – Number of significantly differentially methylated probes (DMPs) between sample groups with adjusted p<0.01, generated using dataset 5.

	P.LGE	P.WGE	C.WGE	L1-MSN	L2-MSN	L11-MSN	F1-MSN	H9-MSN
Primary LGE		0	5920	54407	16456	29923	38026	49210
Primary WGE	0		13581	82643	31197	49605	61989	88751
Cultured WGE	5920	13581		48353	13738	24553	34193	44831
L1-MSNs	54407	82643	48353		0	0	28	3053
L2-MSNs	16456	31197	13738	0		2	6	1089
L11-MSNs	29923	49605	24553	0	2		23	2575
F1-MSNs	38026	61989	34193	28	6	23		1950
H9-MSNs	49210	88751	44831	3053	1089	2575	1950	

Key: P=Primary, C=Cultured

For this reason, a final more conservative dataset (dataset 6) was generated that sought to merge similar groups together to increase the power of this analysis, at the potential loss of some tissue and hPSC line specificity. Primary LGE, Primary WGE, and the Primary MGE samples were pooled into a Primary Ganglionic Eminence group (termed Primary GE; Table 5.1). Cultured WGE samples were not included in this group because their methylation profile was significantly different from both Primary LGE and Primary WGE samples, as evidenced by the significant number of DMPs observed previously (Table 5.4). Furthermore, L1-MSNs, L2-MSNs, and L11-MSNs were pooled into an L-MSN group (Table 5.1). A second DMP analysis was conducted on these newly grouped samples, all DMPs are significant at $p<0.01$ (full results detailed in Table 5.4). Primary GEs were differentially methylated from all hPSC-MSN groups at approximately 25% of all CpG sites tested (114393 DMPs between Primary GE and L-MSNs; 102872 DMPs between Primary GEs and F1-MSNs; 100760 DMPs between Primary GEs and H9-MSNs). By comparison, there were far fewer differences observable between hiPSC-MSNs and hESC-MSNs (4561 DMPs between H9-MSNs and L-MSNs; and 1965 DMPs between H9-MSNs and F1-MSNs). Finally, there were fewest DMPs between hiPSC-MSNs derived from hiPSCs derived from different tissues of origin (276 DMPs between L-MSNs and F1-MSNs), indicating that there is a methylation signature that separates these hiPSCs that is dependent on tissue of origin, but these lines were still the most similar suggesting other features between lines also cause methylation differences.

Table 5.4 – Number of significantly differentially methylated probes (DMPs) between sample groups with adjusted p<0.01, generated using dataset 6.

	Primary GE	L-MSNs	F1-MSNs	H9-MSNs
Primary GE		114393	102872	100760
L-MSNs	114393		276	4561
F1-MSNs	102872	276		1965
H9-MSNs	100760	4561	1965	

5.2.5 Gene Enrichment Analysis of Differentially Methylated Genes

Gene enrichment analysis (Methods 2.5.2.4) was conducted on the genes associated with significant DMPs between groups in dataset 6 (referred to herein as differentially methylated genes). DMPs were filtered to only include those with a Δβ-value difference of 0.5 between groups representing a 50% difference in population methylation between groups, to increase the likelihood that the DMPs analysed were biologically relevant. These DMPs were then divided into hyper- and hypo-methylated groups, and finally DMPs without known associated genes were removed from the analysis. A summary of results are detailed in Tables 5.5-5.7, adjusted *p* value is given for each term in each comparison, greyed out *p* values are non-significant

Table 5.5 – Summary of gene enrichment results from significantly differentially methylated genes in Primary GE and hPSC-MSNs

Methylation Status	Data Base	Term	Adjusted *p* value		
			Vs. L-MSNs	Vs. F1-MSNs	Vs. H9-MSNs
Hypo-methylated in Primary GE	Allen Brain Atlas Up-regulated	Striatum	0.000147682	0.00022169	4.63022E-05
		Putamen	0.000223555	0.00048542	0.000415634
		Caudate nucleus	0.000545271	0.004605231	0.003469212
	ARCHS4 Tissues	Astrocyte	5.60263E-24	1.86313E-25	2.34008E-19
		Cerebral Cortex	1.74043E-24	4.95918E-25	6.86715E-29
		Brain (Bulk)	9.26908E-26	1.64376E-23	1.11572E-31
		Superior Frontal Gyrus	5.54882E-25	1.64319E-21	1.40801E-27
		Motor Neuron	1.74677E-21	4.51428E-21	2.04775E-22
		Fetal Brain	3.63998E-19	2.10352E-19	6.17495E-19
		Prefrontal Cortext	3.63584E-17	2.45371E-17	1.25977E-23
		Dorsal Striatum	9.91063E-12	2.07E-08	1.96558E-12
	GO Biological Process 2018	Nervous system development	3.7425E-10	4.02944E-09	3.21967E-10
		Synapse assembly	3.80871E-07	0.000169691	1.07861E-05
		Axon guidance	2.26731E-05	1.5285E-05	1.51089E-05
		Axonogenesis	8.04354E-05	1.5285E-05	1.24123E-06
		Regulation of cell migration	0.014999505	0.008035587	0.003732031
		Positive regulation of neuron projection development	0.022045683	0.021203901	0.025486359
		Positive regulation of synaptic transmission	0.022297929	0.021203901	0.021522687
		Modulation of excitatory postsynaptic potential	0.022423928	0.07246016	0.11569922
		Neuron projection morphogenesis	0.037239437	0.039023643	0.000101461
		Regulation of neurotransmitter receptor activity	0.044063156	0.067090081	0.285722703
		Positive regulation of cell differentiation	0.057021704	0.080683784	0.004015063
		Cell morphogenesis involved in neuron differentiation	0.164229004	0.059816116	0.001169334
Hyper-methylated in Primary GE	GO Biological Process 2018	Epidermal cell differentiation	N/A	0.03940305	0.025565857
		Peptide cross-linking	N/A	0.03940305	0.019470429
		Sensory perception of smell	N/A	0.03940305	N/A
		Keratinocyte differentiation	N/A	0.03940305	0.023525622
		Skin development	N/A	0.050708058	0.069078235

Table 5.6 – Summary of Gene enrichment results from significantly differentially methylated genes between hPSC-MSN groups

	Database	Term	Adjusted p value	
			Vs. L-MSNs	Vs. F1-MSNs
Hyper-methylated in H9-MSNs	ARCHS4	CEREBRAL CORTEX	0.00115364	0.034358499
		SUPERIOR FRONTAL GYRUS	0.00087737	0.034358499
		BRAIN (BULK)	0.00020233	0.042700832

Table 5.7 – Gene enrichment results from significantly differentially methylated genes between L-MSNs and F1-MSNs

	Database	Term	Vs.F1-MSNs	Genes
Hypo-methylated in L-MSNs	Allen Brain Atlas Down-regulated	Main olfactory bulb, inner plexiform layer	0.10032111	FAM135B;SMARCA2
		Olfactory bulb glomerular layer	0.10032111	CNTNAP1;FAM135B;SMARCA2
		Superficial stratum of olfactory bulb	0.10032111	CNTNAP1;FAM135B;SMARCA2
		Olfactory bulb olfactory fiber layer	0.10032111	CNTNAP1;FAM135B;SMARCA2
Hyper-methylated in L-MSNs	Allen Brain Atlas Up-regulated	Main olfactory bulb, glomerular layer	0.10347247	COL22A1;KIAA1161
		Olfactory bulb glomerular layer	0.10347247	COL22A1;KIAA1161
		Superficial stratum of olfactory bulb	0.10347247	COL22A1;KIAA1161
		Olfactory bulb olfactory fiber layer	0.10347247	COL22A1;KIAA1161
	GO Molecular Process 2018	Sensory perception of chemical stimulus	0.00043704	OR8H2;OR5AS1;OR8U8
		Sensory perception of smell	0.00035449	OR8H2;OR5AS1;OR8U8
		Olfactory receptor activity	0.00025126	OR8H2;OR5AS1;OR8U8

First, the differentially methylated genes between hPSC-MSNs and Primary GEs were analysed. Regardless of cell line, an increased level of methylation was consistently observed at CpG sites in hPSC-MSNs when compared to Primary GEs (number of differentially methylated genes that were more methylated in L-MSNs: 2043; F1-MSNs: 2607; H9-MSNs 2807). Gene enrichment analysis implicated genes that were related to various brain regions and neuronal pathways (Table 5.5). Analysis using the Allen Brain Atlas Up database (which aims to group upregulated genes by neural tissues and regions across rodent, primate and human brains) identified over 400 significant terms for each comparison (456 for L-MSNs, 423 for F1-MSNs, and 467 for H9-MSNs), which included the terms *Striatum*, *Caudate Nucleus*, and *Putamen* (Table 5.6). Analysis using the ARCHS4 Tissues database (which collates publicly available RNA sequencing data from human and mouse tissues) revealed 52 significant terms in the L-MSN comparison, 49 in the F1-MSN comparison and 50 in the H9-MSN comparison. The terms demonstrating the strongest evidence for the enrichment of differentially methylation loci included various neural tissues, including fetal brain and the dorsal striatum (Table 5.5). The GO Biological Function (2018) database also returned several significant terms for each comparison (40 for L-MSNs, 34 for F1-MSNs, and 27 for H9-MSNs) which indicated genes related to related to neuronal development, maturation, and function were differentially methylated (Table 5.5). Conversely, when analysing the few genes associated with more methylated CpGs in Primary GE than hPSC-MSNs (L-MSNs: 4 genes; F1-MSNS: 19 genes; H9-MSNs: 80 genes), it was found that both F1-MSNs and H9-MSNs had unmethylated CpGs associated with non-neural functions. For example, analysis using the GO Biological Function (2018) database with these genes revealed significant associations with epidermal differentiation. L-MSNs did not return terms with more than one associated gene.

Comparisons between hPSC-MSN groups revealed some of the previously observed methylation differences between lines are associated with common pathways and tissues. First, H9-MSNs were compared to L-MSNs (131 differentially methylated genes that were more methylated in L-MSNs; and 294 more methylated in H9-MSNs) and F1-MSNs (102 genes more methylated in F1-MSNs; and 225 genes more methylated in H9-MSNs). Few significant associations were found between these genes, however the ARCHS4 Tissues database found genes more methylated in H9-MSNs were significantly associated with Brain (Bulk), the superior frontal gyrus, and the cerebral cortex (Table 5.6). Finally, L-MSNs were compared to F1-MSNs (7 genes more methylated in L-MSNs; and 13 genes more methylated in F1-MSNs). Only GO Molecular Process (2018) returned significant results, indicating that L-MSNs had increased levels of methylation at three genes (OR8H2, OR5AS1, and OR8U8) which

were significantly associated with olfactory receptor activity and sensory perception of smell and chemical stimulus (Table 5.7). However, the Allen Brain Atlas Up and Down databases highlighted five further genes related to olfactory tissues (COL22A1, KIAA1161, CNTNAP1, FAM135B, and SMARCA2) though these did not remain significant after correction for multiple testing. Examination of these databases indicated that the first three genes are not included in the Allen Brain Atlas libraries, and the other five genes are not included in the GO Molecular Process (2018) library. COL22A1 and KIAA1161 were more methylated in L-MSNs, and were upregulated in olfactory related tissues in the Allen Brain Atlas Up database, whereas CNTNAP1, FAM135B, and SMARCA2 were more methylated in F1-MSNs, and are downregulated in olfactory related tissues in the Allen Brain Atlas Down database.

5.2.6 Methylation at enriched striatal genes

Following the finding that hPSC-MSNs were differentially methylated at CpGs related to the striatum, putamen and caudate nucleus, the methylation status of routinely tested genes associated with striatal MSN phenotypes were individually examined. DMPs between hPSC-MSNs and Primary GEs were explored for CpGs associated with striatal genes. GSX2, DLX2, NKX2.1, DARPP32, DRD1, DRD2, and GABRA1, were all found to have significant differences between hPSC-MSNs and Primary GEs across at least one CpG in all groups, though none exhibited a Δβ-value detectable using the previously defined Δβ-value difference of >0.5. However, FOXP1, FOXP2, CTIP2, and PENK were also found to have at least one CpG that was significantly different between hPSC-MSNs and Primary GEs that was differentially methylated at a Δβ-value difference of >0.5. Examination of all CpGs related to these genes revealed a trend for hPSC-MSNs to exhibit more methylation than Primary GEs across these genes but generally a similar pattern is observed, indicating that classic MSN markers are similarly methylated between hPSC-MSNs and Primary GE tissues (Figure 5.4).

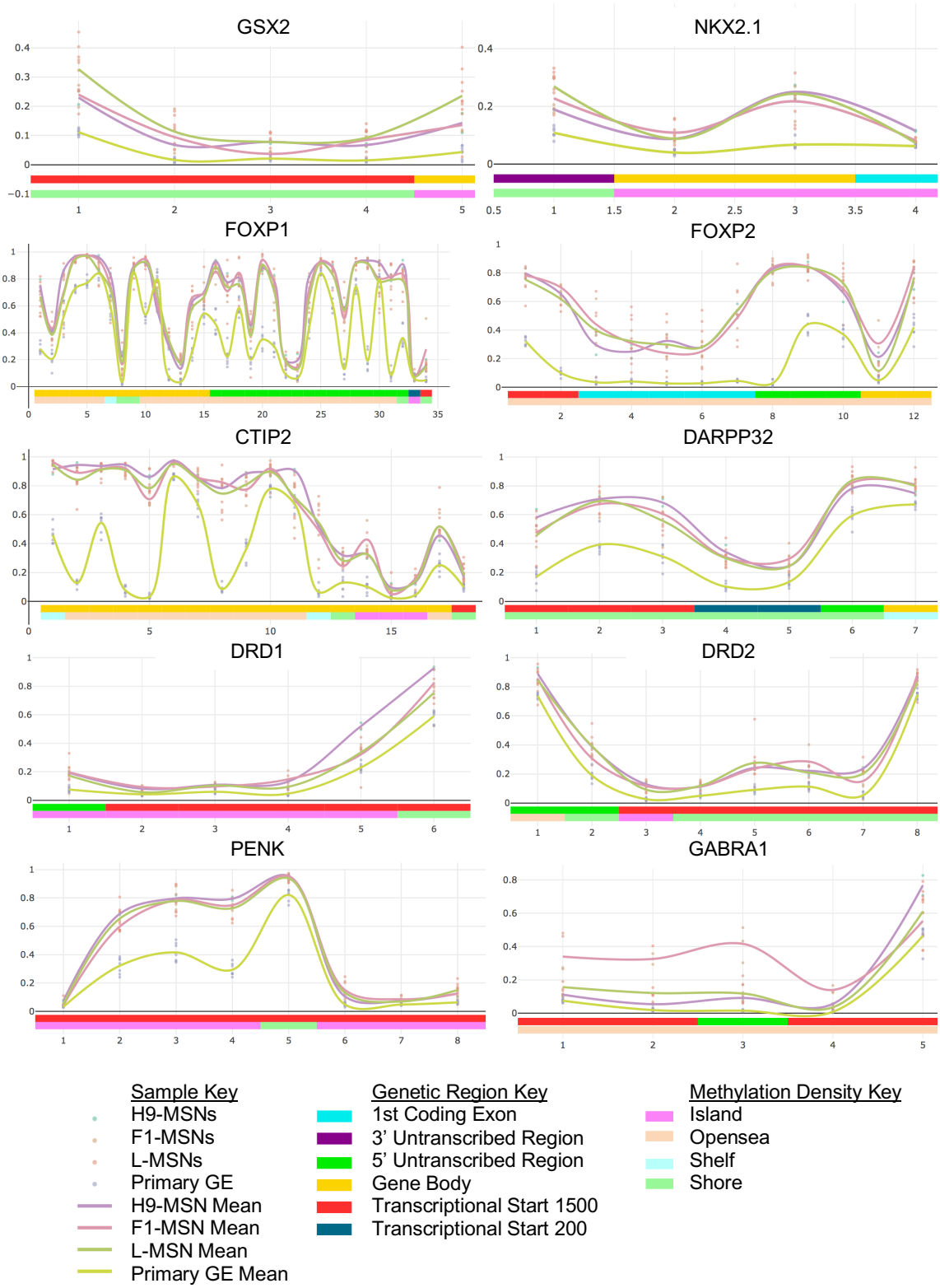


Figure 5.4 Methylation status of genes associated with a striatal MSN phenotype
The Y axis indicates the CpG β-value, and the X axis indicates approximate genetic regions (top bar, see Genetic Region Key), and local CpG density (bottom bar, see Methylation Density Key). Values for each sample are plotted for each CpG examined, and the mean value for each group is indicated by a continuous line across the gene. hPSC-MSNs exhibit similar methylation patterns as Primary GE samples at critical MSN related genes, however there is also an observable trend for increased methylation across these CpGs compared to Primary GE samples.

5.3 Discussion

In this chapter I have explored the methylomes of isogenic hiPSCs following a differentiation into an MSN phenotype to determine the extent and impact of their epigenetic memory of their tissue of origin. I have also compared these cells to Primary GE tissues, which represent the gold standard MSN progenitor tissues that hPSC-MSNs need to attempt to replace to be progressed for clinical application in CRT for HD. Epigenetic memory of tissue of origin is identifiable between isogenic hiPSC-MSNs, but various other features also influence their methylomes.

5.3.1 Genetically identical hiPSCs have a distinct epigenetic signature

During the initial exploration of these data it became apparent that there was a unique DNA methylation signature across 827 CpGs that persisted in all samples derived from SWIFT 2285 tissues. This signature was present in the primary LGE, MGE, cortex and fibroblast tissues indicating it was not reflective of tissue phenotype. It was also highly conserved as it was still present in hiPSCs derived from these tissues and all resulting hiPSC-MSNs. During the QC process all probes associated with SNPs or located on X and Y chromosomes were removed to reduce the genetic and sex differences between samples (Methods 2.5.2.1), however, previous studies examining methylation in the fetal brain have demonstrated that a small number of sex specific methylation differences also occur across autosomal chromosomes (521 CpGs: Spiers *et al.*, 2015). As the sex of SWIFT donors is not tested for or recorded, it was considered possible these differences could be sex based. However, when comparing these 827 CpGs to the 521 CpGs identified in Spiers *et al.*, 2015, there was no overlap between these two lists of CpGs, which suggests these differences are not caused by difference in the sex of the donors. Methylation is known to vary according to genome, and there is evidence that methylation in adult brain can vary according to genetic identity (Zhang, *et al.* 2010). Therefore it is possible this methylation signature is simply unique to SWIFT 2285, and that it arose during data exploration due to the high number of samples derived from this sample and therefore sharing this genetic heritage. Finally, to determine if there was a functional relationship between these differentially methylated CpGs, the genes associated with them (57 genes) were collated and underwent gene enrichment analysis. No tissue associations were observed, but two biological processes (regulation of cellular carbohydrate metabolic process; regulation of glucose metabolic process) were found to be related to three of the differentially methylated genes (SIK3, RORA, and PDK1).

5.3.2 The fetal methylome is highly reflective of tissue phenotype

Following the removal of the CpGs that were differentially methylated in SWIFT 2285 derived samples, exploratory analysis on the full dataset was undertaken. First, samples underwent unsupervised hierarchical clustering analysis which seeks to group samples (and subsequent groups formed of grouped samples) in order of their overall similarity. This was conducted using complete linkage clustering theory, by creating a distance matrix between samples based on the sum of all β-value differences between these samples at each CpG site tested. The most similar samples were then clustered together into a new cluster, and the distance between this cluster and all others was then calculated again to determine which two clusters were the next most similar, which was repeated until a final cluster was formed. This was then complemented with analysis of the 1000 most variable probes across samples, which can highlight more fundamental methylation similarities and differences between samples. Across these two forms of exploratory analysis it became apparent that the samples examined separated primarily by tissue/cellular phenotype, although differences between these two methods of data exploration also indicated that other factors could play a significant role in defining sample methylomes which are discussed below.

The primary fibroblast sample was the last to cluster in the hierarchical clustering analysis (Figure 5.2(1)) and had a distinct methylation profile compared to other samples when comparing the 1000 most variable CpGs (Figure 5.3A-B). As it was the sole sample included from a mesodermal cell lineage, this difference was presumably driven by tissue or germ layer specific methylation patterns which are known to be present in fetal tissues (Roost *et al*., 2017). However, it is yet to be established if the various regions of the fetal brain also exhibit unique methylation profiles. In this project, two neural structures were included (primary cortex and WGE) and two WGE substructures (LGE and MGE). Primary cortex clustered early on with primary LGE, primary MGE, and primary WGE, indicating the relative similarity between the methylation profiles of fetal neural tissues. However, this clustering did not occur at the same level as the LGE, MGE and WGE samples clustered, indicating that there were still some methylation differences that appear to distinguish fetal neural structures (Figure 5.2(8); Figure 5.3A,B). Furthermore, primary WGE samples clustered together before the inclusion of LGE and MGE samples, indicating potential variability between the LGE and MGE substructures. This is further supported by the discrepancies between the number of DMPs observed between hPSC-MSNs and primary WGE and LGE groups, which consistently demonstrated that hPSC-MSNs were more similar to primary LGE than primary WGE tissues. In light of this however, with zero DMPs observed between primary LGE and WGE tissues, it is not possible to conclude

definitively that there exists methylation differences between these two tissues. Collectively these results indicate that fetal tissues possess distinctive tissue methylation patterns associated with their tissue phenotype, and that fetal brain structures also appear to have unique structure-specific methylomes. There is also evidence suggesting fetal brain substructures may have unique methylation patterns, but this will need to be fully explored in future work.

Interestingly, there were distinctions made between primary and cultured WGE tissues, that separated them into distinct clusters during both the unsupervised hierarchical clustering (Figure 5.2(5)) and MDS clustering analysis (Figure 5.3A). This was revealed to be caused by an increase in methylation compared to primary tissue counterparts (Figure 5.3B) and resulted in significant changes at 13581 CpGs (Table 5.3; p=0.01). It is beyond the scope of this thesis to determine what is driving this methylation shift, but it highlights the possibility that culture techniques could be inducing this methylation shift in WGE, and could therefore play a factor in the differences observed between hPSC-MSNs and the primary fetal neural tissues.

5.3.3 hPSC-MSNs exhibit a distinctive methylome which is influenced by genome, hPSC type, and tissue of origin

The second unique cluster in the hierarchical clustering analysis was composed of all three H9-hESC derived samples including the H9-MSNs (Figure 5.2(2)). This indicates a distinction between these cells and primary neural tissues, that is potentially not shared by hiPSC-MSNs, as these were instead observed to first cluster with primary neural tissues (Figure 5.2(4,5,6)) before clustering with their pluripotent counterparts (Figure 5.2(3)). Of note, this difference is not in part due to the removal of SWIFT 2285 specific methylation, as unsupervised hierarchical clustering was unaffected by the removal of these CpGs. Later, H9-hESC derived samples separated by cell phenotype into H9-pluripotent and H9-MSN samples (Figure 5.2(7)) at around the same level of similarity that divisions were made between cultured WGE and many hiPSC-MSNs (Figure 5.2(8)). Collectively, this implies that there exists a methylation signature distinctive for these three H9-hESC derived samples that was sufficiently different from other samples to cause H9-ESC samples to cluster alone, and this could be specific to either the starting cell type (hESC) or the lines genetic identity (H9). Without additional hESC lines included in this analysis, it is not possible to determine which of these components drove this separation from other hPSC-MSNs. Examination of the top 1000 most variable CpGs across all samples did not highlight H9-hESC specific methylation patterns, but examination of the top 1000 most variable CpGs across just hPSC-MSN

samples did indicate a series of CpGs that were differentially methylated between H9-MSNs and hiPSC-MSNs (Figure 5.3D). Furthermore, DMP analysis between hPSC-MSNs indicated there were more DMPs between H9-MSNs and hiPSC-MSNs than between hiPSCs derived from different tissues (Table 5.4). It is important to recognise that the hiPSCs tested here are from the same SWIFT donor, and have been demonstrated in this analysis to share some genome specific methylation. However, the SWIFT 2285 specific significant DMPs were removed from this comparison, and therefore these differences could reflect a difference between MSNs cultured from hiPSCs and hESCs. Both of these analyses complement the separate clustering of H9-hESC derived samples observed in the unsupervised hierarchical clustering analysis, supporting the idea that H9-hESC derived samples have a unique hESC or genetic specific methylation signature, that separates them from primary neural tissues more than the hiPSCs examined here.

As previously mentioned, the hiPSC-MSNs derived from both primary LGE and fibroblasts clustered with fetal neural tissues (Figure 5.2(4,5,6)) before their pluripotent counter parts (Figure 5.2(3)). However, this occurred across multiple clustering points (4-6) indicating some lines were less similar to these cells than others. Three hiPSC-MSNs clustered later during unsupervised hierarchical clustering analysis than the majority (Figure 5.2(4)). This could indicate a high degree of variability in hiPSC-MSN populations, and, as all three samples that clustered later were all derived from LGE derived hiPSCs (lines L2 and L11) it is possible, though unconfirmed, that this is related to tissue of origin. When examining the heatmap expressing the 1000 most variable CpGs between hPSC-MSNs (Figure 5.3D) these same three samples were observed to have less uniform methylation than the other LGE derived hiPSC-MSNs. Collectively this suggests there is more variability between L-hiPSCs undergoing an MSN differentiation than in other lines, but without a more extensive dataset including other hiPSC lines it is difficult to determine this with confidence.

Whilst the global methylation analysis suggests that hESC-MSNs and hiPSC-MSNs are distinctive from each other, examination of the most variable probes reveals that there are key similarities between these cells. When clustering the samples based only on these highly variable CpGs, hPSC-MSNs form a loose cluster away from other samples (Figure 5.3A). However, this appears to be driven by a higher degree of methylation than is observed in primary tissues (Figure 5.3B), as when the primary tissues are removed from this comparison, hPSCs separate instead by cell type (hESC and hiPSC) and tissue of origin (LGE derived hiPSCs and fibroblast derived hiPSCs; Figure 5.3C). Collectively, this demonstrates that hPSC-MSNs share a dominant common methylation profile that is indicative of their cellular phenotype (*in vitro* cultured

MSN at day 30 of a standardised protocol), but secondary line specific differences also exist that persist through the differentiation process which could affect final cell phenotype. Examination of DMPs between these samples, when grouped by cell type and tissue of origin, further supports this hypothesis (Table 5.4), and there is some evidence from the gene enrichment analysis that these DMPs are related to different tissues and pathways. First, hiPSC-MSNs were observed to be less methylated at CpGs associated with genes significantly enriched for spinal cord, brain tissue, and brain regions, compared to hESC-MSNs (Table 5.7), which could indicate more open chromatin structure at these genes in hiPSC-MSNs than hESC-MSNs. When comparing hiPSC-MSNs by tissue of origin (LGE and fibroblast), there was some significant enrichment for olfactory related pathways (Table 5.8), and of the 20 genes examined in total, 8 of them were associated with olfactory pathways or tissues. Furthermore, two of these genes were more heavily methylated in F1-MSNs and these genes are downregulated in olfactory tissues, whereas the other seven were more heavily methylated in L-MSNs and are upregulated in olfactory tissues and pathways. Whilst these terms did not remain significant after correction, investigation of the gene enrichment libraries used revealed that neither library contained all eight of these genes. Additionally, the libraries do not account for the link between both up- and down-regulation and subsequently it is plausible that this lack of significance is a product of the limits of this comparison.

5.3.4 hPSC-MSNs do not share a methylome with Primary GE tissues

The greatest observed differences uncovered in this chapter were not between hPSC-MSNs, but rather between hPSC-MSNs and Primary GE. Outputs from the DMP analysis comparing the hPSC-MSNs and Primary GE revealed there were over 100,000 significant DMPs between these groups, representing significant differences across approximately 25% of the total CpGs examined in this study (Table 5.4). It should be noted that the number of DMPs was lowest in H9-MSNs (100,760 DMPs) and highest in L-MSNs (114,393 DMPs), which could indicate there are more differences between L-MSNs and Primary GE than other comparisons. However, it is likely the number of DMPs that reached significance is in part due to the limitations of this study, specifically the small group sizes. In a previous DMP analysis there was a discrepancy observed in the number of DMPs between similar tissue groups (Table 5.3; Primary LGE and Primary WGE), compared to groups from which they were dissimilar. This indicates a lack of power preventing the identification of differences between similar groups, and therefore similar samples were merged into larger groups to increase the power of this analysis at the potential loss of some specificity. After combining samples into larger

groups, the comparison of L-MSNs to Primary GE had the highest number of samples (n=15) and therefore more statistical power and a greater ability to find less variable DMPs. Conversely, the comparison of H9-MSNs to Primary GE had the lowest number of samples (n=8) and therefore has less power to identify less variable CpGs. Further evidence for this comes from the differentially methylated genes associated with these DMPs. To generate these lists, DMPs were filtered to only include DMPs with a Δβ-value difference between groups of at least 0.5, to indicate a methylation shift in at least 50% of the sample population. Despite having the largest total number of DMPs, L-MSNs had the fewest number of differentially methylated genes associated with DMPs (total 2047 genes) compared to the other groups (Fibro-MSNs: 2628 genes; ESC-MSNs: 2887 genes). This demonstrates that whilst L-MSNs varied significantly from Primary GE, there were fewer instances of Δβ-value differences surpassing 0.5, suggesting the additional DMPs found in the comparison between L-MSNs and Primary GE are likely a product of increased statistical power rather than any true increase in difference. However, this will need to be validated in future bodies of work that can specifically attempt to quantify these differences without losing specificity, and at least within this study it is hard to draw definitive meaning from the number of DMPs or differentially methylated genes. Importantly, that this underpowered study is able to consistently identify over 100,000 DMPs between hPSC-MSNs and Primary GE tissues strongly indicates that there are drastic differences between the authentic gold standard methylome, and that which exists in the hPSC-MSNs generated using a standardised protocol.

To explore the possible biological relevance of these differences in DMPs, gene enrichment analysis was conducted on the genes associated with significant DMPs between groups, divided by direction of methylation (i.e. increased methylation in the first group and decreased in the second, or decreased methylation in the first group and increased in the second). This analysis revealed that genes strongly associated with neuronal tissues and functions, including striatal tissues were more widely methylated in hPSC-MSNs compared to primary GE (Table 5.6). First, the Allen Brain Atlas databases revealed a significant association between these genes and over 400 terms describing neural regions and structures, including the striatum, putamen and caudate nucleus. Indeed, examination of the methylation status of genes commonly used to identify positive MSN populations in hPSC derived cultures confirmed a trend for hPSC-MSNs to exhibit increased, though not complete, methylation across several of these genes (Figure 5.4). The ARCHS4 tissue database confirmed that these genes were more strongly associated with neural tissues, and also fetal brain (Table 5.6). However, it also revealed some enrichment for terms related to tissues across other parts of the body (data not shown). Furthermore, these genes were found to have clear biological

associations each related to neurological function and development as seen by significant associations within the GO Biological Process (2018) database. The strongest association of these processes in all comparisons was nervous system development, which was associated with 100 differentially methylated genes in L-MSNs, 115 differentially methylated genes in F1-MSNs, and 124 differentially methylated genes in H9-MSNs. Conversely, when comparing differentially methylated genes that were more methylated in Primary GE than hPSC-MSNs, significant associations were observed related to biological processes unrelated to neural phenotypes (e.g. epidermal differentiation), but only in F1-MSNs and H9-MSNs. Interestingly, F1-MSNs were found to be uniquely less methylated at genes associated with Sensory Perception of Smell than Primary GE (Table 5.6). This was driven by one of the same genes differentially methylated between L-MSNs and F-MSNs and two additional associated genes (OR8H2, OR10AG1, OR2T34).

Methylation was first established as a repressor of gene transcription, and this is understood to be one of its primary purposes in mediating gene expression (Bird, 2002). Subsequently, the data gathered from this gene enrichment analysis could suggest hPSC-MSNs have a reduced capacity to transcribe genes associated with neural tissues and function, and equally F1-MSNs and H9-MSNs could have an increased capacity to transcribe genes related to epidermal differentiation. However, there is additional evidence that DNA methylation plays a more varied role in gene expression than previously thought, particularly in adult tissues, for instance depending on which gene and specific CpG binding site is methylated, methylation can result in alternative gene splicing instead of transcriptional repression (Lev Maor *et al.*, 2015) and increased tissue specific gene expression (Lokk *et al.*, 2014). Determining the actual biological consequences of these differentially methylated CpGs and genes will require significant further analysis to be conducted in future experiments. Regardless of actual function, DNA methylation is an important marker of variation in development and gene transcription, and the results from this gene enrichment analysis demonstrate hPSC-MSNs possess a widely different methylome to primary fetal striatal tissues.

5.3.5 Conclusions

In this chapter I aimed to explore the methylation profile of *in vitro* cultured MSNs derived from different hPSC lines, to establish if any variability in hiPSC-MSNs could be attributed to their epigenetic memory of their tissue of origin, and to compare their methylomes to authentic LGE tissues. Throughout the analysis it became apparent that the study was not sufficiently powerful to completely achieve these goals, and instead

several compromises have been made to enhance the comparisons which potentially reduced the specificity of the results. Nonetheless, the resulting data has expanded on our understanding of the primary fetal brain, and provided some of the first data observing how the methylome of the human fetal brain structures differ. Furthermore, there is some evidence that the various substructures of the fetal brain may also harbour variable DNA methylation patterns that distinguish them from neighbouring tissues, but a larger experiment is necessary to fully determine this. This work has established that *in vitro* cultured hPSC-MSNs are largely methylated across the genome, in direct contrast to Primary GE tissues. The functional consequences of this remain unknown, but this could be indicating that our current differentiation protocols are insufficient to induce an authentic maturing methylome reflective of authentic primary striatal tissues. I have also uncovered evidence suggesting that the methylome of these cells varies between differentiations, though it remains to be determined if this is specific to conditions fulfilled by LGE derived hiPSCs undergoing an MSN differentiation, or if this indeed a global variability between uniformed directed differentiations. Aside from these differentiation induced variations, hPSC-MSNs also exhibit methylation patterns that are consistent between MSN populations derived from the same cell line, which seem to be related to their tissue of origin, as hPSC-MSNs clustered into groups indicative of tissue of origin, and L-MSNs exhibited fewer total differentially methylated genes compared to Primary GE, compared to other hPSC-MSNs. However, due to the limitations of this study it is difficult to conclude this on this data alone. Finally, in spite of the relatively low power of this study, differentially methylated genes were uncoverable between groups, and were consistently associated with functional pathways between lines, emphasising the need for additional examination of the variability between these lines.

Chapter 6: General Discussion

CRT is a promising therapeutic option for a variety of neurodegenerative disorders including HD. Proof of principle for this intervention has been established by transplanting fetal WGE into the MSN deficient striatum to elicit functional recovery. As such, fetal WGE are the gold standard source of MSN progenitors currently available to us, yet they are inappropriate for widespread clinical use and other cell sources much be considered. hPSCs have the capacity to differentiate into any cell of the body, and therefore are a possible alternative source of MSN progenitors. However, to date, functional recovery using hPSC-derived MNSs has not been consistently demonstrated. One possible reason for this is that these cells are not authentic MSNs. There is evidence to suggest that hiPSCs can retain an epigenetic memory of their tissue of origin, which can enhance their capacity to differentiate towards similar cellular phenotypes to their tissue of origin. In this thesis, I have investigated the potential of hiPSCs derived from fetal LGE tissues to differentiate towards an MSN phenotype, and compared these to isogenic hiPSCs derived from different tissues. I have also explored the epigenome of these cells to establish how much of their epigenome is retained, and have compared them following exposure to an MSN differentiation protocol, to gold standard WGE tissues.

6.1 Discussion

6.1.1 hiPSCs derived from fetal tissues retain an epigenetic memory of their tissue of origin with functional consequences

The primary hypothesis underpinning the work presented here is that iPSCs retain an epigenetic memory of their tissue of origin, and that this will have consequences on the differentiation potential of these cells. This hypothesis is supported by a wider body of research that has consistently demonstrated an effect of epigenetic memory of tissue of origin in iPSCs, but has reported inconsistent results regarding the benefit of this epigenetic memory (Introduction 1.5.4). For instance, hiPSCs with an epigenetic memory of a hepatic tissue of origin have been studied multiple times to try to enhance the efficiency of hepatic differentiations, but consistently this has not provided any observable effect (Ohi, *et al.*, 2011; Liu *et* al., 2011; Heslop et al., 2017). In contrast, and relevant to this thesis, reports of hiPSCs with an epigenetic memory of a neural tissue of origin consistently report some benefit of the epigenetic memory to enhance neural differentiation (Tian *et al.*, 2011; Hargus *et al.*, 2014; Roost

et al., 2017). In this thesis, I have added to this literature by presenting evidence that hiPSCs derived from fetal LGE, MGE, cortex and fibroblast tissues do indeed appear to retain an epigenetic memory of their tissue of origin. Furthermore, I have found consistent differences between these cell lines, demonstrating that there are functional consequences of this epigenetic memory, that alter the differentiation capacity of these cells.

This was first evident in Chapter 3, when examining how the most variably methylated probes clustered pluripotent samples from cell lines L1, L2, L11, F1, and the hESC control line H9. Here, these cells formed three distinct clusters that aligned with their tissue of origin: the first contained all LGE derived hiPSCs (L1, L2, and L11); the second contained the fibroblast derived hiPSCs (F1); and the last contained the hESCs (H9). Later in Chapter 5 these differences were further corroborated, as when examining the hPSC-MSNs resulting from these lines under the same conditions, these samples also clustered in the same pattern. Interestingly, this indicated that these line specific signatures were maintained through the differentiation process, which suggests that the epigenetic variations potentially have long lasting consequences for these cell lines and their final cell phenotype.

It is unfortunate that, due to the cost of the genome wide methylation array, more hiPSC lines could not be examined to further develop this finding. For instance, in Chapter 3, more weight could have been added to the observed tissue of origin specific separation between lines F1 and L1, L2, and L11 by including additional hiPSCs lines derived from the primary fetal fibroblast cells used to generate F1. Additionally, the inclusion of lines M3 and C6 could have further helped our understanding of how specific this retained epigenetic signature is, by providing data on hiPSCs derived from tissues similar to LGE, but still developmentally distinct (MGE and cortex). There is some indirect data from the characterisation work conducted in Chapter 4 that suggests that these lines may have distinct epigenetic profiles that are specific to their tissue subtypes, most notably the trends of the C6 line that are not shared by the L2 and L11 lines. However, as these lines were not included in the DNA methylation analysis, there is no direct observation of their DNA methylation profiles, and it remains unknown whether the epigenetic memory of tissue of origin can be reliably broken down into fetal tissue neural subtypes at this stage of development. Furthermore, there is some work suggesting that the epigenetic variability between hiPSCs is more defined by genome and other variables than by tissue of origin (Choi *et al.*, 2015; Heslop *et al.*, 2017). Some evidence of this may have been observed in this thesis. First, there was a distinct methylation signature in all SWIFT 2285 derived samples which was consistent across tissues and could not be readily attributed to

other variables (e.g. sex differences; Spiers *et al.*, 2015). In this thesis, steps were taken to remove this genome specific epigenetic signature so as to avoid bias in the DNA methylation analysis. But, the effect of genome cannot be completely discounted from the work presented here, as the H9 derived samples (which are genomically distinct from the SWIFT 2285 hiPSCs studied here), continuously clustered further away from the hiPSC samples in all cluster analyses presented in Chapter 3 and Chapter 5. The unique signature of the H9 line could not be removed from this analysis, because it was not possible to directly attribute this distinction to genetic differences, as this may instead have been an effect of pluripotent cell type (hESCs vs hiPSCs). Previous work has indicated that whilst hESCs and hiPSCs do have highly similar methylomes indicative of their pluripotent nature, there are consistent differences between them that seem to be related to pluripotent cell type (Liu *et al.*, 2012; Choi *et al.*, 2015). Future work could be conducted to better understand the relevance of genetic identity on the methylome, and to establish whether this is more important to hiPSC lines than their epigenetic memory of tissue of origin. For example, the current dataset could be complimented with additional hiPSC lines derived from LGE, MGE, cortical and fibroblast tissues collected from new SWIFT fetal donors, generated using the same SeV methods described in Chapter 3. This would enable the comparison of the effect of epigenetic memory of tissue of origin on the methylome, against the effect of genetic variability on the methylome. Therefore, whilst an epigenetic memory of tissue of origin has been directly observed here, it is difficult to draw conclusions about how tissue specific this memory is, or whether genetic variability would overshadow it. Both of these are important considerations for future efforts that attempt to use the epigenetic memory of hiPSCs.

Within Chapter 5 there were further results supporting the conclusion that hiPSCs retained an epigenetic memory of tissue of origin, and that these may have functional consequences. Gene enrichment analysis can be used to establish if there is a relationship between a subset of genes, and this was conducted on differentially methylated genes found between Primary GE tissues and the three groups of hPSC-MSNs: L-MSNs, F1-MSNs and H9-MSNs. Interestingly, examination of the genes that were hypomethylated in hPSC-MSN groups and hypermethylated in Primary GE tissues demonstrated that F1-MSNs and H9-MSNs were less methylated than Primary GE tissues at CpG regions associated with biological processes that were relevant for epidermal differentiation. This suggests an inappropriate methylome for these genes in the F1-MSNs and H9-MSNs, as it was not consistent with the Primary GE tissues, which contain authentic MSN progenitors. Whilst there were genes differentially hypomethylated in L-MSNs compared to Primary GE, none shared function or tissue associations, which directly supports the hypothesis that LGE derived hiPSCs retain an

epigenetic memory of their tissue of origin. This is further bolstered by the finding that, when compared to Primary GE tissues, the L-MSNs had the least number of differentially methylated genes of the hPSC-MSNs. However, due to the low sample size this may require further validation in a larger dataset that examines this effect across multiple fetal donors and their derived hiPSC lines.

Furthermore, gene enrichment analysis conducted on differentially methylated genes between experimental hPSC-MSN groups (also divided by direction of methylation), indicated that L-MSNs and F1-MSNs were differentially methylated at 8 genes, all related to various aspects of olfactory tissues, function, and sensory perception of smell. Following correction for multiple comparisons these were not all significant, but it is possible that this is because this study was underpowered and the databases examined do not have complete gene libraries. There was evidence that this difference was related to tissue of origin, as gene enrichment analysis conducted on the genes associated with significant DMPs between F1-MSNs and Primary GE tissues indicated a similar significant enrichment for sensory perception of smell. This was due to differential methylation at the same genes that were found to be significantly differentially methylated between the L-MSNs and F1-MSNs, and this differential methylation was in the same direction (hypomethylated in F1-MSNs, and hypermethylated in L-MSNs and Primary GE), further supporting the epigenetic memory of tissue of origin of L-MSNs. The functional consequences of these findings are unknown, as the precise role of DNA methylation for these genes has not been established. However, it may be worthwhile to investigate these particular genes further as the developing LGE gives rise to cells that contribute to the olfactory bulb (Stenman *et al.*, 2003; Ma *et al.*, 2013), and therefore understanding the functional relevance of these genes in the context of MSN specification may provide further insight to the findings presented in Chapter 4. In particular, the finding that the F1 line arguably produced the most enriched MSN phenotype as demonstrated through its increased expression of NOLZ1, FOXP1, CTIP2, and DARPP32.

In Chapter 4, the epigenome of hPSC-MSNs was not directly examined, but consistent differences were observed between the isogenic hiPSC lines as they underwent a standardised MSN differentiation that had been optimised for a hESC line. First, there were differences in the capacity of these hiPSCs to undergo a standardised MSN differentiation protocol, which led to two adapted versions of the base protocol used in this thesis. Specifically, morphological changes in lines L1 and M3 appeared to indicate these lines had undergone a more rapid commitment to a neuronal fate during dual SMAD inhibition, a trait which was not shared by the F1 line. Later in the protocol, neural tissue derived lines exhibited reduced viability, particularly following a

freeze/thaw process, which again was not shared by the non-neural derived F1 line. Additional differences were next observed between these cells following transplantation into the QA lesioned rat striatum, where grafts derived from the C6 line exhibited a tendency to overgrow, and to undergo different migratory patterns compared to the grafted cells derived from the F1 line, and the hESC control line, H9. Throughout *in vitro* characterisation, differences were observed in gene expression that were consistent within lines across multiple differentiations. For example, both LGE derived hiPSC lines (L2 and L11) produced cells that highly expressed DARPP32 at day 30, and similar cells were not observed until day 40 in other hiPSC cultures derived from non-WGE tissues (C6 and F1) or a hESC control (H9). Of note, lines L1 and M3 were not included in this study, so it is undetermined if this would have also been observed in these lines. These experiments also identified traits unique to C6, including a significantly higher neuronal fate specificity, and trends for reduced DLX2 expression, and higher NKX2.1 expression than other cell lines. This may be indicative that the epigenetic memory of these hiPSCs is specific between neural tissues, but as previously discussed, without direct evidence from the DNA methylation study it is not possible to conclude this definitively. The last hiPSC line examined here, F1, consistently produced MSN cultures most enriched for a DARPP32 positive phenotype, and exhibited increased expression of critical upstream LGE markers CTIP2, NOLZ1, and FOXP1 compared to other lines. Collectively, the data from Chapter 4 discussed above indicates that these lines differ in consistent ways when undergoing the same differentiation protocol, and due to the methods used to derive these cells in Chapter 3, these differences are presumably related to the epigenetic variation between these lines directly observed in Chapter 3 and 5. Unexpectedly, this did not seem to confer an advantage on the LGE derived hiPSC lines, but instead may have inhibited large portions of these cultures to undergo an MSN differentiation. This calls into question the usefulness of the epigenetic memory of hiPSCs derived from LGE tissues in this context, as the evidence presented in this thesis suggests it may actually inhibit the MSN differentiation capacity of these cells when undergoing the protocols used here.

6.1.2 Further exploration is required to understand if the epigenetic memory of hiPSCs can be utilised to enhance an MSN differentiation

The original purpose in studying the effect of epigenetic memory in hiPSCs, was to establish whether it could be used to enhance hPSC derived MSN differentiations, and whether this would produce a more authentic MSN phenotype. Previously within our research group we have derived hiPSCs from fetal WGE tissues, that when transplanted *in vivo* produced a higher proportion of DARPP32 expressing cells than a

similarly cultured hESC line (Choompoo, 2015). However, these transplants exhibited signs of overgrowth which could not be discounted as a potential cause for this increased DARPP32 population. Furthermore, there were no controls for the effect of genome or hiPSC tissue of origin included in this initial study. Therefore, one aim of the work in this thesis was to attempt to replicate this finding with these controls in place. Unfortunately, this has not been achieved here as only about 25% of the total grafts survived in the *in vivo* experiment, including just 1 of the 14 grafts derived from MSN progenitors differentiated from LGE derived hiPSCs, which prevented the replication of our previous finding. However, differences have been observed between the graft groups that warrant further investigation, as precisely what caused these differences remains undetermined, although some of these findings are potentially consistent with examples in the wider literature. For example, in 2014 Hargus *et al*. reported the generation of hiPSCs from neural stem cells, and fibroblasts. They differentiated these hiPSCs towards a neural progenitor cell fate *in vitro* and conducted *in vitro* characterisation to determine that these cell populations expressed near identical levels of proliferation markers and neural markers. They then transplanted ≈150,000 of these progenitor populations separately into both the mouse striatum and cortex and examined these grafts following 10 weeks of incubation. They did not observe overgrowth, but did observe increased cell survival within both cortical and striatal grafts derived from neural stem cell derived hiPSCs (≈60,000 cells) compared to fibroblast derived hiPSCs (≈2,000 cells). They attributed this increased cell number to an enriched neuronal phenotype in hiPSCs derived from neural stem cells, that increased their capacity to survive within a neural environment. In this thesis, the fibroblast derived line F1 did not produce overgrowing grafts, whereas the cortical derived line, C6, consistently produced grafts that exhibited overgrowth. By the end of the study C6 grafts had the largest estimated graft volume, and the largest estimated number of cells. Whilst this was not the desired outcome, it is possible that a similar effect occurred here that allowed these cells to better survive in a neural environment and facilitated the survival and continued proliferation of these cells. However, it is worth noting that LGE derived line L2 also produced a viable graft, which despite being of neural origin, did not overgrow. Due to the nature of this work, it is hard to establish exactly what caused this discrepancy, however it is interesting this exponential growth was specific to a cortical derived hiPSCs, and there exists the unexplored possibility that this trait is related to its tissue of origin (cortex) which undergoes considerable growth compared to the human LGE from which L2 is derived. Understanding what caused C6, but not other genetically identical lines to overgrow could be particularly useful to the field of CRT, as overgrowth is a safety issue that must be addressed before transition to clinic (Keene, *et al.*, 2009). The later *in vitro* characterisation of these lines as they underwent

an MSN differentiation could hint at a possible opening to such work, as there was some evidence that C6 resisted commitment to a maturing MSN phenotype during the differentiation.

Immunocytochemical analysis indicated that C6 produced the purest neuronal population of the cells examined here. Yet, qPCR analysis of C6 showed that it expressed significantly lower levels of DLX2 than H9 cells at day 20 of the differentiation protocol and exhibited a trend for increased NKX2.1 expression. DLX2 is an early subpallium marker that is understood to be upregulated by TGFβ activity, which in turn in the protocols used in this thesis is regulated by the addition of Activin (Feijen, Goumans and van den Eijnden-van Raaij, 1994; Maira *et al.*, 2010). Conversely, NKX2.1 is also an early subpallium marker, but it is differentially expressed between LGE and MGE tissues, specifically at higher levels in the more ventral MGE (Noakes *et al.*, 2019). Collectively, this data suggest C6 is highly capable of differentiating towards a neuronal phenotype when directed towards a forebrain progenitor fate using dual SMAD inhibition, but that these cells are also less susceptible to Activin treatment as an LGE fate specifier. If this is the case, then this may be because their epigenetic memory of tissue of origin clashes with this part of the differentiation process, which is plausible considering that many of the genes present in the developing forebrain are expressed in a gradient and antagonistic manner. If so, then this could have been an underlying cause of the overgrowth observed in C6 derived grafts, that may have acted in tangent with an increased capacity to survive in a neural environment as observed by Hargus *et al.*, 2014.

C6 was not the only line that seemed to respond differently than expected to the MSN differentiation protocol used here. As mentioned previously, the LGE derived lines L1, L2, and L11 were theorised to have the highest efficiency to produce MSNs over other lines, yet the final MSN differentiation efficiency of these cells was frequently less than both H9 and F1 lines, as measured by the specific markers used here. This is surprising considering the evidence outlined from the DNA methylation data above, which indicated that these cells had an epigenetic memory more similar to primary GE tissues than the other hPSC-MSNs. Additionally, there were trends observed earlier in the differentiation process indicating that these lines exhibited an accelerated capacity for neuronal commitment, and both L2 and L11 developed cells strongly immuno-positive for DARPP32 before non-LGE derived hiPSC lines (C6 and F1) and the hESC line (H9). This conflicting data suggests that there may still be untapped potential in the LGE derived hiPSCs, and it is possible that under the right conditions this could be harnessed properly to enhance an MSN differentiation.

Throughout the experiments conducted in Chapter 4, attempts were made to reduce the variability that could occur due to factors unrelated to the epigenome of these cells, and several of these pertained to the methods used to differentiate these lines towards an MSN fate. Primary amongst these, was the decision to conduct a standardised differentiation and to divide this process into two stages with a freezing step between progenitor fate specification and final MSN maturation. The decision to use a standardised protocol was made because it is challenging to directly compare multiple cell lines undergoing different differentiation protocols, as it is difficult to determine what results are due to differences between cell lines, and what results are due to differences between the protocols. Similarly, the decision to freeze MSN progenitors was made to enable simultaneous culturing of the examined MSNs, to ensure identical environmental conditions, and that variability in sample collection, RNA extraction, cell fixing, immunocytochemical staining, etc, would be minimised.

Previous studies conducted within our lab and in collaborator labs, had found no difficulties with the inclusion of a freeze/thaw cycle on progenitor MSNs derived using a similar differentiation protocol on multiple hESC lines. Indeed, the data presented in this thesis is in agreement with this, as there were no difficulties induced by a freeze/thaw cycle observed in the hESC line used here, and the results I generated from these lines were similar to previously published results (Arber *et al.*, 2015). Yet, hiPSC lines derived from neural tissue sources (L1, L2, L11, M3 and C6) did show consistent viability issues with the inclusion of this freeze/thaw cycle at the same time point, and this, in combination with other data outlined above, suggests that these cells were not following the same developmental time course as the hESC line or even the genomically identical F1 line. As such, it must also be considered that the use of a standardised differentiation may have negatively affected the work presented here, as this could have resulted in inappropriate timing of exposure to critical developmental factors. For example, the timing of Activin exposure to neural precursor cells is understood to play a critical role in final fate specification. In hESCs derived neural precursors generated using a 2D dual SMAD inhibition protocol, exposure to Activin from day 9 reliably induces an MSN phenotype (Arber *et al.*, 2015), but if this exposure is delayed by 8 days, Activin exposure can instead induce an interneuron phenotype more associated with the caudal tail of the GE (Cambray *et al*., 2012). Whilst not examined here, it is possible the accelerated rate of differentiation observed in some lines resulted in fewer MSN lke cells and more caudal GE interneurons, which would in turn prevent the accurate assessment of the true MSN differentiation potential of these lines. This is in agreement with previously published work; both Tian *et al*., 2011 and Roost *et al.*, 2017, published evidence that iPSCs derived from neural tissues (respectively, mouse astrocytes, and human fetal whole brain) differentiated to a βIII-

tubulin expressing population faster than hiPSCs derived from fibroblast cells. Future work should aim to rectify this by closely examining the temporal developmental changes these cell populations undergo, in order to develop optimal protocols for each line being examined. It may also be necessary to adjust the concentrations of factors, as the concentrations used here were also originally determined for a hESC line.

However, this alone would not overcome the issue of comparing lines undergoing different differentiation protocols. Previously, it has been reported that hESCs that undergo dual SMAD inhibition and receive no further fate specification do not produce MSN progenitors (Arber *et al.*, 2015: less than 1%). Therefore, if these hiPSCs have indeed retained an epigenetic memory of their tissue of origin that inherently directs them towards an MSN fate, an alternative experiment could be conducted to compare the MSN differentiation efficiency of these cells to hESCs under sub-optimal conditions for hESCs. This could be achieved by following a similar differentiation protocol to those described here, but by removing Activin from the protocol, or limiting its influence to only 1-2 days. If under these sub-optimal conditions the hiPSCs derived from LGE tissues are more capable of differentiating towards an LGE precursor or MSN fate, then this would be an indication that their epigenetic memory of tissue of origin may indeed have the capacity to produce either more MSNs or potentially a more authentic phenotype.

Finally, future work should also attempt to address other flaws in the design of this project that may have impeded on the potential for this work to test my hypothesis, with focus on three issues in particular: (1) the inherent heterogeneity of the tissues from which these iPSCs were derived; (2) the potential clonal variability between lines was not fully accounted for in the main experimental design, and (3) the set-up of the animal study conducted in Chapter 4.

With regards to the first of these, the developing LGE, MGE and cortex are each composed of a diverse range of both neuronal and glial subtypes, several of which are simply migrating through these structures towards their final site of maturation and thus did not originate within the structure. In this work, hiPSC lines were derived from plated cultures of these tissues that had undergone no further cell sorting following initial dissection. In the case of the LGE cultures, where it was desirable to derived hiPSCs from LGE progeny destined to become MSNs, it is possible that several of these lines could have instead been obtained from off-target cells, such as LGE progeny committing to non-striatal fates such as the olfactory bulb (Stenman *et al.*, 2003; Ma *et al.*, 2013), the migrating interneurons of the MGE, or supporting glial cells which are known to originate in the MGE, the LGE and the cortex in staggered stages of development (Kessaris *et al*., 2006). In an attempt to overcome this issue in this work I

aimed to use multiple clones derived from the LGE tissue, to increase the likelihood that at least one of these lines would have been derived from an MSN progenitor cell. However, it is unclear whether any differences observed between LGE derived hiPSC lines were the result of differences in the original cells they were derived from, or other sources of variation (such as the induction process or differences caused by culture techniques). Furthermore, whilst the LGE derived hiPSC lines were found to be consistently quite similar across most comparisons made here, this alone does not necessarily mean they were all derived from MSN progenitor cells or even a shared phenotypic origin. It is also worth noting that this precaution was not undertaken for hiPSCs derived from other tissues, and future work could also attempt to expand the number of hiPSC clones for other tissues of interest. In future work, to properly select for desirable cells for each tissue (such as LGE progeny committed to an MSN fate) cell sorting methods such as FACS could have been implemented to isolate a desirable cell type before infection with the Sendai virus. Previously within our lab group, we have struggled to maintain viability of human fetal tissues following FACS sorting, but future work could attempt to optimise such techniques, and ensure isolation of only MSN progenitors (by selecting for cells expressing known MSN precursor genes such as CTIP2, or potentially earlier markers such as GSX2 which must be down regulated before gliogenesis can occur within the LGE; Chapman *et al.*, 2013, Chapman *et al.*, 2018).

Otherwise, PSCs are understood to be a highly volatile cell population over long term culture, as they are prone to genetic mutations, duplications, and deletions which results in clonal variability (Laurent *et al.*, 2011, Garitaonandia *et al.*, 2015). These subtle changes are unlikely to cause significant variability in short term culture, but slight competitive advantages in subclones of a pluripotent population will eventually amplify any genetic or epigenetic variations acquired through culture (Brenière-Letuffe *et al.*, 2018). These alone have been shown to drive significant differences in the differentiation potential of a PSC line, and as such may have influenced the results found here. It is generally challenging to monitor and prevent these issues in PSC culture, as they are unpredictable and varied. For this reason, proper management requires regular and in-depth analysis which dramatically increases PSC culture time and costs. An alternative but less vigorous approach that can be used it to restrict the number of passages PSCs undergo, as risks of clonal variability is increased with the number of passages a PSC culture has undergone (Laurent *et al.*, 2011, Garitaonandia *et al.*, 2015, Brenière-Letuffe *et al.*, 2018). In the work prevented here, long term expansion was avoided and only low passage PSCs were used. Furthermore, all data used to examine the differentiation potential of these lines was generated from a restricted number of passages (only between passages 12-18 for hiPSCs and 18-23

for hESCs). As such, any clonal variability that occurred during culture should have minimal effect on the data presented in this thesis. However, this would not overcome any significant genomic or epigenomic events which altered the examined lines prior to this period. As such, future work could consider the inclusion of genome wide SNP and CNV assessment at critical points to determine similarity between samples from within a line, and critically, from between lines to ensure their isogenic nature is maintained.

Lastly, the *in vivo* work presented in Chapter 4 could be further optimised to streamline results in future work, which is especially pertinent as the data, whilst interesting, are difficult to interpret with confidence due to low graft survival. This was most notable for LGE derived hiPSC lines, where only one graft survived out of fourteen. This may be partially due to the issues of viability found when thawing MSN progenitors from neural derived hiPSC lines prior to transplantation. If so, it could have potentially been resolved here by avoiding the freezing step, as unlike the frozen samples, the subset of cells which were simply passaged and maintained in culture from day 18 did not exhibit a reduction in viability. However, as evidenced by the equally low viability in the F1 derived grafts (three surviving grafts out of ten) and hESC derived grafts (four surviving grafts out of ten), this alone would not overcome the issue of low graft survival. It is possible that this low viability was brought about by the reliance on an immune suppressant model, as this requires daily injections to be made accurately, which if not correctly maintained is likely to result in graft rejection. Within our lab group we have previously overcome this issue by developing a tolerization model, which aims to prevent an immune response to the engrafted cells by exposing neonatal rodent pups to the human cells they will be transplanted with during the development of their immune system (Kelly *et al.*, 2009; Roberton *et al.*, 2015). However, this method still requires further optimisation as presently it can still lead to graft rejection if the initial tolerization is not successful (Roberton *et al.*, 2015). Alternatively, immune deficient rats could have potentially been used instead, as they cannot initiate an immune response to grafted cells. However, such animal models are challenging to use for behavioural testing, as they are highly prone to infection whenever exposed to unclean environments, which increases with repeated handling. Future experiments could attempt to enhance graft survival by utilising one of these alternative models, choosing whichever is more appropriate to the desired outcome measures.

6.1.3 hPSC-MSNs exhibited increased methylation at genes associated with striatal phenotypes compared to primary GE tissues

One of the main findings presented in this work is the large discrepancy in the DNA methylation profiles of hPSC-MSNs and Primary GE tissues. It is well established that PSCs have a unique methylome indicative of, and required to maintain, their pluripotent phenotype (Bibikova *et al.*, 2006; Cedar and Bergman, 2012; Arand, *et al.,* 2015; Liao *et al.*, 2015). This is characterised by a typically higher instance of methylation across the genome, particularly at promotors of tissue specific genes, and at this stage DNA methylation is primarily used as a gene silencer for such genes (Bibikova *et al.*, 2006; Cedar and Bergman, 2012). As cells undergo developmental processes, the epigenome shifts, thereby preventing the transcription of genes essential for pluripotency including OCT4 and NANOG, and instead allowing the expression of genes required for the developing cellular phenotype, globally becoming comparatively less methylated (Mohn, *et al.*, 2008; Nazor *et al.*, 2012; Cedar and Bergman, 2012). There is a body of evidence that indicates DNA methylation is still tightly regulated throughout fetal brain development, but the precise role of the developing methylome in neural development is still not completely understood (Numata *et al.*, 2012, Lister *et al.*, 2013; Spiers *et al.*, 2015; Roost *et al.*, 2017). By the time a cell has terminally differentiated into an adult cell, the role of DNA methylation has become broader. The majority of the genome is unmethylated, and the remaining methylated DNA signature occurs in a highly tissue specific manner (Nagae, *et al.*, 2011; Fernandez *et al.*, 2012; Lokk *et al.*, 2014), but depending on locus and direction of methylation, it can be involved in gene silencing, gene splicing, or even gene promotion (Lev Maor *et al.*, 2015). It is not well understood when the role of the methylome changes from strictly a gene silencing mechanism to this more dynamic role, but it is likely a gradual process that occurs throughout development as cells terminally mature into their adult cell phenotype.

In Chapter 5, I have presented evidence that hPSC-MSNs are methylated in a manner distinctly different from the developing primary tissues they are supposed to be replicating. Specifically, gene enrichment analysis revealed that the primary GE tissues were less methylated at CpG regions associated with genes that are relevant to striatal tissues, various brain tissues, and important biological processes related to neural development and function, all of which are presumably necessary for the authentic functionality displayed by grafted WGE tissues. Without full characterisation of the role of hyper- and hypomethylation of these regions in the fetal brain across the time course of striatal development, it is not possible to establish the exact biological consequences of this observed difference. For example, it remains unknown whether the CpGs

associated with these genes need to be differentially methylated between progenitor and mature MSNs. What is clear though, is that DNA methylation is tightly regulated in fetal brain development (as evidenced by the high consistency between GE samples), and several striatal genes were consistently hypomethylated in the primary fetal GE tissues, and hypermethylated in hPSC-MSNs. If hPSC-MSNs are supposed to mimic primary fetal GE tissues, then their methylome, a significant indicator of a cell populations transcriptional capacity, does not appear to be accurately reflecting the gold standard source of MSN progenitors. It is therefore of primary interest to the CRT field to understand precisely why this difference exists between these populations and what the consequences are, as if it is undesirable then it may have to be rectified before hPSCs are ready for CRT for HD. There are two possible reasons for this discrepancy: it could be due to a difference in the developmental maturity of these cells; or it could be that the current protocols used *in vitro* are insufficient to produce an authentic MSN methylome in hPSCs.

It is presently challenging to determine when PSC derived cultures are developmentally equivalent to primary tissues, due to our limited knowledge of how their development and maturation varies between these respective environments. For instance, of the cells examined here, the hPSC-MSNs were differentiated for 30 days, which happens to be the time point that these cells first begin expressing DARPP32 (as evidenced in Chapter 4), additionally the primary fetal tissues examined here were between 59-70 days post conception (Methods, Table 2.2), which is also around the same time that DARPP32 is first identifiable in the developing fetal brain (Onorati *et al.*, 2014 reported DARPP32 in human fetal WGE from day 56). For this reason, the two populations examined in this thesis could be considered developmentally equivalent. Yet, it is unlikely that the hPSC-MSNs were still malleable enough to successfully graft into the QA lesioned rat striatum, as they exhibited similar gene expression and morphology to cells that have failed to graft before (Aubry *et al.*, 2008). In contrast, there is published evidence that primary WGE can still be successfully grafted into the QA lesioned rat striatum and elicit functional recovery up to ≈90 days post conception, a full 20 days of gestation beyond the most developmentally mature WGE sample examined in this thesis (Pundt *et al.*, 1996a). This difference seems to instead suggest that these hPSC-MSNs are more developmentally mature and that their methylome is therefore a reflection of maturity, an argument which is supported by the fact that the WGE is a heterogeneous population that maintains proliferative regions throughout development, and therefore contains a cell population with a wide variety of differently developmental mature cells. However, whilst it is true that there is likely a slight increase in methylation following terminal maturation in the human brain (Numata *et al.* 2012), it is unlikely that this is why the hPSC-MSNs examined here exhibited a more methylated

genome. As evidenced by the various heatmaps in Chapter 5 that first demonstrated this trend, much of the methylated regions in hPSC-MSNs that are hypomethylated in primary GE tissues, are also methylated in the pluripotent populations included in this analysis. Therefore, the data presented in this thesis most likely indicates that these hPSC-MSNs have not yet lost the global gene silencing methylation exhibited in pluripotent cells, and instead only the methylation related to specific genes sought in typical MSN characterisation work (e.g. the genes examined in Chapter 4) have been altered, as emphasised by the relatively similar methylation status of commonly targeted MSN markers (observed in Chapter 5 section 5.2.6). If this is the case, then it strongly suggests that a wide plethora of genetic code is either unavailable to these cells, or expressed in an inauthentic manner, which could hinder authentic MSN development and function.

A more comprehensive DNA methylation study of primary fetal LGE/WGE tissues at different developmental stages, and of hPSC-MSNs as they undergo the various stages of MSN differentiation and maturation, could help further determine which of these possibilities is the underlying cause. Additionally, such work could be further enhanced if it was complimented with RNA sequencing of these samples, to further advance our understanding of the biological consequences of this differential methylation and how the transcriptome of these cells is impacted by their methylome. However, for both of these technologies the heterogeneity of these populations may prevent their detailed characterisation. Recently, advances in single cell sequencing have advanced the possibility of conducting both methylation and RNA analysis at the single cell level, which could provide a far more detailed and accurate picture of the underlying dynamics of these cell populations (Hwang, Lee and Bang, 2018; Karemaker and Vermeulen, 2018). Such comparisons would help build our understanding of the differences observed in this thesis between primary fetal WGE and hPSC-MSNs, and this could be used to identify and explored additional gene targets and pathways to refine our current ability to differentiate MSNs *in vitro*. Furthermore, with the advent of CRISPR mediated technology, it has now become possible to alter the methylome directly, which could be conducted at critical sites to further shape authentic development once the developing methylome of fetal tissues is better understood (Rienecker, Hill and Isles, 2016; Liu *et al.*, 2016). Finally, it is worth mentioning that DNA methylation is not the only epigenetic mechanisms known to play a role in development and the regulation of gene expression. For example, other epigenetic mechanisms that regulate gene expression include histone acetylation, nucleosome positioning, non-coding RNAs, and some self-promoting RNA transcripts. Future work should also account for these mechanisms where possible, to try and develop an understanding of the whole epigenome of authentic fetal tissues.

6.1.4 Some variability in DNA methylation between samples may be due to an effect of culture

Further evidence that the current protocols may be insufficient to produce an authentic MSN methylome comes from the comparison of Primary and Cultured WGE tissues in this thesis. The WGE samples examined here were collected from 3 primary WGE tissues. Half of each donation was preserved as a primary tissue sample, whilst the other half was plated for 4 weeks of spontaneous differentiation *in vitro*. Therefore, one sample in each group is genetically matched with another in the second group. During DNA methylation analysis, it became apparent that the cultured WGE samples had drifted from their original methylome, as detected in the genetically identical Primary WGE samples, and acquired additional methylation at various CpGs, leading to significant differences between the methylome of these groups.

There are several possible causes for this methylation shift. First, it is possible this is an effect of maturation, as the Cultured WGE has had an additional 4 weeks to develop *in vitro*. As previously discussed, it is challenging to determine if this is an effect of maturity because the developing methylome remains relatively understudied. Alternatively, it is possible that the cultured WGE has been able to produce more immature cells through proliferation. At the time of dissection, the WGEs examined here would have still been highly proliferative, as demonstrated by the continued expression of the proliferation marker Ki-67 in similarly aged WGE tissues (Onorati *et al.*, 2014). Recent unpublished work conducted in our lab has found that WGE tissues (undergoing the same culture methods used to spontaneously differentiate these samples) do exhibit a more mature neuronal phenotype. However, a proportion of cells in the cultures also continue proliferating, and there is provisional evidence suggesting that these proliferating cells are more “generic” and are less capable of producing an MSN phenotype. Importantly, there is some evidence that long term expansion of primary fetal WGE results in an altered transcriptome within these cells, and critically, reduces the ability of these cells to develop into grafts (Zietlow *et al.*, 2005; Anderson *et al.*, 2007; Zietlow, *et al.*, 2012). Whilst not studied here, it is possible the observed methylation shift in my data is reflective of this previously published change in gene expression after prolonged culture, which coincides with the provisional data mentioned above in WGE that has been cultured in the same manner as the samples examine here. As the DNA methylation data collected in this thesis is an expression of the methylation status of the whole sample population, an increase in the proportion of methylated cells in a population would theoretically cause a global shift in the output for that population, similar to what has been observed here. It is therefore reasonable to

assume that this is what has occurred in these samples. However, it is unclear why the progenitors expanding in Cultured WGE are more methylated. Is it simply that the most methylated cells are the most proliferative? Or is it instead because these cells have acquired *de novo* methylation across their genome, due to prolonged exposure to the *in vitro* culture environment? The exact cause of this methylation drift remains undetermined here, as it is beyond the scope of this thesis. However, it highlights an important possible contributing factor to the differences between primary GE tissues and the hPSC-MSNs previously outlined above: as an effect of *de novo* methylation during *in vitro* culture.

It is therefore important that this also be determined by future research, as the answer to this question may have important consequences for the future of hPSC differentiations. Specifically, if this shift is not a desirable feature of development and is instead acquired *de novo* due to the culture environment, then until this is resolved it may not be possible to culture authentic MSN progenitors from hPSCs. Furthermore, it may also provide important insight into the previously encountered issues with expanding primary WGE tissues, and open up new avenues to explore the expansion of authentic MSN progenitor tissues. Determining whether this is due to *de novo* acquisition of methylation could be studied by expanding the analysis of DNA methylation to include primary WGE tissues across various developmental time points to first determine if this shift is a feature of typical development, and by comparing this to additional WGE samples, cultured as they have been here, or in the presence of Dnmt3a/b inhibitors to prevent *de novo* methylation (Okano *et al.*, 1998, 1999; Hsieh 1999). This would allow the direct examination of the authentic methylome as it undergoes typical development, and the selective prevention of *de novo* methylation, but not of an increased methylation profile through Dnmt1 regulated progenitor cell division (Bestor 1992; Yoder *et al.*, 1996; Pradhan *et al.*, 1999).

6.1.5 Final Conclusions: are hiPSCs with a "desirable" epigenetic memory of tissue of origin a worthwhile investment for CRT for HD?

The primary goal of this thesis was to establish whether epigenetic differences could be detected between various pluripotent cells lines and their MSN differentiated progeny, and to subsequently consider whether any retained epigenetic memory of tissue of origin in hiPSCs derived from primary LGE tissues could be used to improve the production of hPSC derived MSNs. The experiments detailed in this thesis have demonstrated that there are indeed subtle differences between the cell lines examined here and that hiPSCs derived from primary fetal cells do appear to retain an epigenetic

memory of their tissue of origin. Furthermore, there is some evidence that this epigenetic memory may confer an advantage on LGE derived hiPSCs, as evidenced by the observation that both L2 and L11 hiPSCs generated highly expressing DARPP32 positive cells before other cell lines (Chapter 4 section 4.2.6), and the lack of significant terms in the gene enrichment analysis conducted on hypermethylated genes in the GE and hypomethylated genes in the L-MSNs, for which both other PSCs assessed produced potentially aberrant results (Chapter 5 section 5.2.5 and Table 5.5). However, my experiments have also produced some unexpected findings, such as a greater overall expression of key MSN genes (e.g. DARPP32) in the F1 hiPSC line which were instead derived from fibroblasts. Although this could be perceived as evidence in direct conflict with my hypothesis, it is important to recognise that I also found evidence that the differentiation protocol used in this work was suboptimal for some of these lines. As such, further work is required to properly establish if the aforementioned potential advantage of LGE derived hiPSCs can be used to produce more authentic MSNs for CRT in HD (and by extrapolation in other neurodegenerative diseases). This future work should endeavour to expand on the present work by examining the potential for these lines under optimised differentiation conditions, to examine the resulting MSNs at the single cell level using multi-omnic approaches, and to expand the comparison with authentic human LGE progeny.

This may yet prove especially pertinent, as an unexpected result from this work is the finding that there is a considerable discrepancy in the methylomes between the MSNs we can create in culture, and the current “gold standard” fetal tissues that have been used to demonstrate proof of principle for CRT for HD. It is important to clarify that the data presented in this thesis does not directly indicate that hPSC-MSNs are lacking compared to primary fetal striatal tissues, as the functional consequences of this difference have not been explored here, and further research is required to fully understand the nature of this observation. However, considering the potential implications of this finding for CRT in HD, and other neurodegenerative conditions, building on our understanding of the developing methylome and its role in hPSC-MSNs seems to be an important direction for future research. As such, the findings in this thesis highlight a need to continue studying the normal development of the fetal striatum with particular emphasis on wider aspects of development that are frequently overlooked, including the various mechanisms that underpin the role of the developing epigenome.

Bibliography

Albin, R. L., Young, A. B., and Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neuroscience*, 12, 366–375.

Allegrucci, C., & Young, L. E. (2007). Differences between human embryonic stem cell lines. *Human Reproduction Update*, *13*, 103–120.

Anchan, R., Gerami-Naini, B., Lindsey, J. S., Ho, J. W. K., Kiezun, A., Lipskind, S., … Williams, Z. (2015). Efficient Differentiation of Steroidogenic and Germ-Like Cells from Epigenetically-Related iPSCs Derived from Ovarian Granulosa Cells. *PLOS ONE*, *10*, e0119275.

Anderson, K. D., & Reiner, A. (1991). Immunohistochemical localization of DARPP-32 in striatal projection neurons and striatal interneurons: implications for the localization of D1-like dopamine receptors on different types of striatal neurons. *Brain Research*, *568*, 235–243.

Anderson, L., Burnstein, R. M., He, X., Luce, R., Furlong, R., Foltynie, T., … Caldwell, M. A. (2007). Gene expression changes in long term expanded human neural progenitor cells passaged by chopping lead to loss of neurogenic potential in vivo. *Experimental Neurology*, *204*, 512–524.

Anderson, R. M., Lawrence, A. R., Stottmann, R. W., Bachiller, D., & Klingensmith, J. (2002). Chordin and noggin promote organizing centers of forebrain development in the mouse. *Development (Cambridge, England)*, *129*, 4975–4987.

Andrew, S. E., Paul Goldberg, Y., Kremer, B., Telenius, H., Theilmann, J., Adam, S., … Hayden, M. R. (1993). The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nature Genetics*, *4*, 398–403.

Antequera, F., & Bird, A. (1993). Number of CpG islands and genes in human and mouse. *Proceedings of the National Academy of Sciences of the United States of America*, *90*, 11995–11999.

Arand, J., Wossidlo, M., Lepikhov, K., Peat, J. R., Reik, W., & Walter, J. (2015). Selective impairment of methylation maintenance is the major cause of DNA methylation reprogramming in the early embryo. *Epigenetics & Chromatin*, *8*, 1.

Arber, C., Precious, S. V., Cambray, S., Risner-Janiczek, J. R., Kelly, C., Noakes, Z., … Li, M. (2015). Activin A directs striatal projection neuron differentiation of human pluripotent stem cells. *Development*, *142*, 1375–1386.

Arlotta, P., Molyneaux, B. J., Jabaudon, D., Yoshida, Y., & Macklis, J. D. (2008). Ctip2 Controls the Differentiation of Medium Spiny Neurons and the Establishment of the Cellular Architecture of the Striatum. *The Journal of Neuroscience*, *28*, 622 LP-632.

Aryee, M. J., Jaffe, A. E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A. P., Hansen, K. D., & Irizarry, R. A. (2014). Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*, *30*, 1363–1369.

Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W., and O'reilly, R. C. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nature Neuroscience*, 10, 126–131.

Aubry, L., Bugi, A., Lefort, N., Rousseau, F., Peschanski, M., & Perrier, A. L. (2008). Striatal progenitors derived from human ES cells mature into DARPP32 neurons in vitro and in quinolinic acid-lesioned rats. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 16707–16712.

Bachoud-Lévi, A.-C., Gaura, V., Brugières, P., Lefaucheur, J.-P., Boissé, M.-F., Maison, P., … Peschanski, M. (2006). Effect of fetal neural transplants in patients with Huntington's disease 6 years after surgery: a long-term follow-up study. *The Lancet Neurology*, *5*, 303–309.

Bachoud-Lévi, A.-C., Rémy, P., Nǵuyen, J.-P., Brugières, P., Lefaucheur, J.-P., Bourdet, C., … Peschanski, M. (2000). Motor and cognitive improvements in patients with Huntington's disease after neural transplantation. *The Lancet*, *356*, 1975–1979.

Bagó, J. R., Alfonso-Pecchio, A., Okolie, O., Dumitru, R., Rinkenbaugh, A., Baldwin, A. S., … Hingtgen, S. D. (2016). Therapeutically engineered induced neural stem cells are tumour-homing and inhibit progression of glioblastoma. *Nature Communications*, *7*, 10593.

Bally-Cuif, L., & Boncinelli, E. (1997). Transcription factors and head formation in vertebrates. *BioEssays*, *19*, 127–135.

Ban, H., Nishishita, N., Fusaki, N., Tabata, T., Saeki, K., Shikamura, M., … Nishikawa, S.-I. (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*, 14234–14239.

Bar-Nur, O., Russ, H. A., Efrat, S., & 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*, 17–23.

Barker, R. A., Mason, S. L., Harrower, T. P., Swain, R. A., Ho, A. K., Sahakian, B. J., … NEST-UK collaboration, the N.-U. (2013). The long-term safety and efficacy of bilateral transplantation of human fetal striatal tissue in patients with mild to moderate Huntington's disease. *Journal of Neurology, Neurosurgery, and Psychiatry*, *84*, 657–665.

Bates, G., Tabrizi, S., & Jones, L. (Eds.). (2014). *Huntington's disease* (No. 64). Oxford University Press, USA.

Beal, M. F., Kowall, N. W., Ellison, D. W., Mazurek, M. F., Swartz, K. J., & Martin, J. B. (1986). Replication of the neurochemical characteristics of Huntington's disease by quinolinic acid. *Nature*, *321*, 168–171.

Beddington, R. S., & Robertson, E. J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. *Development*, *105*, 733 LP-737.

Ben-David, U., Gan, Q.-F., Golan-Lev, T., Arora, P., Yanuka, O., Oren, Y. S., … Benvenisty, N. (2013). Selective Elimination of Human Pluripotent Stem Cells by an Oleate Synthesis Inhibitor Discovered in a High-Throughput Screen. *Cell Stem Cell*, *12*, 167–179.

Benraiss, A., & Goldman, S. A. (2011). Cellular therapy and induced neuronal replacement for Huntington's disease. *Neurotherapeutics : The Journal of the American Society for Experimental NeuroTherapeutics*, *8*, 577–590.

Bestor, T. H. (1992). Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. *The EMBO Journal*, *11*, 2611–2617.

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., … Shen, R. (2011). High density DNA methylation array with single CpG site resolution. *Genomics*, *98*, 288–295.

Bibikova, M., Chudin, E., Wu, B., Zhou, L., Garcia, E. W., Liu, Y., … Fan, J.-B. (2006). Human embryonic stem cells have a unique epigenetic signature. *Genome Research*, *16*, 1075–1083.

Bilic, J., & Izpisua Belmonte, J. C. (2012). Concise review: Induced pluripotent stem cells versus embryonic stem cells: close enough or yet too far apart? *Stem Cells (Dayton, Ohio)*, *30*, 33–41.

Bird, A. (2007). Perceptions of epigenetics. *Nature 2007 447:7143*.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development*, *16*, 6–21.

Bird, A., Taggart, M., Frommer, M., Miller, O. J., & Macleod, D. (1985). A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell*, *40*, 91–99.

Björklund, A. and Stenevi, U. (1979) 'Reconstruction of the nigrostriatal dopamine pathway by intracerebral nigral transplants', *Brain Research*, 177, 555–560.

Björklund A., Dunnett S.B., Stenevi U., Lewis M.E., Iversen S.D. (1980) Reinnervation of the denervated striatum by substantia nigra transplants: functional consequences as revealed by pharmacological and sensorimotor testing. *Brain Research*; 199, 307-33.

Boland, M. J., Nazor, K. L., & Loring, J. F. (2014). Epigenetic regulation of pluripotency and differentiation. *Circulation Research*, *115*, 311–324.

Boland, M. J., Hazen, J. L., Nazor, K. L., Rodriguez, A. R., Gifford, W., Martin, G., … Baldwin, K. K. (2009). Adult mice generated from induced pluripotent stem cells. *Nature*, *461*, 91–94.

Boronat-García, A., Guerra-Crespo, M., & Drucker-Colín, R. (2017). Historical perspective of cell transplantation in Parkinson's disease. *World Journal of Transplantation*, *7*, 179–192.

Bradley, A., Evans, M., Kaufman, M. H., & Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. *Nature*, *309*, 255–256.

Brazel, C. Y., Romanko, M. J., Rothstein, R. P., & Levison, S. W. (2003). Roles of the mammalian subventricular zone in brain development. *Progress in Neurobiology*, *69*, 49–69.

Brenière-Letuffe, D., Domke-Shibamiya, A., Hansen, A., Eschenhagen, T., Fehse, B., Riecken, K. and Stenzig, J. (2018). Clonal dynamics studied in cultured induced pluripotent stem cells reveal major growth imbalances within a few weeks. *Stem cell research & therapy*, 9 (1), 165.

Brix, J., Zhou, Y., & Luo, Y. (2015). The Epigenetic Reprogramming Roadmap in Generation of iPSCs from Somatic Cells. *Journal of Genetics and Genomics*, *42*, 661–670.

Bueno, C., Sardina, J. L., Di Stefano, B., Romero-Moya, D., Muñoz-López, A., Ariza, L., … Menendez, P. (2016). Reprogramming human B cells into induced pluripotent stem cells and its enhancement by C/EBPα. *Leukemia*, *30*, 674–682.

Byrne, J. A. (2013). Overcoming clinical hurdles for autologous pluripotent stem cell-based therapies. *OA Stem Cells*, 1, 3.

Cattaneo, E., Zuccato, C., & Tartari, M. (2005). Normal huntingtin function: an alternative approach to Huntington's disease. Nature Reviews Neuroscience, 6 (12), 919-930.

Cambray, S., Arber, C., Little, G., Dougalis, A. G., De Paola, V., Ungless, M. A., Li, M., & Rodríguez, T. A. (2012). Activin induces cortical interneuron identity and differentiation in embryonic stem cell-derived telencephalic neural precursors. *Nature communications*, 3, 841.

Cedar, H., & Bergman, Y. (2012). Programming of DNA Methylation Patterns. *Annual Review of Biochemistry*, *81*, 97–117.

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

Chapman, H., Waclaw, R.R., Pei, Z., Nakafuku, M. and Campbell, K. (2013). The homeobox gene Gsx2 controls the timing of oligodendroglial fate specification in mouse lateral ganglionic eminence progenitors. *Development*, 140(11), 2289-2298.

Chapman, H., Riesenberg, A., Ehrman, L.A., Kohli, V., Nardini, D., Nakafuku, M., Campbell, K. and Waclaw, R.R. (2018). Gsx transcription factors control neuronal versus glial specification in ventricular zone progenitors of the mouse lateral ganglionic eminence. *Developmental biology*, 442(1), 115-126.

Chen, E. Y., Tan, C. M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G., … Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*, *14*, 128.

Chen, Y., Pu, J., & Zhang, B. (2016). Progress and challenges of cell replacement therapy for neurodegenerative diseases based on direct neural reprogramming. *Human Gene Therapy*, hum.2016.078.

Chin, M. H., Mason, M. J., Xie, W., Volinia, S., Singer, M., Peterson, C., … Lowry, W. E. (2009). Induced Pluripotent Stem Cells and Embryonic Stem Cells Are Distinguished by Gene Expression Signatures. *Cell Stem Cell*, *5*, 111–123.

Chiu, A. Y., & Rao, M. S. (2011). Cell-based therapy for neural disorders--anticipating challenges. *Neurotherapeutics : The Journal of the American Society for Experimental NeuroTherapeutics*, *8*, 744–752.

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

Cholewa-Waclaw, J., Bird, A., von Schimmelmann, M., Schaefer, A., Yu, H., Song, H., … Tsai, L.-H. (2016). The Role of Epigenetic Mechanisms in the Regulation of Gene Expression in the Nervous System. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *36*, 11427–11434.

Cisbani, G., & Cicchetti, F. (2014). Review: The fate of cell grafts for the treatment of Huntington's disease: the *post-mortem* evidence. *Neuropathology and Applied Neurobiology*, *40*, 71–90.

Clelland, C. D., Barker, R. A., & Watts, C. (2008). Cell therapy in Huntington disease. *Neurosurgical Focus*, *24*, E9.

Craufurd D, Snowden J. (2002) Neuropsychological and neuropsychiatric aspects of Huntington's disease. In: Bates G, Harper P, Jones L, eds. Huntington's disease. New York: Oxford University Press: 62–94.

Crittenden, J. R. and Graybiel, A. M. (2011) Basal ganglia disorders associated with imbalances in the striatal striosome and matrix compartments. *Frontiers in Neuroanatomy*, 5, 59.

Danesin, C., & Soula, C. (2017). Moving the Shh source over time: what impact on neural cell diversification in the developing spinal cord?. *Journal of developmental biology*, 5(2), 4.

Danjo, T., Eiraku, M., Muguruma, K., Watanabe, K., Kawada, M., Yanagawa, Y., … Sasai, Y. (2011). Subregional Specification of Embryonic Stem Cell-Derived Ventral Telencephalic Tissues by Timed and Combinatory Treatment with Extrinsic Signals. *Journal of Neuroscience*, *31*, 1919–1933.

Darnell, D., & Gilbert, S. F. (2017). Neuroembryology. *Wiley Interdisciplinary Reviews. Developmental Biology*, *6*. https://doi.org/10.1002/wdev.215

Deacon, T. W., Pakzaban, P., & Isacson, O. (1994). The lateral ganglionic eminence is the origin of cells committed to striatal phenotypes: neural transplantation and developmental evidence. *Brain Research*, *668*, 211–219.

Delli Carri, A., Onorati, M., Lelos, M. J., Castiglioni, V., Faedo, A., Menon, R., … Cattaneo, E. (2013). Developmentally coordinated extrinsic signals drive human pluripotent stem cell differentiation toward authentic DARPP-32+ medium-sized spiny neurons. *Development*, *140*, 301 LP-312.

Delong, M. R. (1990). Primate models of movement disorders of basal ganglia origin. *Trends in Neuroscience*, 13, 281–285.

Deng, Y. ., Albin, R. ., Penney, J. ., Young, A. ., Anderson, K. ., & Reiner, A. (2004). Differential loss of striatal projection systems in Huntington's disease: a quantitative immunohistochemical study. *Journal of Chemical Neuroanatomy*, *27*, 143–164.

Deng, J., Shoemaker, R., Xie, B., Gore, A., LeProust, E. M., Antosiewicz-Bourget, J., … Zhang, K. (2009). Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nature Biotechnology*, *27*, 353–360.

Devanna, P., Middelbeek, J., & Vernes, S. C. (2014). FOXP2 drives neuronal differentiation by interacting with retinoic acid signaling pathways. *Frontiers in Cellular Neuroscience*, *8*, 305.

DiFiglia, M., Sapp, E., Chase, K., Schwarz, C., Meloni, A., Young, C., Martin, E., Vonsattle, J. P., Carraway, R., Reeves S., Boyce, F. R., & Aronin, N. (1995). Huntingtin is a cytoplasmic protein associated with vesicles in human and rat brain neurons. *Neuron,* 14, 1075–1081.

Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M. J., Irizarry, R., … Feinberg, A. P. (2009). Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genetics*, *41*, 1350–1353.

Drews, K., Jozefczuk, J., Prigione, A., & Adjaye, J. (2012). Human induced pluripotent stem cells—from mechanisms to clinical applications. *Journal of Molecular Medicine*, *90*, 735–745.

Duke, C. G., Kennedy, A. J., Gavin, C. F., Day, J. J., & Sweatt, J. D. (2017). Experience-dependent epigenomic reorganization in the hippocampus. *Learning & Memory (Cold Spring Harbor, N.Y.)*, *24*, 278–288.

Dunnett, S. B., & Björklund, A. (1994). Mechanisms of function of neural grafts in the injured brain. *Functional neural transplantation*, *2*, 157-195.

Dunnett, S. B., & White, A. (2006). Striatal grafts alleviate bilateral striatal lesion deficits in operant delayed alternation in the rat. *Experimental Neurology*, *199*, 479–489.

Durston, A. J., Timmermans, J. P. M., Hage, W. J., Hendriks, H. F. J., de Vries, N. J., Heideveld, M., & Nieuwkoop, P. D. (1989). Retinoic acid causes an anteroposterior transformation in the developing central nervous system. *Nature*, *340*, 140–144.

Duval, K., Grover, H., Han, L.-H., Mou, Y., Pegoraro, A. F., Fredberg, J., & Chen, Z. (2017). Modeling Physiological Events in 2D vs. 3D Cell Culture. *Physiology (Bethesda, Md.)*, *32*, 266–277.

Duyao, M. P., Auerbach, A. B., Ryan, A., Persichetti, F., Barnes, G. T., McNeil, S. M., … Cattaneo, E. (1995). Inactivation of the mouse Huntington's disease gene homolog Hdh. *Science (New York, N.Y.)*, *269*, 407–410.

Eblen, F., and Graybiel, A. M. (1995). Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. *Journal of Neuroscience*, 15, 5999–6013.

Eddy, S. R. (2013). The ENCODE project: missteps overshadowing a success. *Current Biology : CB*, *23*, R259-61.

Ehrlich, M., & Lacey, M. (2013). DNA methylation and differentiation: silencing, upregulation and modulation of gene expression. *Epigenomics*, *5*, 553–568.

Eiraku, M., Takata, N., Ishibashi, H., Kawada, M., Sakakura, E., Okuda, S., … Sasai, Y. (2011). Self-organizing optic-cup morphogenesis in three-dimensional culture. *Nature*, *472*, 51–56.

Evans, A. E., Kelly, C. M., Precious, S. V., & Rosser, A. E. (2012). Molecular Regulation of Striatal Development: A Review. *Anatomy Research International*, *2012*, 1–14.

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

Faedo, A., Laporta, A., Segnali, A., Galimberti, M., Besusso, D., Cesana, E., … Cattaneo, E. (2017). Differentiation of human telencephalic progenitor cells into MSNs by inducible expression of Gsx2 and Ebf1. *Proceedings of the National Academy of Sciences of the United States of America*, *114*, E1234–E1242.

Fagiolini, M., Jensen, C. L., & Champagne, F. A. (2009). Epigenetic influences on brain development and plasticity. *Current Opinion in Neurobiology*, *19*, 207–212.

Fernández-Santiago, R., & Ezquerra, M. (2016). Epigenetic Research of Neurodegenerative Disorders Using Patient iPSC-Based Models. *Stem Cells International*, *2016*, 9464591.

Fernandez, A. F., Assenov, Y., Martin-Subero, J. I., Balint, B., Siebert, R., Taniguchi, H., … Esteller, M. (2012). A DNA methylation fingerprint of 1628 human samples. *Genome Research*, *22*, 407–419.

Ferré, S., Fredholm, B. B., Morelli, M., Popoli, P., & Fuxe, K. (1997). Adenosine-dopamine receptor-receptor interactions as an integrative mechanism in the basal ganglia. *Trends in Neurosciences*, *20*, 482–487.

Ferré, S., Lluís, C., Justinova, Z., Quiroz, C., Orru, M., Navarro, G., … Goldberg, S. R. (2010). Adenosine-cannabinoid receptor interactions. Implications for striatal function. *British Journal of Pharmacology*, *160*, 443–453.

Fjodorova, M., Noakes, Z., & Li, M. (2015). How to make striatal projection neurons. *Neurogenesis*, *2*, e1100227.

Fraga, M. F., Ballestar, E., Paz, M. F., Ropero, S., Setien, F., Ballestar, M. L., … Esteller, M. (2005). From The Cover: Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences*, *102*, 10604–10609.

Freeman, T. B., Cicchetti, F., Bachoud-Lévi, A. C., & Dunnett, S. B. (2011). Technical factors that influence neural transplant safety in Huntington's disease. *Experimental Neurology*, *227*, 1–9.

Freeman, T. B., Cicchetti, F., Hauser, R. A., Deacon, T. W., Li, X. J., Hersch, S. M., … Isacson, O. (2000). Transplanted fetal striatum in Huntington's disease: phenotypic development and lack of pathology. *Proceedings of the National Academy of Sciences of the United States of America*, *97*, 13877–13882.

French, C. A., Vinueza Veloz, M. F., Zhou, K., Peter, S., Fisher, S. E., Costa, R. M., & De Zeeuw, C. I. (2019). Differential effects of Foxp2 disruption in distinct motor circuits. *Molecular Psychiatry*, *24*, 447–462.

Fricker, R. ., Torres, E. ., Hume, S. ., Myers, R., Opacka-Juffrey, J., Ashworth, S., … Dunnett, S. . (1997). The effects of donor stage on the survival and function of embryonic striatal grafts in the adult rat brain: II. Correlation between positron emission tomography and reaching behaviour. *Neuroscience*, *79*, 711–721.

Fujiyama, F., Sohn, J., Nakano, T., Furuta, T., Nakamura, K. C., Matsuda, W., and Kaneko, T. (2011). Exclusive and common targets of neostriatofugal projections of rat striosome neurons: a single neuron-tracing study using a viral vector. *European Journal of Neuroscience*, 33, 668–677.

Furtado, S., Sossi, V., Hauser, R. A., Samii, A., Schulzer, M., Murphy, C. B., … Stoessl, A. J. (2005). Positron emission tomography after fetal transplantation in Huntington's disease. *Annals of Neurology*, *58*, 331–337.

Gallina, P., Paganini, M., Di Rita, A., Lombardini, L., Moretti, M., Vannelli, G. B., & Di Lorenzo, N. (2008). Human Fetal Striatal Transplantation in Huntington’s Disease: A Refinement of the Stereotactic Procedure. *Stereotactic and Functional Neurosurgery*, *86*, 308–313.

Gallina, P., Paganini, M., Lombardini, L., Mascalchi, M., Porfirio, B., Gadda, D., … Di Lorenzo, N. (2010). Human striatal neuroblasts develop and build a striatal-like structure into the brain of Huntington's disease patients after transplantation. *Experimental Neurology*, *222*, 30–41.

Gallina, P., Paganini, M., Biggeri, A., Marini, M., Romoli, A., Sarchielli, E., … Vannelli, G. B. (2014). Human Striatum Remodelling after Neurotransplantation in Huntington's Disease. *Stereotactic and Functional Neurosurgery*, *92*, 211–217.

Gao, S., Hou, X., Jiang, Y., Xu, Z., Cai, T., Chen, J., & Chang, G. (2017). Integrated analysis of hematopoietic differentiation outcomes and molecular characterization reveals unbiased differentiation capacity and minor transcriptional memory in HPC/HSC-iPSCs. *Stem Cell Research & Therapy*, *8*, 13.

Garitaonandia, I., Amir, H., Boscolo, F.S., Wambua, G.K., Schultheisz, H.L., Sabatini, K., Morey, R., Waltz, S., Wang, Y.C., Tran, H. and Leonardo, T.R. (2015). Increased risk of genetic and epigenetic instability in human embryonic stem cells associated with specific culture conditions. *PloS one*, 10 (2), e0118307.

Gerfen, C. R. (1984). The neostriatal mosaic: compartmentalization of corticostriatal input and striatonigral output systems. *Nature*, 311, 461–464.

Gerfen, C. R., Baimbridge, K. G., & Miller, J. J. (1985). The neostriatal mosaic: compartmental distribution of calcium-binding protein and parvalbumin in the basal ganglia of the rat and monkey. *Proceedings of the National Academy of Sciences of the United States of America*, *82*, 8780–8784.

Gerfen, C. R. (1989). The neostriatal mosaic: striatal patch-matrix organization is related to cortical lamination. *Science*, 246, 385–388.

Ghosh, Z., Wilson, K. D., Wu, Y., Hu, S., Quertermous, T., & 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*, e8975.

Grasbon-Frodl, E. M., Nakao, N., Lindvall, O., & Brundin, P. (1996). Phenotypic development of the human embryonic striatal primordium: a study of cultured and grafted neurons from the lateral and medial ganglionic eminences. *Neuroscience*, *73*, 171–183.

Grasbon-Frodl, E. M., Nakao, N., Lindvall, O., & Brundin, P. (1997). Developmental Features of Human Striatal Tissue Transplanted in a Rat Model of Huntington's Disease. *Neurobiology of Disease*, *3*, 299–311.

Graur, D., Zheng, Y., Price, N., Azevedo, R. B. R., Zufall, R. A., & Elhaik, E. (2013). On the Immortality of Television Sets: "Function" in the Human Genome According to the Evolution-Free Gospel of ENCODE. *Genome Biology and Evolution*, *5*, 578–590.

Grealish, S., Drouin-Ouellet, J., & Parmar, M. (2016). Brain repair and reprogramming: the route to clinical translation. *Journal of Internal Medicine*, *280*, 265–275.

Guillemin, G. J. (2012). Quinolinic acid, the inescapable neurotoxin. *FEBS Journal*, *279*, 1356–1365.

Gunhaga, L., Marklund, M., Sjödal, M., Hsieh, J.-C., Jessell, T. M., & Edlund, T. (2003). Specification of dorsal telencephalic character by sequential Wnt and FGF signaling. *Nature Neuroscience*, *6*, 701–707.

Guo, L., Karoubi, G., Duchesneau, P., Shutova, M. V., Sung, H.-K., Tonge, P., … Waddell, T. K. (2017). Generation of Induced Progenitor-like Cells from Mature Epithelial Cells Using Interrupted Reprogramming. *Stem Cell Reports*, *9*, 1780–1795.

Gupta, R., Nagarajan, A., & Wajapeyee, N. (2010). Advances in genome-wide DNA methylation analysis. *BioTechniques*, *49*, iii–xi.

Gusella, J. F., Wexler, N. S., Conneally, P. M., Naylor, S. L., Anderson, M. A., Tanzi, R. E., ... & Young, A. B. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. Nature, 306 (5940), 234-238.

Halder, R., Hennion, M., Vidal, R. O., Shomroni, O., Rahman, R.-U., Rajput, A., … Bonn, S. (2016). DNA methylation changes in plasticity genes accompany the formation and maintenance of memory. *Nature Neuroscience*, *19*, 102–110.

Halliday, G., McRitchie, D., Macdonald, V., Double, K., Trent, R. and McCusker, E. (1998) Regional Specificity of Brain Atrophy in Huntington's Disease, *Experimental Neurology*, 154, 663–672.

Hargus, G., Ehrlich, M., Araúzo-Bravo, M. J., Hemmer, K., Hallmann, A.-L., Reinhardt, P., … Zaehres, H. (2014). Origin-dependent neural cell identities in differentiated human iPSCs in vitro and after transplantation into the mouse brain. *Cell Reports*, *8*, 1697–1703.

Hauser, R. A., Furtado, S., Cimino, C. R., Delgado, H., Eichler, S., Schwartz, S., … Freeman, T. B. (2002). Bilateral human fetal striatal transplantation in Huntington's disease. *Neurology*, *58*, 687–695.

Heard, E., Clerc, P., & Avner, P. (1997). X-CHROMOSOME INACTIVATION IN MAMMALS. *Annual Review of Genetics*, *31*, 571–610.

Henikoff, S., & Greally, J. M. (2016). Epigenetics, cellular memory and gene regulation. *Current Biology*, *26*, R644–R648.

Heslop, J. A., Kia, R., Pridgeon, C. S., Sison-Young, R. L., Liloglou, T., Elmasry, M., … Park, B. K. (2017). Donor-Dependent and Other Nondefined Factors Have Greater Influence on the Hepatic Phenotype Than the Starting Cell Type in Induced Pluripotent Stem Cell Derived Hepatocyte-Like Cells. *Stem Cells Translational Medicine*, *6*, 1321–1331.

Hiler, D., Chen, X., Hazen, J., Kupriyanov, S., Carroll, P. A., Qu, C., … Dyer, M. A. (2015). Quantification of Retinogenesis in 3D Cultures Reveals Epigenetic Memory and Higher Efficiency in iPSCs Derived from Rod Photoreceptors. *Cell Stem Cell*, *17*, 101–115.

Hindley, C., & Philpott, A. (2013). The cell cycle and pluripotency. *Biochem. J*, *451*, 135–143.

Hon, G. C., Rajagopal, N., Shen, Y., McCleary, D. F., Yue, F., Dang, M. D., & Ren, B. (2013). Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nature Genetics*, *45*, 1198–1206.

Hsieh, C. L. (1999). In vivo activity of murine de novo methyltransferases, Dnmt3a and Dnmt3b. *Molecular and Cellular Biology*, *19*, 8211–8218.

Hu, B.-Y., Weick, J. P., Yu, J., Ma, L.-X., Zhang, X.-Q., Thomson, J. A., & Zhang, S.-C. (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*, 4335–4340.

Hu, Q., Friedrich, A. M., Johnson, L. V., & Clegg, D. O. (2010). Memory in Induced Pluripotent Stem Cells: Reprogrammed Human Retinal-Pigmented Epithelial Cells Show Tendency for Spontaneous Redifferentiation. *STEM CELLS*, *28*, 1981–1991.

Huang, H., & Kim, K. (n.d.). *Unsupervised Clustering Analysis of Gene Expression*. Retrieved from http://www.stat.berkeley.edu/~hhuang/STAT141/CLustering-analysis.pdf

Huang, K., Shen, Y., Xue, Z., Bibikova, M., April, C., Liu, Z., … Fan, G. (2014). A Panel of CpG Methylation Sites Distinguishes Human Embryonic Stem Cells and Induced Pluripotent Stem Cells. *Stem Cell Reports*, *2*, 36–43.

Hughes, C. S., Postovit, L. M., & Lajoie, G. A. (2010). Matrigel: A complex protein mixture required for optimal growth of cell culture. *PROTEOMICS*, *10*, 1886–1890.

Hwang, B., Lee, J. H., & Bang, D. (2018). Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental & Molecular Medicine*, *50*, 96.

Isacson, O. (2009). Cell Replacement Therapy for Huntington's Disease. In *Encyclopedia of Neuroscience* (pp. 639–642).

Ishiy, F. A. A., Fanganiello, R. D., Griesi-Oliveira, K., Suzuki, A. M., Kobayashi, G. S., Morales, A. G., … Passos-Bueno, M. R. (2015). Improvement of In Vitro Osteogenic Potential through Differentiation of Induced Pluripotent Stem Cells from Human Exfoliated Dental Tissue towards Mesenchymal-Like Stem Cells. *Stem Cells International*, *2015*, 249098.

Ivkovic, S., & Ehrlich, M. E. (1999). Expression of the striatal DARPP-32/ARPP-21 phenotype in GABAergic neurons requires neurotrophins in vivo and in vitro. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *19*, 5409–5419.

Iyer, L. M., Zhang, D., & Aravind, L. (2016). Adenine methylation in eukaryotes: Apprehending the complex evolutionary history and functional potential of an epigenetic modification. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *38*, 27–40.

Jerussi, T. P., & Glick, S. D. (1975). Apomorphine-induced rotation in normal rats and interaction with unilateral caudate lesions. *Psychopharmacologia*, *40*, 329–334.

Ji, H., Ehrlich, L. I. R., Seita, J., Murakami, P., Doi, A., Lindau, P., … Feinberg, A. P. (2010). Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature*, *467*, 338–342.

Ji, H., Zhang, X., Oh, S., Mayhew, C. N., Ulm, A., Somineni, H. K., … Khurana Hershey, G. K. (2015). Dynamic transcriptional and epigenomic reprogramming from pediatric nasal epithelial cells to induced pluripotent stem cells. *The Journal of Allergy and Clinical Immunology*, *135*, 236–244.

Jimenez-Castellanos, J., and Graybiel, A. M. (1987). Subdivisions of the dopamine-containing A8-A9-A10 complex identified by their differential mesostriatal innervation of striosomes and extrastriosomal matrix. *Neuroscience*, 23, 223–242.

Kang, L., & Gao, S. (2012). Pluripotency of induced pluripotent stem cells. *Journal of Animal Science and Biotechnology*, *3*, 5.

Karemaker, I. D., & Vermeulen, M. (2018). Single-Cell DNA Methylation Profiling: Technologies and Biological Applications. *Trends in Biotechnology*, *36*, 952–965.

Keene, C. D., Chang, R. C., Leverenz, J. B., Kopyov, O., Perlman, S., Hevner, R. F., … Montine, T. J. (2009). A patient with Huntington's disease and long-surviving fetal neural transplants that developed mass lesions. *Acta Neuropathologica*, *117*, 329–338.

Kelly, C. M., Dunnett, S. B., Rosser, A. E., Harper, P. S., Factor, S. A., Wictorin, K., … Campbell, K. (2009). Medium spiny neurons for transplantation in Huntington's disease. *Biochemical Society Transactions*, *37*, 323–328.

Kelly, C. M., Precious, S. V., Scherf, C., Penketh, R., Amso, N. N., Battersby, A., Allen, N. D., Dunnett, S. B. and Rosser, A. E. (2009). Neonatal desensitization allows long-term survival of neural xenotransplants without immunosuppression. *Nature methods*, 6(4), 271.

Kelly, C. M., Precious, S. V., Torres, E. M., Harrison, A. W., Williams, D., Scherf, C., … Rosser, A. E. (2011). Medical Terminations of Pregnancy: A Viable Source of Tissue for Cell Replacement Therapy for Neurodegenerative Disorders. *Cell Transplantation*, *20*, 503–513.

Kessaris, N., Fogarty, M., Iannarelli, P., Grist, M., Wegner, M. and Richardson, W. D. (2006). Competing waves of oligodendrocytes in the forebrain and postnatal elimination of an embryonic lineage. *Nature neuroscience*, 9(2), 173.

Kim, J. B., Greber, B., Araúzo-Bravo, M. J., Meyer, J., Park, K. I., Zaehres, H., & Schöler, H. R. (2009). Direct reprogramming of human neural stem cells by OCT4. *Nature*, *461*, 649–3.

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

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

Kim, S., & Kaang, B.-K. (2017). Epigenetic regulation and chromatin remodeling in learning and memory. *Experimental & Molecular Medicine*, *49*, e281.

Kincaid, A. E., and Wilson, C. J. (1996). Corticostriatal innervation of the patch and matrix in the rat neostriatum. *Journal of Comparative Neurology*, 374, 578–592.

Knoepfler, P. S. (2015). From bench to FDA to bedside: US regulatory trends for new stem cell therapies. *Advanced Drug Delivery Reviews*, *82–83*, 192–196.

Kravitz, A. V., and Kreitzer, A. C. (2011). Optogenetic manipulation of neural circuitry in vivo. *Current Opinions in Neurobiology*, 21, 433–439.

Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., … Ma'ayan, A. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, *44*, W90–W97.

Kyttälä, A., Moraghebi, R., Valensisi, C., Kettunen, J., Andrus, C., Pasumarthy, K. K., … Trokovic, R. (2016). Genetic Variability Overrides the Impact of Parental Cell Type and Determines iPSC Differentiation Potential. *Stem Cell Reports*, *6*, 200–212.

Lamb, T. M., & Harland, R. M. (1995). Fibroblast growth factor is a direct neural inducer, which combined with noggin generates anterior-posterior neural pattern. *Development*, *121*, 3627 LP-3636.

Lamb, T., Knecht, A., Smith, W., Stachel, S., Economides, A., Stahl, N., … Harland, R. (1993). Neural induction by the secreted polypeptide noggin. *Science*, *262*, 713–718.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., … International Human Genome Sequencing Consortium. (2001). Initial sequencing and analysis of the human genome. *Nature*, *409*, 860–921.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C. T., … Wei, C.-L. (2010). Dynamic changes in the human methylome during differentiation. *Genome Research*, *20*, 320–331.

Laurent, L. C., Ulitsky, I., Slavin, I., Tran, H., Schork, A., Morey, R., Lynch, C., Harness, J. V., Lee, S., Barrero, M. J. and Ku, S. (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.

Lee, K. J., Dietrich, P., & Jessell, T. M. (2000). Genetic ablation reveals that the roof plate is essential for dorsal interneuron specification. *Nature*, *403*, 734–740.

Lelos, M. J., Roberton, V. H., Vinh, N.-N., Harrison, C., Eriksen, P., Torres, E. M., … Dunnett, S. B. (2016). Direct Comparison of Rat- and Human-Derived Ganglionic Eminence Tissue Grafts on Motor Function. *Cell Transplantation*, *25*, 665–675.

Leptin, M. (2005). Gastrulation Movements: the Logic and the Nuts and Bolts. *Developmental Cell*, *8*, 305–320.

Lev Maor, G., Yearim, A., & Ast, G. (2015). The alternative role of DNA methylation in splicing regulation. *Trends in Genetics : TIG*, *31*, 274–280.

Levesque, M., and Parent, A. (2005). The striatofugal fiber system in primates: a reevaluation of its organization based on single-axon tracing studies. *Proceedings of the National Academy of Sciences of the United States of America*, 102, 11888–11893.

Levy, M., Boulis, N., & Svendsen, C. N. (2016). Regenerative cellular therapies for neurologic diseases. *Brain Research*, *1638*, 88–96.

Lewandowski, J., & Kurpisz, M. (2016). Techniques of Human Embryonic Stem Cell and Induced Pluripotent Stem Cell Derivation. *Archivum Immunologiae et Therapiae Experimentalis*, *64*, 349–370.

Li, E., Beard, C., & Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. *Nature*, *366*, 362–365.

Li, E., Bestor, T. H., & Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell*, *69*, 915–926.

Li, J.-Y., Christophersen, N. S., Hall, V., Soulet, D., & Brundin, P. (2008). Critical issues of clinical human embryonic stem cell therapy for brain repair. *Trends in Neurosciences*, *31*, 146–153.

Li, K., Kong, Y., Zhang, M., Xie, F., Liu, P., & Xu, S. (2016). Differentiation of pluripotent stem cells for regenerative medicine. *Biochemical and Biophysical Research Communications*, *471*, 1–4.

Li, X.-J., Zhang, X., Johnson, M. A., Wang, Z.-B., LaVaute, T., Zhang, S.-C., … Li, M. (2009). Coordination of sonic hedgehog and Wnt signaling determines ventral and dorsal telencephalic neuron types from human embryonic stem cells. *Development*, *136*, 4055–4063.

Liao, J., Karnik, R., Gu, H., Ziller, M. J., Clement, K., Tsankov, A. M., … Meissner, A. (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nature Genetics*, *47*, 469–478.

Lister, R., Mukamel, E. A., Nery, J. R., Urich, M., Puddifoot, C. A., Johnson, N. D., … Ecker, J. R. (2013). Global Epigenomic Reconfiguration During Mammalian Brain Development. *Science*, *341*, 1237905–1237905.

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

Liu, H., Kim, Y., Sharkis, S., Marchionni, L., & Jang, Y.-Y. (2011). In vivo liver regeneration potential of human induced pluripotent stem cells from diverse origins. *Science Translational Medicine*, *3*, 82ra39.

Liu, X. S., Wu, H., Ji, X., Stelzer, Y., Wu, X., Czauderna, S., … Jaenisch, R. (2016). Editing DNA Methylation in the Mammalian Genome. *Cell*, *167*, 233–247.e17.

Liu, Y., Cheng, D., Li, Z., Gao, X., & Wang, H. (2012). The gene expression profiles of induced pluripotent stem cells (iPSCs) generated by a non-integrating method are more similar to embryonic stem cells than those of iPSCs generated by an integrating method. *Genetics and Molecular Biology*, *35*, 693–700.

Lokk, K., Modhukur, V., Rajashekar, B., Märtens, K., Mägi, R., Kolde, R., … Tõnisson, N. (2014). DNA methylome profiling of human tissues identifies global and tissue-specific methylation patterns. *Genome Biology*, *15*, r54.

Lopez-Leon, M., Reggiani, P. C., Herenu, C. B., & Goya, R. G. Regenerative Medicine for the Aging Brain. *Enliven. Journal of Stem Cell Research & Regenerative Medicine*, *1*, 1–9.

Lund, R., Hauschka, S., Hoffer, B., Seiger, A., Olson, L., & Wyatt, R. (1976). Transplanted neural tissue develops connections with host rat brain. *Science*, *193*, 582–584.

Ma, L., Hu, B., Liu, Y., Vermilyea, S. C., Liu, H., Gao, L., … Zhang, S.-C. (2012). Human embryonic stem cell-derived GABA neurons correct locomotion deficits in quinolinic acid-lesioned mice. *Cell Stem Cell*, *10*, 455–464.

Ma, T., Wang, C., Wang, L., Zhou, X., Tian, M., Zhang, Q., … Yang, Z. (2013). Subcortical origins of human and monkey neocortical interneurons. *Nature Neuroscience*, *16*, 1588–1597.

MacArthur, C. C., Fontes, A., Ravinder, N., Kuninger, D., Kaur, J., Bailey, M., … Lieu, P. T. (2012). Generation of human-induced pluripotent stem cells by a nonintegrating RNA Sendai virus vector in feeder-free or xeno-free conditions. *Stem Cells International*, *2012*, 564612.

MacDonald, M. E., Ambrose, C. M., Duyao, M. P., Myers, R. H., Lin, C., Srinidhi, L., ... MacFarlane, H. (1993). A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell, 72(6), 971-983.

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

Maherali, N., Sridharan, R., Xie, W., Utikal, J., Eminli, S., Arnold, K., … Hochedlinger, K. (2007). Directly Reprogrammed Fibroblasts Show Global Epigenetic Remodeling and Widespread Tissue Contribution. *Cell Stem Cell*, *1*, 55–70.

Mahla, R. S. (2016). Stem Cells Applications in Regenerative Medicine and Disease Therapeutics. *International Journal of Cell Biology*, *2016*, 1–24.

Malik, N., & Rao, M. S. (2013). A review of the methods for human iPSC derivation. *Methods in Molecular Biology (Clifton, N.J.)*, *997*, 23–33.

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

Marchetto, M. C. N., Yeo, G. W., Kainohana, O., Marsala, M., Gage, F. H., & Muotri, A. R. (2009). Transcriptional Signature and Memory Retention of Human-Induced Pluripotent Stem Cells. *PLoS ONE*, *4*, e7076.

Marión, R. M., Strati, K., Li, H., Murga, M., Blanco, R., Ortega, S., … Blasco, M. A. (2009). A p53-mediated DNA damage response limits reprogramming to ensure iPS cell genomic integrity. *Nature*, *460*, 1149–1153.

Marthaler, A. G., Tiemann, U., Araúzo-Bravo, M. J., Wu, G., Zaehres, H., Hyun, J. K., … Tapia, N. (2013). Reprogramming to Pluripotency through a Somatic Stem Cell Intermediate. *PLoS ONE*, *8*, e85138.

Martín-Ibáñez, R., Crespo, E., Esgleas, M., Urban, N., Wang, B., Waclaw, R., … Canals, J. M. (2012). *Helios* Transcription Factor Expression Depends on *Gsx2* and *Dlx1&2* Function in Developing Striatal Matrix Neurons. *Stem Cells and Development*, *21*, 2239–2251.

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*, 7634–7638.

Matsumoto, T., Fujimori, K., Andoh-Noda, T., Ando, T., Kuzumaki, N., Toyoshima, M., … Okano, H. (2016). Functional Neurons Generated from T Cell-Derived Induced Pluripotent Stem Cells for Neurological Disease Modeling. *Stem Cell Reports*, *6*, 422–435.

Mizuno, K., Carnahan, J., & Nawa, H. (1994). Brain-Derived Neurotrophic Factor Promotes Differentiation of Striatal GABAergic Neurons. *Developmental Biology*, *165*, 243–256.

Mohn, F., Weber, M., Rebhan, M., Roloff, T. C., Richter, J., Stadler, M. B., … Schübeler, D. (2008). Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. *Molecular Cell*, *30*, 755–766.

Molero, A. E., Gokhan, S., Gonzalez, S., Feig, J. L., Alexandre, L. C., & Mehler, M. F. (2009). Impairment of developmental stem cell-mediated striatal neurogenesis and pluripotency genes in a knock-in model of Huntington's disease. *Proceedings of the National Academy of Sciences*, *106*, 21900–21905.

Monte, S. M., Vonsattel, J. P., & Richardson, E. P. (1988). Morphometric demonstration of atrophic changes in the cerebral cortex, white matter, and neostriatum in Huntington's disease. *Journal of Neuropathology and Experimental Neurology*, *47*, 516–525.

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology*, *38*, 23–38.

Morris, T. J., & Beck, S. (2015). Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data. *Methods (San Diego, Calif.)*, *72*, 3–8.

Morris, T. J., Butcher, L. M., Feber, A., Teschendorff, A. E., Chakravarthy, A. R., Wojdacz, T. K., & Beck, S. (2014). ChAMP: 450k Chip Analysis Methylation Pipeline. *Bioinformatics*, *30*, 428–430.

Morton, A. J. (2018). *Large-Brained Animal Models of Huntington's Disease: Sheep*. Humana Press, New York, NY.

Myers, R. H., Leavitt, J. L. A. F., Farrer, L. A., Jagadeesh, J., McFarlane, H., Mastromauro, C. A., ... & Gusella, J. F. (1989). Homozygote for Huntington disease. American journal of human genetics, 45(4), 615.

Nagae, G., Isagawa, T., Shiraki, N., Fujita, T., Yamamoto, S., Tsutsumi, S., … Aburatani, H. (2011). Tissue-specific demethylation in CpG-poor promoters during cellular differentiation. *Human Molecular Genetics*, *20*, 2710–2721.

Nakao, N., Grasbon-Frodl, E. M., Widner, H., & Brundin, P. (1996). DARPP-32-rich zones in grafts of lateral ganglionic eminence govern the extent of functional recovery in skilled paw reaching in an animal model of Huntington's disease. *Neuroscience*, *74*, 959–970.

Nasir, J., Floresco, S. B., O'Kusky, J. R., Diewert, V. M., Richman, J. M., Zeisler, J., … Hayden, M. R. (1995). Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. *Cell*, *81*, 811–823.

Nazor, K. L., Altun, G., Lynch, C., Tran, H., Harness, J. V., Slavin, I., … Laurent, L. C. (2012). Recurrent Variations in DNA Methylation in Human Pluripotent Stem Cells and Their Differentiated Derivatives. *Cell Stem Cell*, *10*, 620–634.

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

Nicholas, C. R., Chen, J., Tang, Y., Southwell, D. G., Chalmers, N., Vogt, D., … Kriegstein, A. R. (2013). Functional Maturation of hPSC-Derived Forebrain Interneurons Requires an Extended Timeline and Mimics Human Neural Development. *Cell Stem Cell*, *12*, 573–586.

Nicoleau, C., Varela, C., Bonnefond, C., Maury, Y., Bugi, A., Aubry, L., … Perrier, A. L. (2013). Embryonic stem cells neural differentiation qualifies the role of Wnt/β-Catenin signals in human telencephalic specification and regionalization. *STEM CELLS*, *31*, 1763–1774.

Nicoleau, C., Viegas, P., Peschanski, M., & Perrier, A. L. (2011). Human pluripotent stem cell therapy for Huntington's disease: technical, immunological, and safety challenges human pluripotent stem cell therapy for Huntington's disease: technical, immunological, and safety challenges. *Neurotherapeutics : The Journal of the American Society for Experimental NeuroTherapeutics*, *8*, 562–576.

Nishihara, T., Okahashi, N., & Ueda, N. (1993). Activin A Induces Apoptotic Cell Death. *Biochemical and Biophysical Research Communications*, *197*, 985–991.

Nishikawa, S., Goldstein, R. A., & Nierras, C. R. (2008). The promise of human induced pluripotent stem cells for research and therapy. *Nature Reviews Molecular Cell Biology*, *9*, 725–729.

Nishino, K., Arai, Y., Takasawa, K., Toyoda, M., Yamazaki-Inoue, M., Sugawara, T., … Umezawa, A. (2018). Epigenetic-scale comparison of human iPSCs generated by retrovirus, Sendai virus or episomal vectors. *Regenerative Therapy*, *9*, 71–78.

Nishino, K., Toyoda, M., Yamazaki-Inoue, M., Fukawatase, Y., Chikazawa, E., Sakaguchi, H., … Umezawa, A. (2011). DNA Methylation Dynamics in Human Induced Pluripotent Stem Cells over Time. *PLoS Genetics*, *7*, e1002085.

Nishizawa, M., Chonabayashi, K., Nomura, M., Tanaka, A., Nakamura, M., Inagaki, A., … Yoshida, Y. (2016). Epigenetic Variation between Human Induced Pluripotent Stem Cell Lines Is an Indicator of Differentiation Capacity. *Cell Stem Cell*, *19*, 341–354.

Niwa, H. (2007). How is pluripotency determined and maintained? *Development (Cambridge, England)*, *134*, 635–646.

Noakes, Z., Keefe, F., Tamburini, C., Kelly, C. M., Cruz Santos, M., Dunnett, S. B., … Li, M. (2019). Human Pluripotent Stem Cell-Derived Striatal Interneurons: Differentiation and Maturation In Vitro and in the Rat Brain. *Stem Cell Reports*, *12*, 191–200.

Nordlund, J., Bäcklin, C. L., Wahlberg, P., Busche, S., Berglund, E. C., Eloranta, M.-L., … Syvänen, A.-C. (2013). Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome Biology*, *14*, r105.

Nukaya, D., Minami, K., Hoshikawa, R., Yokoi, N., & Seino, S. (2015). Preferential gene expression and epigenetic memory of induced pluripotent stem cells derived from mouse pancreas. *Genes to Cells*, *20*, 367–381.

Numata, S., Ye, T., Hyde, T. M., Guitart-Navarro, X., Tao, R., Wininger, M., … Lipska, B. K. (2012). DNA methylation signatures in development and aging of the human prefrontal cortex. *American Journal of Human Genetics*, *90*, 260–272.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., Dolan, R. J. (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, 304, 452-454.

Ohi, Y., Qin, H., Hong, C., Blouin, L., Polo, J. M., Guo, T., … Ramalho-Santos, M. (2011). Incomplete DNA methylation underlies a transcriptional memory of somatic cells in human iPS cells. *Nature Cell Biology*, *13*, 541–549.

Okano, M., Bell, D. W., Haber, D. A., & Li, E. (1999). DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell*, *99*, 247–257.

Okano, M., Xie, S., & Li, E. (1998). Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics*, *19*, 219–220.

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

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

Olsson, M., Björklund, A., & Campbell, K. (1998). Early specification of striatal projection neurons and interneuronal subtypes in the lateral and medial ganglionic eminence. *Neuroscience*, *84*, 867–876.

Olsson, M., Campbell, K., Wictorin, K., & Björklund, A. (1995). Projection neurons in fetal striatal transplants are predominantly derived from the lateral ganglionic eminence. *Neuroscience*, *69*, 1169–1182.

Onorati, M., Castiglioni, V., Biasci, D., Cesana, E., Menon, R., Vuono, R., … Cattaneo, E. (2014). Molecular and functional definition of the developing human striatum. *Nature Neuroscience*, *17*, 1804–1815.

Ouimet, C. C., Langley-Gullion, K. C., & Greengard, P. (1998). Quantitative immunocytochemistry of DARPP-32-expressing neurons in the rat caudatoputamen. *Brain Research*, *808*, 8–12.

Paavilainen, T., Pelkonen, A., Mäkinen, M. E.-L., Peltola, M., Huhtala, H., Fayuk, D., & Narkilahti, S. (2018). Effect of prolonged differentiation on functional maturation of human pluripotent stem cell-derived neuronal cultures. *Stem Cell Research*, *27*, 151–161.

Papp, B., & Plath, K. (2013). Epigenetics of reprogramming to induced pluripotency. *Cell*, *152*, 1324–1343.

Park, I.-H., Arora, N., Huo, H., Maherali, N., Ahfeldt, T., Shimamura, A., … Daley, G. Q. (2008). Disease-Specific Induced Pluripotent Stem Cells. *Cell*, *134*, 877–886.

Park, I.-H., Zhao, R., West, J. A., Yabuuchi, A., Huo, H., Ince, T. A., … Daley, G. Q. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. *Nature*, *451*, 141–146.

Parnavelas, J. G. (2000). The origin and migration of cortical neurones: new vistas. *Trends in Neurosciences*, *23*, 126–131.

Pauly, M.-C., Döbrössy, M. D., Nikkhah, G., Winkler, C., & Piroth, T. (2013). Organization of the human fetal subpallium. *Frontiers in Neuroanatomy*, *7*, 54.

Pennarossa, G., Zenobi, A., Gandolfi, C. E., Manzoni, E. F. M., Gandolfi, F., & Brevini, T. A. L. (2015). Erase and Rewind: Epigenetic Conversion of Cell Fate. *Stem Cell Reviews*.

Perlow, M. J., Freed, W. J., Hoffer, B. J., Seiger, A., Olson, L., & Wyatt, R. J. (1979). Brain grafts reduce motor abnormalities produced by destruction of nigrostriatal dopamine system. *Science (New York, N.Y.)*, *204*, 643–647.

Peschanski, M., Cesaro, P., & Hantraye, P. (1995). Rationale for intrastriatal grafting of striatal neuroblasts in patients with Huntington's disease. *Neuroscience*, *68*, 273–285.

Petit, I., Salman Kesner, N., Karry, R., Robicsek, O., Aberdam, E., Müller, F. J., … Ben-Shachar, D. (2012). Induced pluripotent stem cells from hair follicles as a cellular model for neurodevelopmental disorders. *Stem Cell Research*, *8*, 134–140.

Pfaff, N., Lachmann, N., Kohlscheen, S., Sgodda, M., Araúzo-Bravo, M. J., Greber, B., … Moritz, T. (2012). Efficient Hematopoietic Redifferentiation of Induced Pluripotent Stem Cells Derived from Primitive Murine Bone Marrow Cells. *Stem Cells and Development*, *21*, 689–701.

Phetfong, J., Supokawej, A., Wattanapanitch, M., Kheolamai, P., U-pratya, Y., & Issaragrisil, S. (2016). Cell type of origin influences iPSC generation and differentiation to cells of the hematoendothelial lineage. *Cell and Tissue Research*, *365*, 101–112.

Pierani, A., Brenner-Morton, S., Chiang, C., & Jessell, T. M. (1999). A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. *Cell*, *97*, 903–915.

Pilz, G.-A., Shitamukai, A., Reillo, I., Pacary, E., Schwausch, J., Stahl, R., … Götz, M. (2013). Amplification of progenitors in the mammalian telencephalon includes a new radial glial cell type. *Nature Communications*, *4*, 2125.

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

Polo, J. M., Anderssen, E., Walsh, R. M., Schwarz, B. A., Nefzger, C. M., Lim, S. M., … Hochedlinger, K. (2012). A Molecular Roadmap of Reprogramming Somatic Cells into iPS Cells. *Cell*, *151*, 1617–1632.

Porfirio, B., Morelli, A., Conti, R., Vannelli, G. B., & Gallina, P. (2015). A commentary on Differentiation of pluripotent stem cells into striatal projection neurons: a pure MSN fate may not be sufficient. *Frontiers in Cellular Neuroscience*, *9*, 177.

Pradhan, S., Bacolla, A., Wells, R. D., & Roberts, R. J. (1999). Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *The Journal of Biological Chemistry*, *274*, 33002–33010.

Precious, S. V., Kelly, C. M., Reddington, A. E., Vinh, N. N., Stickland, R. C., Pekarik, V., … Rosser, A. E. (2016). FoxP1 marks medium spiny neurons from precursors to maturity and is required for their differentiation. *Experimental Neurology*, *282*, 9–18.

Precious, S. V., & Kelly, C. M. (2017). Transplantation in HD: Are We Transplanting the Right Cells? In *Huntington's Disease - Molecular Pathogenesis and Current Models*. InTech.

Pundt, L. L., Kondoh, T., Conrad, J. A., & Low, W. C. (1996 a). Transplantation of human fetal striatum into a rodent model of Huntington's disease ameliorates locomotor deficits. *Neuroscience Research*, *24*, 415–420.

Pundt, L. L., Kondoh, T., Conrad, J. A., & Low, W. C. (1996 b). Transplantation of human striatal tissue into a rodent model of Huntington's disease: phenotypic expression of transplanted neurons and host-to-graft innervation. *Brain Research Bulletin*, *39*, 23–32.

Quattrocelli, M., Palazzolo, G., Floris, G., Schöffski, P., Anastasia, L., Orlacchio, A., … Sampaolesi, M. (2011). Intrinsic cell memory reinforces myogenic commitment of pericyte-derived iPSCs. *The Journal of Pathology*, *223*, 593–603.

Quiroz, C., Luján, R., Uchigashima, M., Simoes, A. P., Lerner, T. N., Borycz, J., … Ferré, S. (2009). Key Modulatory Role of Presynaptic Adenosine A $_{2A}$ Receptors in Cortical Neurotransmission to the Striatal Direct Pathway. *The Scientific World JOURNAL*, *9*, 1321–1344.

Rallu, M., Machold, R., Gaiano, N., Corbin, J. G., McMahon, A. P., & Fishell, G. (2002). Dorsoventral patterning is established in the telencephalon of mutants lacking both Gli3 and Hedgehog signaling. *Development (Cambridge, England)*, *129*, 4963–4974.

Reddington, A. E., Rosser, A. E., & Dunnett, S. B. (2014). Differentiation of pluripotent stem cells into striatal projection neurons: a pure MSN fate may not be sufficient. *Frontiers in Cellular Neuroscience*, *8*, 398.

Reik, W., Dean, W., & Walter, J. (2001). Epigenetic Reprogramming in Mammalian Development. *Science*, *293*, 1089–1093.

Repair HD (2013). European Consortium: FP7 Heath – 2013 – 1.4 – 1; http://www.repair-hd.eu/component/content/category/83-repair-hd

Rienecker, K. DA, Hill, M. J., & Isles, A. R. (2016). Methods of epigenome editing for probing the function of genomic imprinting. *Epigenomics*, *8*, 1389–1398.

Rizzi, R., Di Pasquale, E., Portararo, P., Papait, R., Cattaneo, P., Latronico, M. V. G., … Bearzi, C. (2012). Post-natal cardiomyocytes can generate iPS cells with an enhanced capacity toward cardiomyogenic re-differentation. *Cell Death & Differentiation*, *19*, 1162–1174.

Roberton, V.H., Rosser, A.E. and Kelly, C.M. (2015). Neonatal desensitization for the study of regenerative medicine. *Regenerative medicine*, 10(3), 265-274.

Roberton, V. H., Rosser, A. E., McGorrian, A.-M., & Precious, S. V. (2018). *Dissection and Preparation of Human Primary Fetal Ganglionic Eminence Tissue for Research and Clinical Applications*.

Roberts, C. L., Chen, S. S., Murchison, A. C., Ogle, R. A., Francis, M. P., Ogle, R. C., & Sachs, P. C. (2017). Preferential Lineage-Specific Differentiation of Osteoblast-Derived Induced Pluripotent Stem Cells into Osteoprogenitors. *Stem Cells International*, *2017*, 1513281.

Robson, N. C., Wei, H., McAlpine, T., Kirkpatrick, N., Cebon, J., & Maraskovsky, E. (2009). Activin-A attenuates several human natural killer cell functions. *Blood*, *113*, 3218–3225.

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

Ross, C. A., & Tabrizi, S. J. (2011). Huntington's disease: from molecular pathogenesis to clinical treatment. *The Lancet Neurology*, 10 (1), 83-98.

Rosser, A. E., Barker, R. A., Harrower, T., Watts, C., Farrington, M., Ho, A. K., … NEST-UK. (2002). Unilateral transplantation of human primary fetal tissue in four patients with Huntington's disease: NEST-UK safety report ISRCTN no 36485475. *Journal of Neurology, Neurosurgery, and Psychiatry*, *73*, 678–685.

Rosser, A. E., & Dunnett, S. B. (2003). Neural Transplantation in Patients with Huntington's Disease. *CNS Drugs*, *17*, 853–867.

Rosser, A. E., & Bachoud-Lévi, A.-C. (2012). Clinical trials of neural transplantation in Huntington's disease. *Progress in Brain Research*, *200*, 345–371.

Rubenstein, J. L. R., Shimamura, K., Martinez, S., & Puelles, L. (1998). REGIONALIZATION OF THE PROSENCEPHALIC NEURAL PLATE. *Annual Review of Neuroscience*, *21*, 445–477.

Ruiz, S., Diep, D., Gore, A., Panopoulos, A. D., Montserrat, N., Plongthongkum, N., … Izpisua Belmonte, J. C. (2012). Identification of a specific reprogramming-associated epigenetic signature in human induced pluripotent stem cells. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 16196–16201.

Sackett, S. D., Brown, M. E., Tremmel, D. M., Ellis, T., Burlingham, W. J., & Odorico, J. S. (2016). Modulation of human allogeneic and syngeneic pluripotent stem cells and immunological implications for transplantation. *Transplantation Reviews*, *30*, 61–70.

Sadri-Vakili, G., & Cha, J.-H. J. (2006). Mechanisms of Disease: histone modifications in Huntington's disease. *Nature Clinical Practice Neurology*, *2*, 330–338.

Sanberg, P. R., Borlongan, C. V, Koutouzis, T. K., Norgren, R. B., Cahill, D. W., & Freeman, T. B. (1997). Human fetal striatal transplantation in an excitotoxic lesioned model of Huntington's disease. *Annals of the New York Academy of Sciences*, *831*, 452–460.

Sanchez-Freire, V., Lee, A. S., Hu, S., Abilez, O. J., Liang, P., Lan, F., … Wu, J. C. (2014). Effect of human donor cell source on differentiation and function of cardiac induced pluripotent stem cells. *Journal of the American College of Cardiology*, *64*, 436–448.

Sareen, D., Saghizadeh, M., Ornelas, L., Winkler, M. A., Narwani, K., Sahabian, A., … Ljubimov, A. V. (2014). Differentiation of human limbal-derived induced pluripotent stem cells into limbal-like epithelium. *Stem Cells Translational Medicine*, *3*, 1002–1012.

Savas, J. N., Ma, B., Deinhardt, K., Culver, B. P., Restituito, S., Wu, L., … Tanese, N. (2010). A role for huntington disease protein in dendritic RNA granules. *The Journal of Biological Chemistry*, *285*, 13142–13153.

Saxonov, S., Berg, P., & Brutlag, D. L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters.

Proceedings of the National Academy of Sciences of the United States of America, *103*, 1412–1417.

Schackel, S., Pauly, M.-C., Piroth, T., Nikkhah, G., & Döbrössy, M. D. (2013). Donor age dependent graft development and recovery in a rat model of Huntington's disease: Histological and behavioral analysis. *Behavioural Brain Research*, *256*, 56–63.

Schwarcz, R., & Köhler, C. (1983). Differential vulnerability of central neurons of the rat to quinolinic acid. *Neuroscience Letters*, *38*, 85–90.

Schwarcz, R., Whetsell, W. O., & Mangano, R. M. (1983). Quinolinic acid: an endogenous metabolite that produces axon-sparing lesions in rat brain. *Science (New York, N.Y.)*, *219*, 316–318.

Sela, Y., Molotski, N., Golan, S., Itskovitz-Eldor, J., & Soen, Y. (2012). Human Embryonic Stem Cells Exhibit Increased Propensity to Differentiate During the G1 Phase Prior to Phosphorylation of Retinoblastoma Protein. *STEM CELLS*, *30*, 1097–1108.

Sen, G. L., Reuter, J. A., Webster, D. E., Zhu, L., & Khavari, P. A. (2010). DNMT1 maintains progenitor function in self-renewing somatic tissue. *Nature*, *463*, 563–567.

Sheaffer, K. L., Kim, R., Aoki, R., Elliott, E. N., Schug, J., Burger, L., … Kaestner, K. H. (2014). DNA methylation is required for the control of stem cell differentiation in the small intestine. *Genes & Development*, *28*, 652–664.

Shi, Y., Inoue, H., Wu, J. C., & Yamanaka, S. (2017). Induced pluripotent stem cell technology: a decade of progress. *Nature Reviews. Drug Discovery*, *16*, 115–130.

Shtrichman, R., Germanguz, I., & Eldor, J. I.-. (2013). Induced Pluripotent Stem Cells (iPSCs) Derived from Different Cell Sources and their Potential for Regenerative and Personalized Medicine. *Current Molecular Medicine*, *13*, 792–805.

Simmons, R. K., Stringfellow, S. A., Glover, M. E., Wagle, A. A., & Clinton, S. M. (2013). DNA methylation markers in the postnatal developing rat brain. *Brain Research*, *1533*, 26–36.

Sjöholm, L. K., Ransome, Y., Ekström, T. J., & Karlsson, O. (2018). Evaluation of Post-Mortem Effects on Global Brain DNA Methylation and Hydroxymethylation. *Basic & Clinical Pharmacology & Toxicology*, *122*, 208–213.

Smyth, G. K. (2004). Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments. *Statistical Applications in Genetics and Molecular Biology*, *3*, 1–25.

Snyder, B. R., & Chan, A. W. S. (2018). Progress in developing transgenic monkey model for Huntington's disease. *Journal of Neural Transmission*, *125*, 401–417.

Sparrow, S., Manning, J. R., Cartier, J., Anblagan, D., Bastin, M. E., Piyasena, C., … Boardman, J. P. (2016). Epigenomic profiling of preterm infants reveals DNA methylation differences at sites associated with neural function. *Translational Psychiatry*, *6*, e716.

Sproul, A. A. (2015). Being human: The role of pluripotent stem cells in regenerative medicine and humanizing Alzheimer's disease models. *Molecular Aspects of Medicine*, *43–44*, 54–65.

Srinageshwar, B., Maiti, P., Dunbar, G. L., & Rossignol, J. (2016). Role of Epigenetics in Stem Cell Proliferation and Differentiation: Implications for Treating Neurodegenerative Diseases. *International Journal of Molecular Sciences*, *17*. https://doi.org/10.3390/ijms17020199

Stadtfeld, M., & Hochedlinger, K. (2010). Induced pluripotency: history, mechanisms, and applications. *Genes & Development*, *24*, 2239–2263.

Stadtfeld, M., Maherali, N., Breault, D. T., & Hochedlinger, K. (2008). Defining Molecular Cornerstones during Fibroblast to iPS Cell Reprogramming in Mouse. *Cell Stem Cell*, *2*, 230–240.

Starr, P. A., Kang, G. A., Heath, S., Shimamoto, S., & Turner, R. S. (2008). Pallidal neuronal discharge in Huntington's disease: support for selective loss of striatal cells originating the indirect pathway. Experimental neurology, 211 (1), 227-233.

Stenman, J., Toresson, H., & Campbell, K. (2003). Identification of two distinct progenitor populations in the lateral ganglionic eminence: implications for striatal and olfactory bulb neurogenesis. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*, 167–174.

Straccia, M., Garcia-Diaz Barriga, G., Sanders, P., Bombau, G., Carrere, J., Mairal, P. B., … Canals, J. M. (2015). Quantitative high-throughput gene expression profiling of human striatal development to screen stem cell–derived medium spiny neurons. *Molecular Therapy - Methods & Clinical Development*, *2*, 15030.

Sweatt, J. D. (2016). Dynamic DNA methylation controls glutamate receptor trafficking and synaptic scaling. *Journal of Neurochemistry*, *137*, 312–330.

Tabata, T. (2001). Genetics of morphogen gradients. *Nature Reviews Genetics*, *2*, 620–630.

Tabrizi, S., Leavitt, B., Kordasiewicz, H., Czech, C., Swayze, E., Norris, D. A., ... & Lane, R. (2018). Effects of IONIS-HTTRx in patients with early Huntington's disease, results of the first HTT-lowering drug trial (CT. 002).

Tabrizi, S. J., Leavitt, B. R., Landwehrmeyer, G. B., Wild, E. J., Saft, C., Barker, R. A., ... & Rosser, A. (2019). Targeting huntingtin expression in patients with Huntington's disease. *New England Journal of Medicine*, 380 (24), 2307-2316.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., & Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. *Cell*, *131*, 861–872.

Takahashi, K., & Yamanaka, S. (2006). Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*, *126*, 663–676.

Tapia, N., & Schöler, H. R. (2016). Molecular Obstacles to Clinical Translation of iPSCs. *Cell Stem Cell*, *19*, 298–309.

Tate, M. C., García, A. J., Keselowsky, B. G., Schumm, M. A., Archer, D. R., & LaPlaca, M. C. (2004). Specific β1 integrins mediate adhesion, migration, and differentiation of neural progenitors derived from the embryonic striatum. *Molecular and Cellular Neuroscience*, *27*, 22–31.

Tesarova, L., Simara, P., Stejskal, S., & Koutna, I. (2016). The Aberrant DNA Methylation Profile of Human Induced Pluripotent Stem Cells Is Connected to the

Reprogramming Process and Is Normalized During In Vitro Culture. *PLOS ONE*, *11*, e0157974.

Teschendorff, A. E., Marabita, F., Lechner, M., Bartlett, T., Tegner, J., Gomez-Cabrero, D., & Beck, S. (2013). A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics*, *29*, 189–196.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., & Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science (New York, N.Y.)*, *282*, 1145–1147.

Thomson, J. A., Kalishman, J., Golos, T. G., Durning, M., Harris, C. P., Becker, R. A., & Hearn, J. P. (1995). Isolation of a primate embryonic stem cell line. *Proceedings of the National Academy of Sciences of the United States of America*, *92*, 7844–7848.

Tian, C., Wang, Y., Sun, L., Ma, K., & Zheng, J. C. (2011). Reprogrammed mouse astrocytes retain a “memory” of tissue origin and possess more tendencies for neuronal differentiation than reprogrammed mouse embryonic fibroblasts. *Protein & Cell*, *2*, 128–140.

Tokuno, H., Chiken, S., Kametani, K., and Moriizumi, T. (2002). Efferent projections from the striatal patch compartment: anterograde degeneration after selective ablation of neurons expressing mu-opioid receptor in rats. *Neuroscience Letters*, 332, 5–8.

Toresson, H., Potter, S. S., & Campbell, K. (2000). Genetic control of dorsal-ventral identity in the telencephalon: opposing roles for Pax6 and Gsh2. *Development (Cambridge, England)*, *127*, 4361–4371.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., … Okano, M. (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes to Cells*, *11*, 805–814.

Tung, P.-Y., & Knoepfler, P. S. (2015). Epigenetic mechanisms of tumorigenicity manifesting in stem cells. *Oncogene*, *34*, 2288–2296.

Urbán, N., Martín-Ibáñez, R., Herranz, C., Esgleas, M., Crespo, E., Pardo, M., … Canals, J. M. (2010). Nolz1 promotes striatal neurogenesis through the regulation of retinoic acid signaling. *Neural Development*, *5*, 21.

Valjent, E., Bertran-Gonzalez, J., Herve, D., Fisone, G., and Girault, J. A. (2009). Looking BAC at striatal signaling: cell-specific analysis in new transgenic mice. *Trends in Neuroscience*, 32, 538–547.

Vallier, L., Alexander, M., & Pedersen, R. A. (2005). Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *Journal of Cell Science*, *118*, 4495–4509.

Ventimiglia, R., Mather, P. E., Jones, B. E., & Lindsay, R. M. (1995). The NeurotroDhins BDNF, NT-3 and NT-4/5 Promote Survival and Morphological and Biochemical Differentiation of Striatal Neurons *In Vitro*. *European Journal of Neuroscience*, *7*, 213–222.

Victor, M. B., Richner, M., Hermanstyne, T. O., Ransdell, J. L., Sobieski, C., Deng, P.-Y., … Yoo, A. S. (2014). Generation of Human Striatal Neurons by MicroRNA-Dependent Direct Conversion of Fibroblasts. *Neuron*, *84*, 311–323.

Walker, F. (2007). Huntington's Disease. *Lancet*, 369, 218-28.

Walsh, C. P., Chaillet, J. R., & Bestor, T. H. (1998). Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature Genetics*, *20*, 116–117.

Wang, L., Wang, L., Huang, W., Su, H., Xue, Y., Su, Z., … Pei, D. (2013). Generation of integration-free neural progenitor cells from cells in human urine. *Nature Methods*, *10*, 84–89.

Watts, C., Brasted, P. J., Eagle, D. M., & Dunnett, S. B. (2000). Embryonic Donor Age and Dissection Influences Striatal Graft Development and Functional Integration in a Rodent Model of Huntington's Disease. *Experimental Neurology*, *163*, 85–97.

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

Wettenhall, J. M., & Smyth, G. K. (2004). limmaGUI: A graphical user interface for linear modeling of microarray data. *Bioinformatics*, *20*, 3705–3706.

Wexler, N. S. et al. (1987). Homozygotes for Huntington's disease. *Nature*, 326, 194–197.

Wexler, N. S., Lorimer, J., Porter, J., Gomez, F., Moskowitz, C., … Landwehrmeyer, B. (2004). Venezuelan kindreds reveal that genetic and environmental factors modulate Huntington's disease age of onset. *Proceedings of the National Academy of Sciences*, *101*, 3498–3503.

Wu, M., Zhang, D., Bi, C., Mi, T., Zhu, W., Xia, L., … Wu, Y. (2018). A Chemical Recipe for Generation of Clinical-Grade Striatal Neurons from hESCs. *Stem Cell Reports*, *11*, 635–650.

Wu, S. M., & Hochedlinger, K. (2011). Harnessing the potential of induced pluripotent stem cells for regenerative medicine. *Nature Cell Biology*, *13*, 497–505.

Wu, T. P., Wang, T., Seetin, M. G., Lai, Y., Zhu, S., Lin, K., … Xiao, A. Z. (2016). DNA methylation on N6-adenine in mammalian embryonic stem cells. *Nature*, *532*, 329–333.

Xiao, B., Ng, H. H., Takahashi, R., & Tan, E.-K. (2016). Induced pluripotent stem cells in Parkinson's disease: scientific and clinical challenges. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*, 697–702.

Xu, H., Yi, B. A., Wu, H., Bock, C., Gu, H., Lui, K. O., … Chien, K. R. (2012). Highly efficient derivation of ventricular cardiomyocytes from induced pluripotent stem cells with a distinct epigenetic signature. *Cell Research*, *22*, 142–154.

Xu, Z., Liang, Q., Song, X., Zhang, Z., Lindtner, S., Li, Z., … Yang, Z. (2018). SP8 and SP9 coordinately promote D2-type medium spiny neuron production by activating Six3 expression. *Development*, *145*, dev165456.

Yap, M. S., Nathan, K. R., Yeo, Y., Lim, L. W., Poh, C. L., Richards, M., … Heng, B. C. (2015). Neural differentiation of human pluripotent stem cells for nontherapeutic applications: Toxicology, pharmacology, and in vitro disease modeling. *Stem Cells International*, Vol. 2015. https://doi.org/10.1155/2015/105172

Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. European Journal of Neuroscience. 19, 181–189.

Yoder, J. A., Yen, R. W., Vertino, P. M., Bestor, T. H., & Baylin, S. B. (1996). New 5' regions of the murine and human genes for DNA (cytosine-5)-methyltransferase. *The Journal of Biological Chemistry*, *271*, 31092–31097.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J. L., Tian, S., … Thomson, J. A. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science*, *318*, 1917 LP-1920.

Zhang, D., Cheng, L., Badner, J. A., Chen, C., Chen, Q., Luo, W., … Liu, C. (2010). Genetic Control of Individual Differences in Gene-Specific Methylation in Human Brain. *The American Journal of Human Genetics*, *86*, 411–419.

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

Zhou, W., Laird, P. W., & Shen, H. (2016). Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Research*, *45*, e22.

Zhou, Y., Wang, L., Liu, Z., Alimohamadi, S., Yin, C., Liu, J., & Qian, L. (2017). Comparative Gene Expression Analyses Reveal Distinct Molecular Signatures between Differentially Reprogrammed Cardiomyocytes. *Cell Reports*, *20*, 3014–3024.

Zietlow, R., Pekarik, V., Armstrong, R. J. E., Tyers, P., Dunnett, S. B., & Rosser, A. E. (2005). The survival of neural precursor cell grafts is influenced by in vitro expansion. *Journal of Anatomy*, *207*, 227–240.

Zietlow, R., Precious, S. V, Kelly, C. M., Dunnett, S. B., & Rosser, A. E. (2012). Long-term expansion of human foetal neural progenitors leads to reduced graft viability in the neonatal rat brain. *Experimental Neurology*, *235*, 563–573.

Zuccato, C., Ciammola, A., Rigamonti, D., Leavitt, B. R., Goffredo, D., Conti, L., ... & Timmusk, T. (2001). Loss of huntingtin-mediated BDNF gene transcription in Huntington's disease. Science, 293(5529), 493-498.