Characterising the molecular heterogeneity within human inhibitory interneurons – at single-cell resolution

Francesca Keefe

Thesis submitted for the degree of Doctor of Philosophy (PhD)

Cardiff University, 4-year Wellcome Trust funded PhD in Integrative Neuroscience

September 2019

Acknowledgements

Most importantly I would like to thank my primary supervisor, Meng Li, for her patience, encouragement and guidance over the past three years. There were occasions when either reassurance, space or gentle pushing were warranted, and I truly appreciate Meng's help to see me through these times. Although there are many people I have had the joy to meet and work with during the last three years, there are a few who deserve mentioning by name. First, Zoe Noakes for all the time she spent in the early days training me in stem cell/ tissue culture work. There were certainly rocky times, but Zoe's advice was always on point and greatly appreciated. Second, Adam Errington who provided me with training on whole-cell patch clamp technique and optogenetics. Adam's willingness to help and genuine interest in my optogenetics work made this element of my PhD possible in the short time span of my degree. Third, Sarah Edkins for her remarkable resource of knowledge concerning RNA sequencing. Sarah's guidance when setting up my single-cell RNA sequencing work helped the study to successfully develop, and become a core part of my PhD. Finally, the Wellcome Trust funding has made this work and my time in Cardiff possible, and I will be forever thankful for this opportunity that has brought me to Cardiff and to all the new people that have become part of my life as a result.

Abstract

Inhibitory interneurons are a heterogeneous class of GABAergic neurons, classified into discrete subtypes based on the expression of molecular markers, physiology, morphology and connectivity. However, the mechanism of interneuron diversification is poorly understood. hPSC-derived interneurons may help solve the mystery, by providing an unrestricted trajectory of interneuron differentiation, which is not confounded by species differences. Although, how well the model recapitulates native development remains under scrutiny. Moreover, the full extent of interneuron subtype diversity that can be achieved in vitro is unclear.

Here an evaluation of hPSC model of interneuron differentiation has been presented. To tease apart the molecular heterogeneity, longitudinal single-cell RNA sequencing was conducted. The analysis revealed a heterogeneous MGE-like enriched progenitor population that differentiated into a diverse population of interneuron-like cells. On examining the authenticity of the model, using a novel cross-comparison to the native human MGE (10WPC-15WPC), encouraging similarities were found, especially between post-mitotic hPSC-derived interneurons and 15WPC MGE. However, several shortcomings of the model were also indicated. First, under-representation of specific genes encoding for transcription factors. Second, absent or low expression of late MGE markers and subtypes. Enforced expression of the under-expressed transcription factors may provide a means to enrich the interneuron-like population with these subtypes.

Furthermore, the molecular resemblance to foetal-derived interneurons was reflected in the immature intrinsic and evoked neuronal responses of hPSC-derived interneurons, recorded using whole-cell patch clamp. Nevertheless, monosynaptic viral tracing demonstrated the capability of the interneuron-like cells to make synaptic connections, which (although not extensive) may be enhanced in a co-culture system.

Overall, the interneuron differentiation of hPSC holds promise as a neurodevelopmental model. The novel characterisation of hPSC-derived interneurons and human MGE, at single-cell resolution, has simultaneously revealed the limitations and refinements necessary to generate more authentic and defined interneuron subtypes in vitro. Thereby, widening the scope and validity of the model.

Contents list

DECLARATION..I

ACKNOWLEDGEMENTS III

ABSTRACT... IV

CONTENTS LIST VI

FIGURE LIST XIII

ABBREVIATION LIST...XV

CHAPTER 1 - GENERAL INTRODUCTION 1

1.1 Introduction.. 1

1.2 Interneurons in pathology 1

1.3 Interneuron diversity.. 2

1.4 The region of origin.. 2

1.4.1 Spatial influence 3
  1.4.1.1 Sonic hedgehog signalling 5
  1.4.1.2 WNT signalling 7
  1.4.1.3 Retinoic acid signalling..8
1.4.2 Temporal influence 9
  1.4.2.1 Early and late-born progenitors...9
  1.4.2.2 Cell cycle 10

1.5 The region of destination 11

1.5.1 Tangential migration 11
1.5.2 Circuit integration.. 11
1.5.3 Molecular and functional developmental trajectory 12
1.5.4 Region-dependent interneuron subtype composition 14
1.5.5 Significance of extrinsic cues on lineage commitment.................. 15

1.6 Lineage commitment models 16

1.7 Single-cell transcriptome analysis 19

1.7.1 Principle 19
1.7.2 Methodology... 19
1.7.3 Applications.. 22

1.8 The connectivity of neurons.. 23

1.8.1 How to measure neuronal connectivity 23
1.8.2 Principles of the monosynaptic rabies viral tracing system 24
1.8.3 Applications 25

1.9 Research models 29
1.9.1 Principles of neuron differentiation of hPSC 29
1.9.2 Interneuron differentiation of hPSC 33
1.9.3 Application of hPSC-derived neuronal models 34

1.10 General aims 34

CHAPTER 2 - MATERIAL AND METHODS 36

2.1 Stem cell and tissue culture work 36
2.1.1 General cell and primary tissue culture practice 36
2.1.2 Cell lines 36
2.1.3 Cell maintenance for hESC lines 36
2.1.4 Cell maintenance for HEK and BHK lines 37
2.1.5 Cell maintenance for human foetal tissue 38
2.1.6 Interneuron differentiation of hESC 39
2.1.7 Sequential serial expansion of hESC-derived interneuron progenitors 41

2.2 Immunohistochemistry 42

2.3 Stages prior to quantitative polymerase chain reaction 43
2.3.1 RNA extraction 43
2.3.2 DNase treatment of RNA 44
2.3.3 Reverse transcription 44

2.4 Quantitative polymerase chain reaction 45

2.5 Molecular characterisation using single-cell RNA sequencing 45
2.5.1 Fluidigm C1 – single-cell capture and generation of single-cell cDNA 45
2.5.2 Quality control checks 46
2.5.2.1 Bioanalyzer 46
2.5.2.2 Picogreen assay 47
2.5.2.3 Fluidigm Biomark - single-cell QPCR 47
2.5.3 Library preparation 48
2.5.3.1 Tagmentation and indexing 48
2.5.3.2 Bead clean up 48
2.5.3.3 Qubit assay 49

2.6 Bacterial work leading to transfections 50
2.6.1 General practice in bacterial work 50
2.6.2 Bacterial transformation 50

2.6.3 Colony selection 50
2.6.4 DNA plasmid extraction 51
2.6.5 Restriction enzyme digestion 52
2.6.6 Gel electrophoresis 53

2.7 Lentiviral vector: pBOB-synP-HTB 54
2.7.1 Transfection of HEK cells and viral harvest 54
2.7.2 Titering of lentivirus in hESC-derived interneurons 55

2.8 Pseudotyped ΔG ChR2-mCherry rabies viral tracer 55
2.8.1 Transfection of B7GG cells 56
2.8.2 Viral harvest of unpseudotyped ΔG ChR2-mCherry viral tracer 56
2.8.3 Generating EnvA pseudotyped ΔG ChR2-mCherry viral tracer 57
2.8.4 Titering of pseudotyped and unpseudotyped ΔG ChR2-mCherry viral tracer 57

2.9 Application of the completed viral tracing system 58

2.10 Flow cytometry 58

2.11 Functional characterisation of hESC-derived interneurons 59
2.11.1 Electrophysiology solutions for whole-cell patch clamp 59
2.11.2 Immediate steps prior to patch-clamp recordings 59
2.11.3 Whole-cell patch clamp 60
2.11.4 Optogenetics 60

2.12 Data processing and statistics 61
2.12.1 Figures 61
2.12.2 Statistics 61
2.12.3 Single-cell transcriptome analysis 61

CHAPTER 3 - INTERNEURON DIFFERENTIATION OF HPSC 63

3.1 Introduction 63
3.1.1 Extrinsic cue based interneuron differentiation paradigms 63
3.1.2 Shortcomings of in vitro interneuronal models 63
3.1.3 Refining interneuron differentiation paradigms that use extrinsic cues 64
  3.1.3.1 Sonic hedgehog signalling 64
  3.1.3.2 Temporal manipulation 65

3.2 Aims 66

3.3 Results 66
3.3.1 SHH treatment efficiently generated interneuron-like cells 66
3.3.2 The day 20 population was abundant in MGE-like progenitors 67

3.3.3 Terminal differentiation gave rise to diverse interneuron-like subtypes 68
3.3.4 Serial expansions generated molecularly distinct interneuron-like progenitor populations 77
3.3.5 Sequentially generated cultures showed progressively elevated SST: CR on terminal differentiation 85

3.4 Discussion 92
3.4.1 Cross-comparison of hPSC-interneuron differentiation paradigms . 92
3.4.2 The differentiation trajectory of interneuron-like cells 93
3.4.3 Extended production of interneuron-like progenitors 94
3.4.4 Early vs late-born interneuron-like progenitors 94
3.4.5 Mechanisms underpinning the elevated SST: CR interneuron 95
3.4.6 Conclusions 96

CHAPTER 4 - SINGLE-CELL TRANSCRIPTOME ANALYSIS OF HESC-DERIVED INTERNEURONS........ 97

4.1 Introduction 97
4.1.1 Single-cell transcriptomics – applications........ 97
4.1.2 Single-cell transcriptomics – interneurons 97
4.1.3 Single-cell transcriptomics – hPSC-derived interneurons 99

4.2 Aims 99

4.3 Results 99
4.3.1 hESC-derived interneuron subtypes resembled those originating from dorsal MGE and CGE 100
4.3.2 hESC-derived interneurons showed subtype-dependent expression patterns........ 100
4.3.3 Known interneuron markers showed temporal dynamics in expression profile 101
4.3.4 Differentiation trajectory of hPSC-derived interneurons was underpinned by the differential expression of specific transcription factors 109
4.3.5 Single-cell RNA sequencing data was validated by quality control checks........ 114

4.4 Discussion 119
4.4.1 Examining the subtype diversity within hPSC-derived interneurons 119
4.4.2 The diversification of hPSC-derived interneurons 121
4.4.3 Candidate lineage determining genes........ 122
4.4.4 Conclusions 123

CHAPTER 5 - SINGLE-CELL TRANSCRIPTOME ANALYSIS OF PRIMARY HUMAN FOETAL MGE 124

5.1 Introduction 124
5.1.1 Single-cell transcriptomics 124
5.1.2 Molecular characterisation of primary mouse MGE and CGE 125
5.1.3 Age-dependent changes in primary mouse MGE and CGE 126
5.1.4 Species differences and importance of primary human tissue studies 127

5.2 Aims 127

5.3 Results 127
5.3.1 The human MGE samples had consistently low contamination by neighbouring CGE and LGE regions 128
5.3.2 A discrete set of interneuron subtypes emerged within the human MGE 135
5.3.3 The human MGE was composed of a diverse population 136
5.3.4 Subtype dictated subpopulations become more abundant over time, adding to the age-dependent heterogeneity 137
5.3.5 The developmental trajectory of the MGE could be distilled down to a shortlist of differentially expressed genes 138
5.3.6 The MGE derived single-cell cDNA passed quality control checks 148
5.3.7 Cross-comparison to hESC-derived interneurons revealed a large source-dependent difference in expression profiles 153
5.3.8 Post-mitotic interneuron-like cells showed the greatest resemblance to 15WPC MGE 154

5.4 Discussion 161
5.4.1 Human MGE developmental trajectory 161
5.4.2 Interneuron lineage commitment in the MGE 162
5.4.3 hPSC-derived interneuronal models – the good and the bad 163
5.4.4 Conclusion 165

CHAPTER 6 - PHYSIOLOGY AND CONNECTIVITY OF HESC DERIVED INTERNEURONS 166

6.1 Introduction 166
6.1.1 The connectivity of neurons 166
6.1.2 Neuronal physiology 167
6.1.3 Physiology of primary tissue derived interneurons 168
6.1.4 Physiology of hPSC-derived neurons 168
6.1.5 Improving the functional maturity of hPSC-derived neurons 169

6.2 Aims 169

6.3 Results 170

6.3.1 The hESC-derived interneurons had a heterogeneous, but overall immature physiology.. 170
6.3.2 The differentiation trajectory of hESC-derived interneurons was reflected in the expression of markers associated with neuronal physiology.. 179
6.3.3 hESC-derived interneurons showed subtype-dependent expression pattern of markers associated with neuronal physiology................... 180
6.3.4 Lentiviral transduction efficiently generated hESC-derived starter neuron population 187
6.3.5 EnvA pseudotyped ΔG ChR2-mCherry viral tracer showed restricted uptake to TVA expressing cells 191
6.3.6 Monosynaptic viral tracer revealed the existence of small neuronal circuits within hESC-derived interneuronal cultures 198
6.3.7 The monosynaptic ChR2-mCherry viral tracer system had no consequence on the intrinsic physiology of hESC-derived interneurons 199
6.3.8 Photostimulation efficiently and consistently evoked action potential response in hESC-derived traced neurons..................................... 199

6.4 Discussion 205
6.4.1 Comparative assessment of the physiology of hPSC-derived interneurons... 205
6.4.2 Molecular and functional similarities of hPSC-derived interneurons and primary foetal tissue 206
6.4.3 Bridging the gap between neuron molecular and functional profiles 206
6.4.4 Viral tracers and insight into in vitro network activity and connectivity.. 210
6.4.5 Conclusion 211

CHAPTER 7 - GENERAL DISCUSSION... 212

7.1 Summary 212

7.2 Questions, limitations and future studies 213
7.2.1 Exploring the changes driven by serial expansions 213
7.2.1.1 Combining lineage tracing and single-cell RNA sequencing ... 214
7.2.1.2 Dual reporter hESC lines 217
7.2.2 Exploring the means to better bridge the gap between transcriptome and physiology 217
7.2.2.1 Patch-seq 218
7.2.3 Exploring the reasons behind the MGE and in vitro model discrepancies.. 220
7.2.3.1 MGE developmental period 220
7.2.3.2 MGE vs CGE 221
7.2.4 Exploring ways to improve hESC-derived interneuron functional maturity 221

7.2.4.1 Co-cultures of native partners 222

7.3 Wider application 223

7.3.1 Understanding the vulnerability of interneurons in neurodevelopmental disorders 223
7.3.2 Refining the interneuron differentiation of hPSC 223
7.3.3 In vitro disease modelling and drug development 228
7.3.4 Regenerative medicine 228

7.4 Concluding remarks 232

REFERENCES 233

APPENDIX 249

Figure list

CHAPTER 1. GENERAL INTRODUCTION

Figure 1.1 The main properties defining the interneuron subtype classification system. 4

Figure 1.2 Schematic coronal section of the human telencephalon labelled to define the regions of origin and destination of different interneuron subtypes. 6

Figure 1.3 Two alternative interneuron lineage commitment models. 17-18

Figure 1.4 Step-by-step guide to single-cell RNA sequencing methodology. 21

Figure 1.5 Schematic of the retrograde monosynaptic viral tracer system: step-by-step guide. 27

Figure 1.6 Methodologies used to generate hPSC-derived neurons. 31-32

CHAPTER 3. INTERNEURON DIFFERENTIATION OF HPSC

Figure 3.1 Longitudinal molecular characterisation of hESC-derived interneurons. 70-76

Figure 3.2 Effect of sequential serial expansions of hESC-derived interneuron progenitors on molecular profile of progenitors. 79-84

Figure 3.3 Effect of sequential serial expansions of hESC-derived interneuron progenitors on molecular profile of the post-mitotic interneuronal population. 86-91

CHAPTER 4. SINGLE-CELL TRANSCRIPTOME ANALYSIS OF HPSC-DERIVED INTERNEURONS

Figure 4.1 Single-cell RNA sequencing analysis of hESC-derived interneurons. 103-107

Figure 4.2 Differentially expressed genes marking the developmental trajectory of hESC-derived interneurons. 110-112

Figure 4.3 Single-cell QPCR analysis: quality control check data. 114-117

CHAPTER 5– SINGLE-CELL TRANSCRIPTOME ANALYSIS OF PRIMARY HUMAN FOETAL MGE

Figure 5.1 Primary human foetal tissue characterisation of MGE, CGE and LGE: pilot data. 129-132

Figure 5.2 Single cell RNA sequencing analysis of primary human foetal MGE. 139-146

Figure 5.3 Single-cell QPCR analysis: quality control check data. 148-151

Figure 5.4 Cross-comparison to hESC-derived interneuron single-cell RNA sequencing data. 155-159

CHAPTER 6 – PHYSIOLOGY AND CONNECTIVITY OF HESC DERIVED INTERNEURONS

Figure 6.1 Physiological parameters of hESC-derived interneurons. 171-177

Figure 6.2 Longitudinal trajectories of molecular markers associated with neuronal physiology, obtained from single-cell RNA sequencing of hESC-derived interneurons. 182-185

Figure 6.3 hESC-derived starter neurons: lentiviral transduction and titering. 187-189

Figure 6.4 The glycoprotein deficient (ΔG) ChR2-mCherry rabies viral tracer: viral transduction and titering. 192-196

Figure 6.5 Application of EnvA pseudotyped ΔG ChR2-mCherry rabies viral tracer system on hESC-derived interneurons. 200-203

CHAPTER 7 – GENERAL DISCUSSION

Figure 7.1 Clonal lineage and fate restrictions. 214-215

Figure 7.2 The principle of patch-seq. 218

Figure 7.3 Application of the dCas9 – transcriptional effector system for targeted manipulation of gene expression. 225-226

Figure 7.4 The application of hPSC-derived neurons. 229-230

Abbreviation list

ΔG	Glycoprotein deficient
5HT-R 3A	Serotonin receptor 3A
AP1S2	Adaptor related protein complex 1 subunit sigma 2
ARX1	Aristaless related homeobox 1
ASD	Autism spectrum disorders
ATP1A1	ATPase Na^+/K^+ transporting subunit alpha 1
ATP1B3	ATPase Na^+/K^+ transporting subunit beta 3
BDNF	Brain derived neurotrophic factor
BHK	Baby hamster kidney cells
BMP	Bone morphogenetic protein
CACNA1C	Voltage-gated l-type calcium channel Cav1.2 alpha 1 subunit
CACNA1D	Voltage-gated l-type calcium channel Cav1.3 alpha 1 subunit
CACNA1F	Voltage-gated l-type calcium channel Cav1.4 alpha 1 subunit
CACNA1S	Voltage-gated l-type calcium channel Cav1.1 alpha 1 subunit
CALB1	Calbindin
CALM3	Calmodulin 3
CALY	Calcyon neuron specific vesicular protein
CAMK2B	Calcium/calmodulin dependent protein kinase II beta
CAP2	Cyclase associated actin cytoskeleton regulatory protein 2
CCDC144NL	Coiled-coil domain containing 144 family N-terminal like
CCK	Cholecystokinin
CCL2	Chemokine C-C motif ligand 2
CCND1/2	Cyclin D1/2
CDH6	Cadherin 6
CDK5	Cyclin dependent kinase 5
CDKN2C	Cyclin dependent kinase inhibitor 2C
CGE	Caudal ganglionic eminence
Ch	Chromosome
ChAT	Choline acetyltransferase
CHR2	Channelrhodopsin
CKS2	CDC28 protein kinase regulatory subunit 2

CLCN3	Chloride voltage-gated channel 3
COUPTF I (NR2F1)	Coup transcription factor 1
COUPTF II (NR2F2)	Coup transcription factor 2
CPLX3	Complexin 3
CR	Calretinin
CRABP1	Cellular retinoic acid binding protein 1
CREB5	Camp responsive element binding protein 5
CRH	Corticotropin-releasing hormone
CRL	Crown-rump length
CTCF	CCCTC-binding factor
CTIP2 (BCL11B)	Coup-TF-interacting protein 2
CXCR	C-X-C motif chemokine receptor
DARPP32 (PPP1R1B)	Dopamine and cAMP-regulated neuronal phosphoprotein 32
dCas9	Endonuclease inactive/dead cas9
DISC1	Disrupted in schizophrenia 1
DKK1	Dickkopf WNT signalling pathway inhibitor 1
DLEU1	Long intergenic non-protein coding RNA 21
DLK1	Delta like non-canonical Notch ligand 1
DLX 1/2/5	Distal-less homeobox 1/2/5
DLX6-AS1	DLX6 antisense RNA 1
DPBS	Dulbecco's phosphate-buffered saline (Ca^{2+} -free, Mg^{2+} -free)
DPPA4	Developmental pluripotency-associated protein 4
DRD1/2	Dopamine receptor D1/2
E	Embryonic day
E2F1/2	E2f transcription factor 1/2
EMX1	Empty spiracles homeobox 1
EPHA4	Eph receptor a4
ER81 (ETV1)	Ets-related protein 81
ERBB4	Erb-b2 receptor tyrosine kinase 4
FABP3	Fatty acid binding protein 3
FGF	Fibroblast growth factor
FOXA2	Forkhead box A2
FOXG1	Forkhead box G1
FOXJ1	Forkhead box J1
FOXP1/2	Forkhead box P1/2
GABA	Gamma-aminobutyric acid
GABARAP	GABA type A receptor-associated protein

GABRD	Gamma-aminobutyric acid type A receptor delta subunit
GAD1	Glutamate decarboxylase
GADD45G	Growth arrest and DNA damage inducible gamma
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
GDNF	Glial derived neurotrophic factor
GEM	Gel beads in emulsion
GFP	Green fluorescent protein
GLI1/2	Gli family zinc finger 1/2
GRIA3	Glutamate ionotropic receptor AMPA type subunit 3
GRIN2 A/B/C	Glutamate ionotropic receptor NMDA type subunit 2A/B/C
GSX2 (GSH2)	Genomic screened homeobox 2
H2AFZ	H2A.Z variant histone 1
HEK	Human embryonic kidney cells
HES 4/5	Hes family BHLH transcription factor 4/5
hESC	Human embryonic stem cell
HMGB2	High mobility group box 2
hPSC	Human pluripotent stem cells
IFC	Integrated fluidic circuit
IFITM2	Interferon induced transmembrane protein 2
IGFBP-3	Insulin-like growth factor binding protein 3
IHC	Immunohistochemistry
iPSC	Induced pluripotent stem cells
ISL1	ISL LIM homeobox 1
ISYNA1	Inositol-3-phosphate synthase 1
ITPR1	Inositol 1,4,5-trisphosphate receptor type 1
KCNC1	Voltage-gated potassium channel subunit Kv3.1
KCNC2	Voltage-gated potassium channel subunit Kv3.2
KCNK1	Potassium two pore domain channel subfamily k member 1
KCNQ2	Voltage-gated potassium channel subunit Kv7.2
KCNQ3	Voltage-gated potassium channel subunit Kv7.3
KCNQ4	Voltage-gated potassium channel subunit Kv7.4
KCNQ5	Voltage-gated potassium channel subunit Kv7.5
KO	Knock-out
LB	Luria bertani

LGE	Lateral ganglionic eminence
LHX6	LIM homeobox 6/7/8
LINC0	Long intergenic non-protein coding RNA
LPPR1	Phospholipid phosphatase related 1
MAF	MAF BZIP transcription factor
MAP2	Microtubule associated protein 2
MASH1 (ASCL1)	Achaete-scute family BHLH transcription factor 1
MEA	Multi-electrode array
MEF2C	Myocyte enhancer factor 2C
MEG3	Maternally expressed 3
MEIS2	MEIS homeobox 2
mESC	Mouse embryonic stem cells
MFAP4	Microfibril associated protein 4
MGE	Medial ganglionic eminence
- RT	Reverse transcriptase free
MIR	MicroRNA
MSN	Medium spiny neurons
NDC80 COMPLEX	NDC80 kinetochore complex component
NEUN	Neuronal nuclei
NEUROD1	Neurogenic differentiation factor 1
NFI B	Nuclear factor I/B
NFI X	Nuclear factor X
NGN2	Neurogenin-2
NKX2.1 (TTF1)	NK2 homeobox 1
NKX2.2	NK2 homeobox 2
NKX6.1	NK6 homeobox 1
NKX6.2	NK6 homeobox 2
NLGN	Neuroligins
NNAT	Neuronatin
NNOS (NOS1)	Nitric oxide synthase 1
NOLZ1 (ZNF503)	Zinc finger protein 503
NPY	Neuropeptide Y
NR2E1	Nuclear receptor subfamily 2 group E member 1
NRG	Neuregulin
NRP1	Neuropilin 1
NRP2	Neuropilin 2
OLIG2	Oligodendrocyte transcription factor 2
OTX2	Orthodenticle homeobox 2
P	Postnatal day
PAX6	Paired box protein 6
PCR	Polymerase chain reaction

PDL	Poly-d-lysine
PENK	Proenkephalin
PLTP	Phospholipid transfer protein
PNOC	Prepronociceptin
POU3F4	POU class 3 homeobox 4
PROX1	Prospero homeobox protein 1
PSC	Pluripotent stem cell
PTCH1	Patched 1
PTHLH	Parathyroid hormone like hormone
PVALB	Parvalbumin
QPCR	Quantitative polymerase chain reaction
RALDH3	Retinaldehyde dehydrogenase
RB	Rabies virus
RBMS1	RNA binding motif single stranded interacting protein 1
RELN	Reelin
ROBO2	Roundabout guidance receptor 2
RRM2	Ribonucleotide reductase regulatory subunit M2
RT	Reverse transcription
S100 A11/13	Calcium signalling protein S100 calcium binding protein A11/13
S100B	S100 calcium binding protein b
SAD	Street Alabama Dufferin
SAG	Smoothened agonist
SATB1	SATB homeobox 1
SCN1A	Voltage-gated sodium channel subunit alpha Nav1.1
SCN2A	Voltage-gated sodium channel subunit alpha Nav1.2
SCN8A	Voltage-gated sodium channel subunit alpha Nav1.8
SE	Serial expansion
SEM	Standard error of the mean
SERPINF1	Serpin family f member
SFRP2	Secreted frizzled related protein 2
SHANK	SH3 and multiple ankyrin repeat domains
SHH	Sonic hedgehog
SIX3	Six homeobox 3
SKP1	S-phase kinase associated protein 1
SLC12A5	Neuronal K^+ -Cl^- ion cotransporter
SMASH	Small molecule–assisted shutoff

SMIM26	Small integral membrane protein 26
SMO	Smoothened
SNRPD2	Small nuclear ribonucleoprotein D2 polypeptide
SOX2	SRY-box 2
SOX6	SRY-box 6
SP8	Specificity protein 8
SPARC	Secreted protein acidic and cysteine rich
SPRY2	Sprouty RTK signalling antagonist 2
SST	Somatostatin
STMN2	Stathmin 2
SV2C	Synaptic vesicle glycoprotein 2C
TAC3	Tachykinin 3
TE	Transcriptional effector
TGF-B	Transforming growth factor-beta
TMBIM4	Transmembrane BAX inhibitor motif containing 4
TPM	Transcripts per million
TPM4	Tropomyosin 4
TSPAN7	Tetraspanin 7
TUJ1 (TUBB3)	Tubulin beta 3 class III
TYH	Tyrosine hydroxylase
UV	Ultra-violet
VAX1	Ventral anterior homeobox 1
VGAT	Vesicular GABA transporter
VGLUT1	Vesicular glutamate transporter 1
VIP	Vasoactive intestinal peptide
WPC	Weeks post-conception
XIST	X inactive specific transcript
YIF1B	YIP1 interacting factor homolog B
ZCCHC12	Zinc finger CCHC domain-containing protein 12

Chapter 1 - General introduction

1.1 Introduction

Inhibitory gamma-aminobutyric acid (GABA)-ergic interneurons represent a small proportion of the neuronal population (20-30%) in the brain, but have a dominant control over network activity (Lewis *et al.* 2005; Yizhar *et al.* 2011; Marín 2012). Primarily, inhibitory interneurons are required for maintaining the appropriate excitation-inhibition balance within the brain (Rubenstein & Merzenich 2003; Yizhar *et al.* 2011; Marín 2012). Moreover, the connectivity of a single interneuron onto multiple post-synaptic targets allows a synchronised, rhythmic pattern of firing. The synchronised firing pattern generates network oscillations necessary for stable, long distance transmission (Hirano *et al.* 2008; Sohal *et al.* 2009).

1.2 Interneurons in pathology

The brain-wide distribution of inhibitory interneurons means that this class of GABAergic neurons have been implicated in many different neuropathologies, including autism spectrum disorders (ASD), schizophrenia, and epilepsies (Powell *et al.* 2003; Rubenstein & Merzenich 2003; Lewis *et al.* 2005; Hirano *et al.* 2008; Cunningham *et al.* 2014). A common mechanism is believed to underpin these neurological disorders: excitation – inhibition imbalance. If interneuron function is compromised (either from developmental insults or neurodegeneration) the inappropriate shift towards excitation can occur, along with the disruption of synchronised network activity (Rubenstein & Merzenich 2003; Hirano *et al.* 2008). Depending on the circuits affected, this can manifest into affective, cognitive and motor deficits. Currently, the therapeutic treatments for these neurological disorders (if one exists) often have limited efficacy and severe side effects (Rubenstein & Merzenich 2003). However, the excitation – inhibition imbalance hypothesis has opened up an alternative, and potentially widely beneficial, therapeutic approach to pursue, namely enhancing inhibitory output (Rubenstein & Merzenich 2003). Although caution is still warranted, with questions regarding whether inhibitory interneuron deficit/dysfunction is a primary cause of the pathology (or a secondary effect) remaining unanswered. Nonetheless, the recognition of the critical role of inhibitory interneurons has encouraged research into interneuron development and function in health and disease (Rubenstein & Merzenich 2003; Marín 2012; Arber & Li 2013).

1.3 Interneuron diversity

Inhibitory interneurons represent one of the most heterogeneous class of neurons in the brain (Darmanis *et al.* 2015; Lake *et al.* 2016; Boldog *et al.* 2018). This has led to the development of a multi-factorial classification system to assign interneurons into different subtypes, with the system being refined over time to a smaller definitive set of categories (DeFelipe *et al.* 2013; Kepecs & Fishell 2014; Bandler *et al.* 2017). These categories are: morphology, physiology, site of connectivity onto post-synaptic targets, and expression of specific molecular markers (such as neuropeptides and calcium binding proteins) (Figure 1.1). With recognition of interneuron diversity came the awareness of interneuron composition being brain region-specific (Zeisel *et al.* 2015; Lake *et al.* 2016; Nowakowski *et al.* 2017; Fan *et al.* 2018). This has, in turn, focused attention on the differential impact of subtypes on neuronal activity in health and disease (Robbins *et al.* 1991; Reynolds *et al.* 2002; Lewis *et al.* 2005; Lin & Sibille 2013; Ito-Ishida *et al.* 2015). Moreover, piecing together the intricate steps of interneuron development and diversification remains a core component of interneuron research. To date the where, when and how of interneuron diversification remains under investigation (Kepecs & Fishell 2014; Kessaris *et al.* 2014; Wamsley & Fishell 2017).

1.4 The region of origin

Inhibitory interneurons are born within neighbouring regions within the ventral telencephalon, called the caudal and medial ganglionic eminences (CGE and MGE, respectively) (Zecevic *et al.* 2011; Hansen *et al.* 2013; Al-Jaberi *et al.* 2015). Within these regions there is a conserved spatial patterning of cells, which marks the ventricular, subventricular and mantle zones (Flames *et al.* 2007; Wonders *et al.* 2008; Zechel *et al.* 2014). The between-species conservation of the three zones has subsequently been confirmed, although in humans there is a further division of the subventricular zone into inner and outer layers (Zecevic *et al.* 2011; Hansen *et al.* 2013; Al-Jaberi *et al.* 2015). The cells within the ventricular to subventricular zone have a high expression of cell cycle regulation and DNA replication genes, reflecting the highly proliferative nature of the early progenitors within these zones. Hence, the expansion of the subventricular zone in humans likely accounts for the larger abundance of interneurons within the human brain compared to rodents (Wu & Parent 2000). As the progenitors exit the cell cycle and become replaced by newborn progenitors, the cells proceed into the mantle zone to become post-mitotic interneuron precursors, ready to migrate to their region of destination (Flames *et al.* 2007; Wonders *et al.* 2008; Zechel *et al.* 2014).

1.4.1 Spatial influence

Investigations regarding the potential differences in the lineage commitment of progenitors derived from the MGE and CGE have been carried out in rodent models (Flames *et al.* 2007; Butt *et al.* 2008; Wonders *et al.* 2008; Miyoshi *et al.* 2010). One approach was to modify the proportion of CGE- and MGE-derived interneurons, using a mouse line with tamoxifen-inducible oligodendrocyte transcription factor 2 (*Olig2)::Cre-ER* dependent knock-out (KO) of the transcription factor NK2 homeobox 1 (*Nkx2.1*) (Butt *et al.* 2008). The MGE boundary is marked by NKX2.1 positive cells, which become less abundant further into the neighbouring regions of the CGE and lateral ganglionic eminence (LGE) (Zecevic *et al.* 2011; Hansen *et al.* 2013; Onorati *et al.* 2014). Therefore, on administrating tamoxifen to pregnant dams at different times of gestation (corresponding to embryonic day E9.5, E10.5, or E12.5) a deficiency in *Nkx2.1* was induced (Butt *et al.* 2008). The *Nkx2.1* deficiency was reflected in the depletion of MGE committed cells, and ectopic expansion of CGE and LGE regions in the mouse embryos (irrespective of timing of tamoxifen administration) (Butt *et al.* 2008). As a result, in the adult brain the proportion of interneurons derived from the CGE, at the expense of the MGE, was elevated (Butt *et al.* 2008). But did this impact the interneuron subtype diversity in the adult mouse brain? In short, yes. The cohort of mice had significantly reduced number of somatostatin (SST) and parvalbumin (PVALB) interneurons, with sparing/ heightened levels of the calretinin (CR), vasoactive intestinal peptide (VIP) and reelin (RELN) interneurons (Butt *et al.* 2008). This led to the conclusion that the spared

Figure 1.1

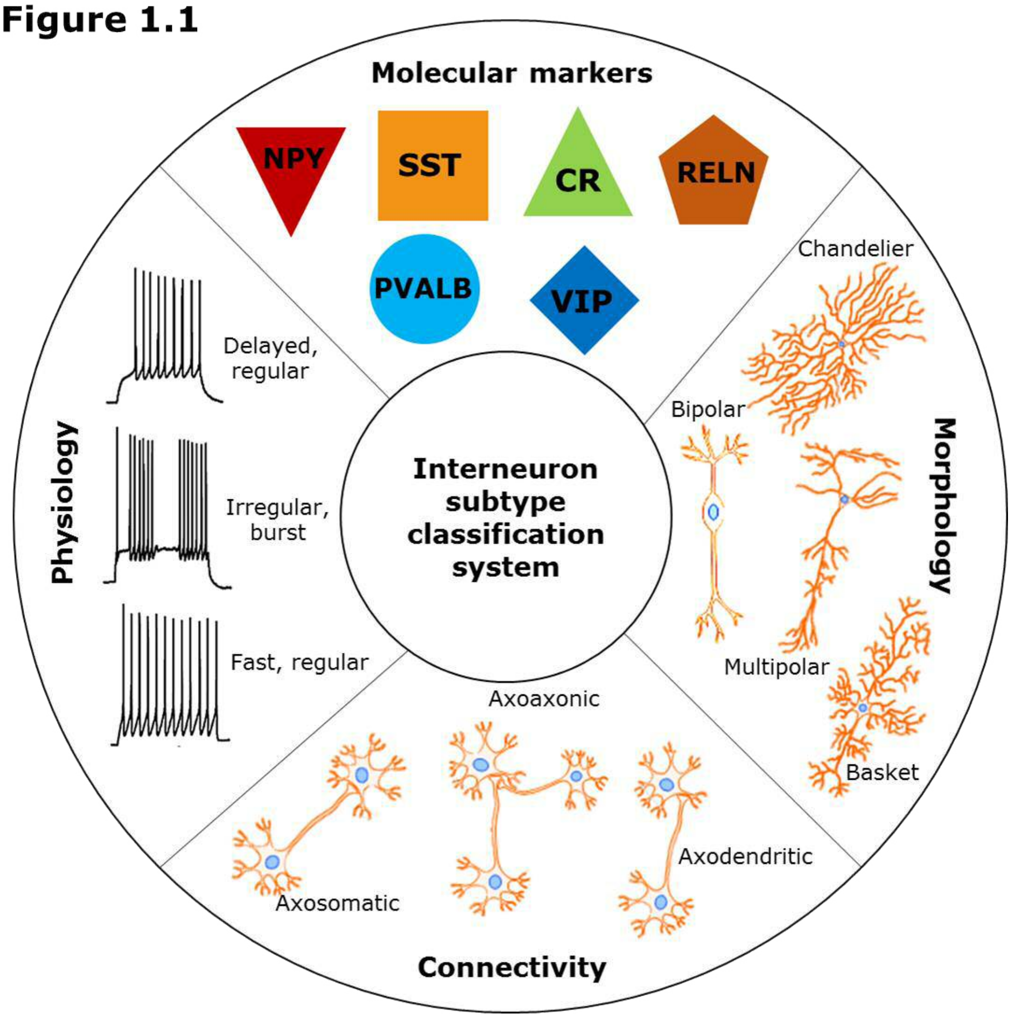

Molecular markers
NPY
SST
CR
RELN
PVALB
VIP
Chandelier
Bipolar
Morphology
Multipolar
Basket
Axoaxonic
Axodendritic
Axosomatic
Connectivity
Physiology
Delayed, regular
Irregular, burst
Fast, regular
Interneuron subtype classification system

Figure 1.1 The main properties defining the interneuron subtype classification system.

The molecular and functional heterogeneities within inhibitory interneurons has resulted in this neuronal class being divided into distinct subtypes. The assignment to a particular subtype is determined using a complex multifactorial classification system that takes into account: the expression of molecular markers (calcium binding proteins and neuropeptides), morphology, connectivity onto post-synaptic partners, and physiology of the interneuron. Cartoon illustrations of neurons were collected from https://www.motifolio.com/neuroscience-all.html (Motifolio 2018).

Neuropeptide Y, NPY. Somatostatin, SST. Parvalbumin, PVALB. Calretinin, CR. Vasoactive intestinal peptide, VIP. Reelin, RELN.

subtypes must be generated through *Nkx2.1* independent pathway within the CGE, whereas the depleted subtypes were predominantly MGE-derived (Butt *et al.* 2008). Furthermore, Butt *et al.* investigated whether this shift in the interneuron subtype diversity led to any aberrant behaviours in adulthood, possibly reflective of an excitation-inhibition imbalance (Butt *et al.* 2008). This was indeed the case, with mice displaying motor coordination impairments (dyskinesia and ataxia) and myoclonic seizures (Butt *et al.* 2008). The phenotype was most severe, and often resulted in premature death (by postnatal day P11), when *Nkx2.1* deficiency was induced at the earliest time point (E9.5 administration of tamoxifen to dams) (Butt *et al.* 2008). This hinted at the non-overlapping roles of different interneuron subtypes in the brain, and the differential contributions in neuropathologies (Robbins *et al.* 1991; Reynolds *et al.* 2002; Lewis *et al.* 2005; Ito-Ishida *et al.* 2015).

With these findings and fate-mapping experiments in rodents, the following guide on the origin of interneuron subtypes has become established (Figure 1.2). Classically, CR, VIP and RELN are CGE-derived interneurons, and neuropeptide Y (NPY), SST and PVALB are MGE-derived (Xu *et al.* 2004; Butt *et al.* 2005; Flames *et al.* 2007; Wonders *et al.* 2008; Miyoshi *et al.* 2010). In summary, the location of birth influences the lineage commitment of interneuron precursors, but how?

1.4.1.1 Sonic hedgehog signalling

Sonic hedgehog (SHH) is an endogenous morphogen released under a dorsal-ventral gradient across the ventral telencephalon (Flames *et al.* 2007; Wonders *et al.* 2008; Xu *et al.* 2010; Danjo *et al.* 2011; Anderson *et al.* 2016). SHH signalling is highest within the MGE, peaking in the dorsal MGE (Flames *et al.* 2007; Wonders *et al.* 2008). To determine whether changes in the relative strength of SHH signalling affected the lineage commitment of cells, a mouse line with SIX homeobox 3 (*Six3)::Cre*-dependent KO of SHH receptor, Smoothened (*Smo*) was used (Xu *et al.* 2010). The MGE marker *Six3* is expressed by early proliferating progenitors in MGE ventricular zone. Therefore, the reduction of SHH signalling activation was restricted to early MGE progenitors, and resulted in the MGE

Figure 1.2

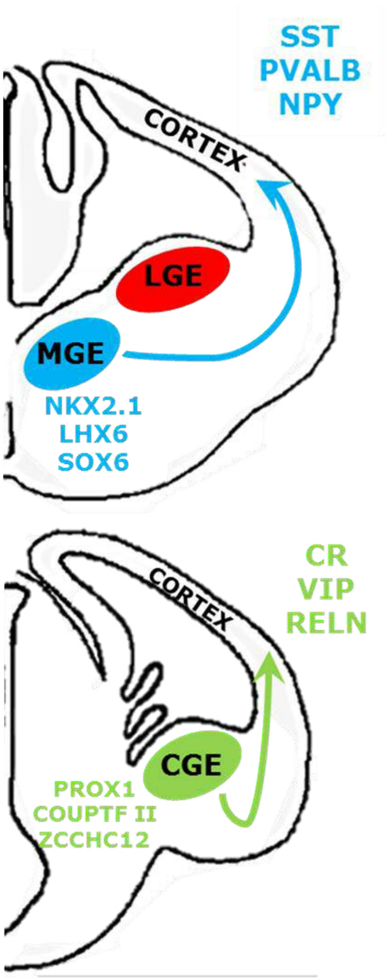


Figure 1.2 Schematic coronal section of the telencephalon labelled to define the regions of origin and destination of different interneuron subtypes.

Within the ventral telencephalon, interneurons are born within focal neighbouring regions, the MGE and CGE. Region of origin (CGE/MGE) specific molecular markers are illustrated for CGE- and MGE-derived progenitors in green and blue respectively. Interneuron progenitors reach their appropriate region of destination (e.g. cerebral cortex) through tangential migration, and then integrate to form neuronal circuits. Classic CGE- and MGE- derived interneuron subtype markers are highlighted in green and blue respectively. Coronal outline of the telencephalon was initially derived from mouse telencephalon figure in Brandão & Romcy-Pereira 2015.

Neuropeptide Y, NPY. Somatostatin, SST. Parvalbumin, PVALB. Calretinin, CR. Vasoactive intestinal protein, VIP. Reelin, RELN. NK 2 Homeobox 1, NKX2.1. LIM homeobox protein 6, LHX6. SRY-box 6, SOX6. Prospero homeobox protein 1, PROX1. COUP transcription factor 2, COUPTF II. Zinc finger CCHC domain-containing protein 12, ZCCHC12. MGE, medial ganglionic eminence. CGE, caudal ganglionic eminence. LGE, lateral ganglionic eminence.

region shrinking and ectopic CGE expansion during embryonic development (Xu *et al.* 2010). This, in turn, led to the higher abundance of CGE-derived CR interneurons at the expense of MGE-derived SST and PVALB interneurons in the adult mouse brain (Xu *et al.* 2010). The similarities in the *Olig2::Cre-ER* dependent KO *Nkx2.1* mice and *Six3::Cre*-dependent KO *Smo* mice led to the conclusion that high SHH signalling activates the expression of *Nkx2.1*, which in turn promotes MGE commitment and represses CGE commitment (Butt *et al.* 2008; Xu *et al.* 2010).

Additionally, the differential levels of SHH activation within the ventral and dorsal MGE ventricular zone leads to the establishment of molecularly distinct progenitor pools within these subdomains (Flames *et al.* 2007; Wonders *et al.* 2008). The distinction between progenitors within the dorsal and ventral mouse MGE has been demonstrated at RNA and protein level, with many of the differentially expressed genes corresponding to transcription factors, including the dorsally elevated genes, GLI family zinc finger 1/2 (*Gli1/2*), NK6 homeobox 2 (*Nkx6.2*), COUP transcription factor 2 (*Couptf II*) and ventrally elevated genes forkhead box J1 (*Foxj1*), Ets-related protein 81 (*Er81*) and LIM homeobox family genes (including *Lhx 6/7/8*) (Flames *et al.* 2007; Wonders *et al.* 2008; Mi *et al.* 2018).

Moreover, evidence for the differential lineage commitment of ventral and dorsal MGE derived progenitors has been demonstrated using utero-transplantation of reporter labelled whole, ventral or dorsal MGE ventricular zones (at E13.5), into age-matched wild-type mouse embryos (Flames *et al.* 2007; Wonders *et al.* 2008). On analysing the interneuron composition of each graft at P14 in the mouse cortex, a significant preferential bias towards SST fate at the expense of PVALB was observed in dorsal MGE-derived grafts (Flames *et al.* 2007; Wonders *et al.* 2008). Whereas, PVALB bias was observed for ventral MGE-derived grafts (Flames *et al.* 2007; Wonders *et al.* 2008).

1.4.1.2 WNT signalling

Another spatial patterning morphogen in the foetal brain is WNT, which is under a caudal-rostral and dorsal-ventral gradient (WNT gradient highest caudally and dorsally). To counteract WNT signalling, WNT signalling inhibitors, such as the endogenous morphogen dickkopf WNT signalling pathway inhibitor 1 (DKK1), are released along a reciprocal gradient (Nicoleau *et al.* 2013). The restriction of canonical WNT signalling is critical for maintaining the pallium/ subpallium boundary (Backman *et al.* 2005). Disrupting the fine-balance between canonical WNT signalling activation and inhibition along caudal-rostral and dorsal-ventral gradients, as seen in *Dkk1* KO mouse, led to gross abnormalities, including the loss of the anterior brain, additional limb formation, and was ultimately embryonic lethal (Backman *et al.* 2005). Due to the severe consequence of manipulating WNT signalling during development, conditional loss/gain-of-function rodent models have been developed to limit the WNT signalling disruption to a restricted time-window. Canonical WNT/ β-catenin signalling was ectopically activated using a conditional gain-of-function β-catenin mouse model (at E8 or E11), which resulted in the depletion of ventral telencephalon committed cells (Backman *et al.* 2005). Furthermore, the subpallium region boundary was partially distorted, with the invasion of dorsal pallial cells (paired box protein 6, *Pax6* and neurogenin-2, *Ngn2* co-expression) within this region (Backman *et al.* 2005). Whereas, the loss-of-function β-catenin mouse model at E8 (but not E11) had an expanded subpallium region. In conclusion, canonical WNT signalling appears to act at early stages of development to both repress ventral fate, and promote dorsal fate, with loss/gain of canonical WNT signalling past E11 having no apparent consequence on the fate of these progenitors (Backman *et al.* 2005).

1.4.1.3 Retinoic acid signalling

Retinoic acid signalling contributes to many pathways, including promoting caudal fate commitment of neuroepithelial cells and driving the maturation and survival of post-mitotic neurons (Li *et al.* 2009; Goulburn *et al.* 2011; Cambray *et al.* 2012). The role of retinoic acid signalling has been investigated in rodent models through the KO of retinaldehyde dehydrogenase (*Raldh3*), the enzyme responsible for retinoic acid production in the ventral telencephalon (Chatzi *et al.* 2011). Chatzi *et al.* demonstrated that retinoic acid signalling was required for the induction of glutamate decarboxylase (*Gad1*) expression (Chatzi *et al.* 2011). The *Raldh3* KO mice failed to generate ventral telencephalon derived GABAergic neurons, including LGE-derived medium spiny neurons (MSN) and MGE/CGE-derived inhibitory interneurons due to the lack of expression of *Gad1*, and subsequent production of GABA (Chatzi *et al.* 2011).

1.4.2 Temporal influence

In addition to the place of birth, differences in the birth date of CGE and MGE progenitors has been implicated in dictating the lineage commitment of interneurons. During MGE/CGE development the ventricular zone is enriched in proliferating progenitors, with a pool of newborn progenitors emerging through sequential divisions of cells from a common ancestral parent cell (Arendt *et al.* 2016). Certain developmental models suggest that diversification of cell types during development revolve around changes in the intrinsic programming of multipotent neural progenitors that occur during sequential divisions (Arendt *et al.* 2016). These models would suggest that each division holds the potential to generate temporally distinct progenitor pools that have different lineage commitments (Arendt *et al.* 2016).

1.4.2.1 Early and late-born progenitors

Evidence of the temporal influence over interneuron lineage commitment has come about from fate-mapping studies, tracking the differentiation of different aged CGE/MGE grafts (E13.5 and E15.5) (Butt *et al.* 2005; Inan *et al.* 2012). Utero-transplantation of the reporter labelled CGE/MGE-derived grafts into wild-type hosts (of the matching age) were left to terminally differentiate until P14-21 (Butt *et al.* 2005; Inan *et al.* 2012). The reporter labelled cells were then identified at both the molecular and functional level using immunohistochemistry and whole-cell patch clamp recordings respectively (Butt *et al.* 2005; Inan *et al.* 2012). The grafts derived from the CGE at the earliest age (E13.5) had the greatest proportion of SST interneurons, and to a lesser extent generated CR, VIP and NPY subtypes (and even fewer PVALB) (Butt *et al.* 2005). Whereas, CGE E15.5 progenitors differentiated into an interneuron population equally abundant across all subtype markers (20-30% abundance of SST, CR, VIP, NPY and PVALB) (Butt *et al.* 2005). The composition of the MGE-derived grafts after terminal differentiation were distinct from that reported for CGE-derived grafts, supporting the spatial influence over lineage commitment described above (Section 1.4.1). Namely, a high proportion of SST and PVALB interneurons were established in the MGE-derived grafts (at both E13.5 and E15.5), with a very small proportion of alternative subtypes (with <10% VIP, CR, RELN). Although the proportion of SST and PVALB interneurons derived from E13.5 and E15.5 MGE grafts were similar, a specific subtype of PVALB interneurons (the Chandelier cells), were only obtained from E15.5 MGE grafts (Inan *et al.* 2012). The difference between the differentiation potential of the newborn MGE/CGE progenitors at E13.5 and E15.5 provided novel evidence for early and late-born progenitors from the same region having differential interneuron lineage commitments (Butt *et al.* 2005; Inan *et al.* 2012). However, as these findings have only been obtained in mouse, whether the same temporal bias in lineage commitment is replicated in human MGE/CGE development is still unknown.

1.4.2.2 Cell cycle

The intrinsic clock mechanism describes the dual influence of when the progenitor is born, and for how long the progenitor has been proliferating, has over the differentiation potential of the progenitor (Bandler *et al.* 2017). The period of time the progenitor remains proliferating in the ventricular zone must be strictly controlled in order to maintain the appropriate numbers of early and late-born progenitors. When the cell cycle regulation becomes aberrant, as seen in *Nxk2.1::Cre*-dependent KO of *Couptf II*, the proportion of progenitors belonging to early and late-born pools can inappropriately shift. In the case of *Couptf II* deficiency, the late-born born progenitor pools are elevated at the expense of the early-born progenitors. The subsequent impact in adult mouse brain was the increase in PVALB interneurons at the expense of SST and choline acetyltransferase (ChAT) interneurons (Hu *et al.* 2017a). This has been linked to the loss of *Couptf II* expression in MGE progenitors resulting in the increased expression of cyclin D2 (*Ccnd2)* (Huard *et al.* 1999; Glickstein *et al.* 2007; Gilani *et al.* 2014; Hu *et al.* 2017a). CCND2 acts to shorten the duration of G1-S phase of the cell cycle, thereby enhancing the rate of proliferation (Huard *et al.* 1999; Mitsuhashi *et al.* 2001; Glickstein *et al.* 2007). As a result, global KO of *Ccnd2* in mice causes severe microcephaly (Huard *et al.* 1999). On closer evaluation of changes in interneuron composition and abundance in the adult *Ccnd2* KO mice, a selective deficit in PVALB interneurons was observed in multiple brain regions (cortex, hippocampus, cerebellum), with no changes reported in other subtypes (Glickstein *et al.* 2007). The selective deficit in PVALB interneurons complements findings from conditional *Couptf II* KO mice that had enhanced *Ccdn2* expression and increased abundance of PVALB interneurons (Hu *et al.* 2017a). Therefore, *Couptf II* was concluded to have a critical role in downregulating *Ccnd2* in order to maintain the appropriate balance of early and late-born progenitor pools, which will give rise to the correct SST:PVALB ratio (Hu *et al.* 2017a). Moreover, this work provided evidence for late-born MGE progenitors being biased towards PVALB fate (Hu *et al.* 2017a).

1.5 The region of destination

1.5.1 Tangential migration

During early stages of gestation, typically from 8 weeks post-conception (WPC) in humans, immature interneuron precursors within the MGE/CGE have begun tangential migration to the appropriate region of destination (Kang *et al.* 2011; Onorati *et al.* 2014; Zhong *et al.* 2018). Neuronal migration is guided by cues released under a concentration gradient that act to either attract or repress migratory neurons (Elias *et al.* 2008; Elbert *et al.* 2019). The interneuron precursors predestined for a particular location are selectively receptive to the appropriate guidance cues by the expression of the complementary set of guidance receptors (such as neuropilin-2, *Nrp2* and C-X-C motif chemokine receptor *Cxcr* family) (Elias *et al.* 2008). When this expression pattern becomes altered this can lead to subtype and brain region selective deficits in interneuron numbers in the adult brain. These changes were observed in *Nestin::Cre*-dependent CCCTC-binding factor (*Ctcf*) KO mice (Elbert *et al.* 2019). The *Ctcf* deficiency impacted selective pools of SST and PVALB committed precursors, and led to the poor expression of C-X-C motif chemokine receptors (*Cxcr4* and *Cxcr7*), thus stunting the ability of these precursors to migrate to the cortex (Elbert *et al.* 2019). This led to the selective depletion of PVALB and SST interneuron numbers in the adult cortex (Elbert *et al.* 2019).

1.5.2 Circuit integration

On arrival to the region of destination, the interneuron precursors must integrate appropriately with synaptic partners. Primarily, interneurons connect directly onto the principal cells of the region (Varga *et al.* 2015).

Alternatively, a small proportion of interneurons (notedly VIP) innervate other inhibitory interneurons to indirectly elevate excitation through disinhibition (Wolff *et al.* 2014; Lee *et al.* 2019). The integration of interneurons to form viable synaptic connections to influence neuronal network signalling marks the final stages of interneuron development. During this process both the abundance and position of interneurons within the region becomes refined, in a mechanism based on the interplay of extrinsically driven factors (e.g. contact-dependent signalling between interneurons and principal neurons), and innately driven differential expression of layer markers (Darmanis *et al.* 2015; Lake *et al.* 2016; Frazer *et al.* 2017; Nowakowski *et al.* 2017; Mickelsen *et al.* 2019).

1.5.3 Molecular and functional developmental trajectory

The molecular transition of interneuron precursors within the region of destination has been monitored longitudinally for human cortical foetal tissue across 14WPC to 26WPC period (Nowakowski *et al.* 2017; Fan *et al.* 2018; Zhong *et al.* 2018; Mayer *et al.* 2019). These studies provided evidence for an age-dependent rise in the abundance and heterogeneity of the cortical interneuron population, that was region and subtype-specific (Nowakowski *et al.* 2017; Fan *et al.* 2018; Zhong *et al.* 2018; Mayer *et al.* 2019). Moreover, although large changes occurred in the expression profile of cortical interneurons, certain genes displayed constant expression throughout interneuron development, remaining in adulthood (Fan *et al.* 2018). These specific markers are believed to act as an early molecular signature of this subtype (Fan *et al.* 2018).

The functional maturity of cortical interneurons is marked by the progressive increase in expression of voltage-gated ion channels, glutamate receptor (AMPA and NMDA) subunits and synaptic proteins (vesicular GABA transporter, synaptophysin, synapsin) (Moore *et al.* 2009; Moore *et al.* 2011; Mayer *et al.* 2019). As the expression of these physiological markers rise in the foetal interneuron, the physiology of the foetal interneuron transitions into that known of their adult counterparts (Varga *et al.* 2015; Lake *et al.* 2016; Boldog *et al.* 2018).

Intracellular calcium imaging recordings detect transient rises in intracellular Ca^{2+} ion levels, using an intracellular fluorescent calcium indicator that has a fluorescence intensity dependent on Ca^{2+} ion concentration (Sun *et al.* 2016; Xiang *et al.* 2017; Mayer *et al.* 2019). The changes in fluorescence intensity are used as indications of spike activity (Sun *et al.* 2016; Xiang *et al.* 2017; Mayer *et al.* 2019). Ex vivo calcium imaging recordings were performed to determine the responsiveness of foetal cortical neurons (14WPC-22WPC) to different neurotransmitters (GABA, glutamate, cholinergic and serotonergic receptor agonists) in the bathing solution. For the cortical interneurons, the calcium responses were generally absent or weak (Mayer *et al.* 2019). The only response detected was AMPA glutamatergic receptor-mediated, and this was restricted to MGE-derived cortical interneurons (absent in CGE-derived subtypes) (Mayer *et al.* 2019). As the expression of AMPA receptor subunits increased progressively with functional maturity, this implied a potential region of origin dependent maturation trajectory of cortical interneurons (Mayer *et al.* 2019). Alternatively, this could be linked to MGE-derived SST interneuron subtypes being the first to arrive in the cortex, thereby providing additional time to make and strengthen connections to excitatory cortical neurons, leading to a stronger AMPA evoked response (Zhong *et al.* 2018; Mayer *et al.* 2019).

Furthermore, the functional immaturity of foetal cortical neurons was determined using a direct measure of neuronal physiology, whole-cell patch clamp (Moore *et al.* 2009; Moore *et al.* 2011; Calvigioni *et al.* 2016; Pan *et al.* 2016; Zhong *et al.* 2018). In a similar manner to the calcium imaging studies, ex vivo recordings were performed on human foetal cortical tissue up to age 26WPC (Moore *et al.* 2009; Moore *et al.* 2011; Zhong *et al.* 2018). Even by the oldest age sampled, foetal cortical neurons typically remained passive to current injections and had limited spontaneous activity. In addition, foetal tissue cortical neurons did not display intrinsic physiology equivalent to adult primary tissue-derived cortical neurons, with a high input resistance, low capacitance, and weakly hyperpolarised resting membrane potential (Moore *et al.* 2009; Moore *et al.* 2011). Although, the identity of the cortical neurons remained undetermined (potentially masking any differences between cortical pyramidal cells and interneurons) ex vivo whole-cell patch clamp recordings on reporter labelled interneurons in the prenatal and early postnatal mouse brain slices described a similar profile (Moore *et al.* 2009; Moore *et al.* 2011; Calvigioni *et al.* 2016; Pan *et al.* 2016).

1.5.4 Region-dependent interneuron subtype composition

The brain-wide distribution of interneurons led to investigation determining if interneuron composition and abundance is region-specific. So far this analysis has only been conducted on regions within the human adult cortex, derived from a single tissue source (51 year old female) (Lake *et al.* 2016). Nonetheless, this analysis provided evidence for between-region differences in interneuron subtypes. A prime example was SST/RELN interneurons (also co-expressing synaptic vesicle glycoprotein 2C *SV2C*), which are a widely abundant layer III cortical interneuron subtype, but absent from primary visual and auditory cortices (Lake *et al.* 2016). To explore between-region differences in a more extensive number of brain regions, molecular and functional analysis of adult (P21-P35) mouse primary tissue has been conducted. These brain regions included: primary somatosensory cortex, hippocampus (CA1 region), and dorsal striatum (Zeisel *et al.* 2015; Földy *et al.* 2016; Gokce *et al.* 2016; Muñoz-Manchado *et al.* 2018).

On characterising the interneuron subpopulations, a high degree of overlap was detected in the interneuron subtype composition between the mouse hippocampus, dorsal striatum and somatosensory cortex (Zeisel *et al.* 2015; Földy *et al.* 2016; Gokce *et al.* 2016; Muñoz-Manchado *et al.* 2018). Although, the dorsal striatal interneuron subtype composition appeared to be the least diverse. In addition, the relative abundance of common subtypes was also region-dependent, for example fast spiking PVALB interneurons were more abundant in the hippocampus CA1 region (Zeisel *et al.* 2015). Whereas, regular-spiking CGE-derived (serotonin receptor 3A, *5ht-r 3a* expressing) interneuron subtypes were more abundant in somatosensory cortex (Földy *et al.* 2016). Moreover, each brain region had a unique region-specific set of interneuron subtypes (Zeisel *et al.* 2015; Földy *et al.* 2016; Gokce *et al.* 2016; Muñoz-Manchado *et al.* 2018). For example, VIP/CR interneurons (also co-expressing Proenkephalin *Penk,* and corticotropin-releasing hormone *Crh*), and RELN interneurons (co-express *Lhx6*, GABA type A receptor delta subunit *Gabrd*) were restricted to somatosensory cortex and hippocampus respectively (Zeisel *et al.* 2015; Földy *et al.* 2016). The mouse dorsal striatum had the largest number of unique region-specific subtypes including: parathyroid hormone like hormone (*Pthlh*)/*Pvalb* co-expressing interneurons (most abundant), cholinergic interneurons (*Gad1* negative) and tyrosine hydroxylase (*Tyh*)/ *Gad1* co-expressing interneurons (Muñoz-Manchado *et al.* 2018). Also of interest was the region-specific differences in expression of genes related to formation and maintenance of perineuronal nets and synaptic plasticity (Muñoz-Manchado *et al.* 2018). This led to the conclusion that interneurons within a particular region express a unique set of genes that complement their surroundings. However, whether the region-appropriate gene expression pattern was already innately pre-programmed, or moulded within the region of destination remains unclear.

1.5.5 Significance of extrinsic cues on lineage commitment

Fate-mapping studies have been used to investigate the influence that region-specific cues have over determining the terminal fate of the interneurons (Bele *et al.* 1995; Rolando *et al.* 2010; Ishino *et al.* 2017; Quattrocolo *et al.* 2017). In brief, the experimental design centred around tracking the fate of interneuron precursors that originated from one region (donor region), but had been transplanted into an alternative region (host region). These so-called heterotopic transplanted grafts were compared to controls, which had been transplanted back into the donor region (homotopic transplant) (Ishino *et al.* 2017; Quattrocolo *et al.* 2017). For example, P0-2 hippocampal and cortical tissue-derived interneuron precursors were isolated from *Nkx2.1*-driven reporter mouse line, and transplanted into wild-type mice of matched age. On 30 days post-transplantation, some interneuron subtypes specific to the donor region manifested in the grafts independent of the site of transplantation. This included the establishment of hippocampal-specific type II nitric oxide synthase 1 (nNOS) interneurons in hippocampal-derived grafts transplanted into the cortex and hippocampus (Quattrocolo *et al.* 2017). Whereas, the abundance of other interneuron subtypes (e.g. *PVALB* expressing Chandelier cells) appeared to be dependent on host region (Ishino *et al.* 2017; Quattrocolo *et al.* 2017). This could be due to several reasons. First, the subtype commitment of the interneuron precursors was subject to modifications by extrinsic cues, and by altering the region of destination the final lineage commitment of the interneuron was changed. The extrinsic influences mainly refer to the shaping of interneuron fate by the inputs/outputs the interneuron precursor is subjected to within the region of destination (an activity-dependent mechanism) (Wamsley & Fishell 2017). However, this theory is questioned due to independent studies finding an incomplete/absent conversion of interneuron fates to match the host region (Bele *et al.* 1995; Rolando *et al.* 2010). The second mechanism favours an intrinsically driven lineage commitment, with interneuron fate restricted early in development (in the region of origin). Therefore, to explain changes observed in interneuron subtype composition reported by Ishino *et al.* and Quattrocolo *et al.* an alternative theory has been proposed, namely the selective survival bias of interneuron precursors committed to subtypes that matched the host region (Southwell *et al.* 2012). To date, which explanation was correct or whether the mechanism is subtype-dependent remains unknown (Quattrocolo *et al.* 2017). Moreover, there is evidence to suggest that responsiveness to local environmental cues progressively decreases with developmental age of the interneuron precursors, adding additional complexities to untangling the mechanism (Grimaldi *et al.* 2005).

1.6 Lineage commitment models

Ultimately, these findings have led to the proposal of two alternative interneuron lineage commitment models illustrated in Figure 1.3. Model one favours an intrinsically guided fate restriction, with the place and time of birth acting to preprogram the fate of the cell within the region of origin, and this remains largely independent of extrinsic cues in the region of destination (Bele *et al.* 1995; Rolando *et al.* 2010) (Figure 1.3). However, the conflicting evidence that indicates a more dominant extrinsic influence over dictating interneuron lineage has led to the proposal of the second model (Ishino *et al.* 2017; Quattrocolo *et al.* 2017). In this model interneuron precursors have a certain degree of flexibility in their lineage commitment, and differential extrinsic cues (synaptic inputs/outputs) determine the final fate. As a result, progenitors born with common intrinsic factors can diverge into different subtypes (Mukhopadhyay *et al.* 2009; Ishino *et al.* 2017; Quattrocolo *et al.* 2017; Wamsley & Fishell 2017) (Figure 1.3).

Figure 1.3

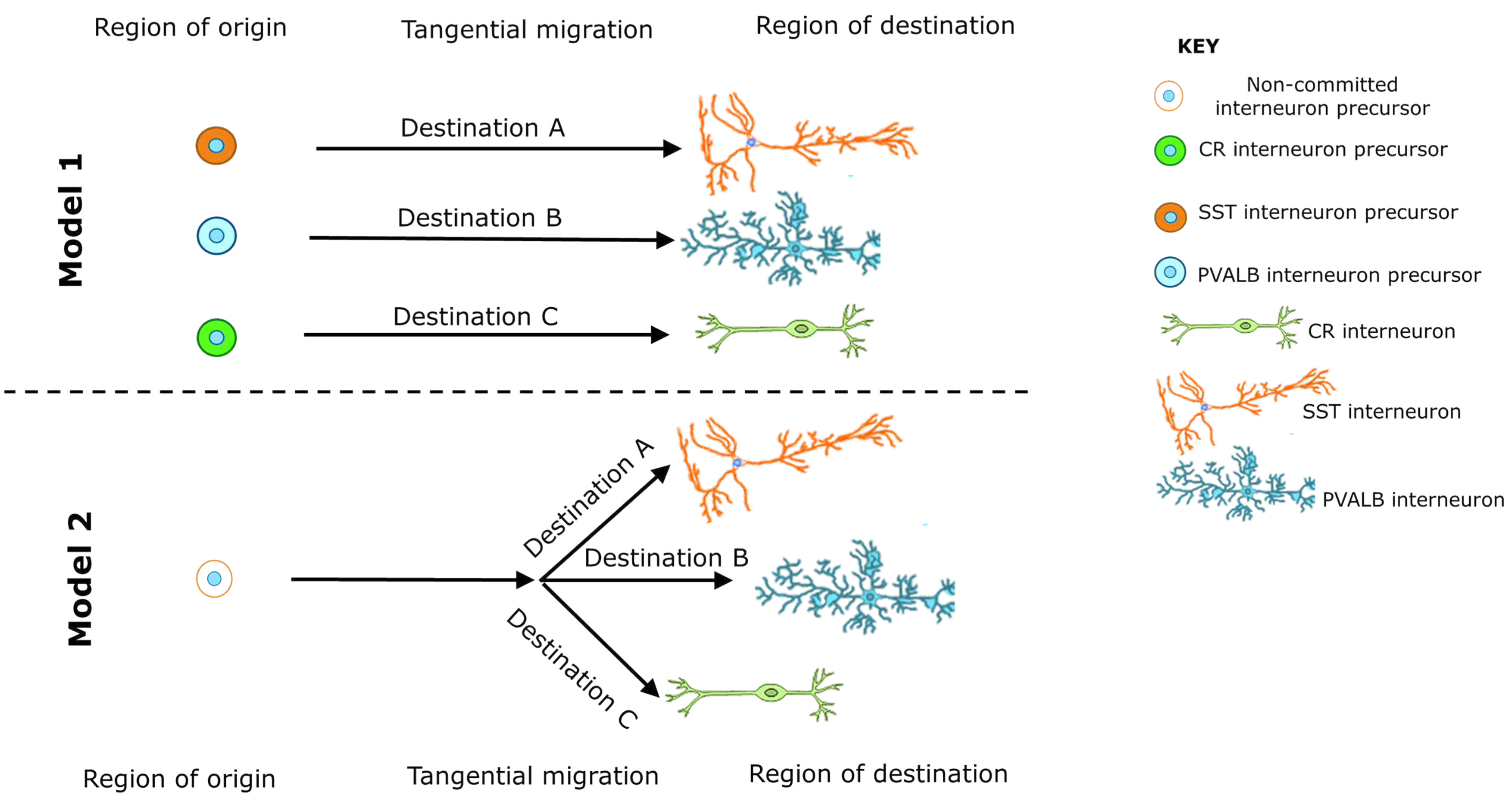

Region of origin
Tangential migration
Region of destination
Model 1
Destination A
Destination B
Destination C
Model 2
Destination A
Destination B
Destination C
Region of origin
Tangential migration
Region of destination
KEY
Non-committed interneuron precursor
CR interneuron precursor
SST interneuron precursor
PVALB interneuron precursor
CR interneuron
SST interneuron
PVALB interneuron

Figure 1.3 Two alternative interneuron lineage commitment models.

Interneurons are born within two neighbouring regions, the CGE and MGE. Interneurons are then dispersed across multiple region of destinations (e.g. the cerebral cortex, striatum, hippocampus) through tangential migration, and integrate to form neuronal circuits. **Model 1** proposes that the subtype the MGE/CGE progenitor will ultimately become is determined through specific intrinsic instructions, dictated by the location and time of birth within the region of origin. Therefore, within the MGE/CGE there are distinct progenitor pools with differential lineage commitments. **Model 2** considers the influences that the region of destination has on interneuron fate. This second model proposes that interneuron precursors from the same progenitor pool maintain the potential to become one of many interneuron subtypes, depending on the final location of the precursor. In this model, the lineage commitment is moulded by the specific cues within the region of destination, such as the region-dependent extracellular matrices and inputs from region-specific principle cells (activity-dependent mechanism). Cartoon illustrations of neurons were adapted from figures taken from https://www.motifolio.com/neuroscience-all.html (Motifolio 2018).

Somatostatin, SST. Parvalbumin, PV. Calretinin, CR. MGE, medial ganglionic eminence. CGE, caudal ganglionic eminence.

1.7 Single-cell transcriptome analysis

1.7.1 Principle

To better understand interneuron diversification a method is required that allows one to characterise the molecular heterogeneity of the population, and the transcriptome dynamics of the population over time. However, when using bulk-level analysis to explore heterogeneity in a population, there is the risk that some differences may be masked due to a similar proportion of the population upregulating and downregulating the same gene (Wills *et al.* 2013). Moreover, bulk-level transcriptomics prevents one from distinguishing between a transcript that is weakly detected due to low expression in a large subpopulation or high expression in a rare subpopulation (Ofengeim *et al.* 2017). Single-cell transcriptomics provides an ideal method to tease apart and define the molecular heterogeneity, by detecting between-cell differences in gene expression, and using this to assign cells with similar expression profiles to discrete subpopulations (Liu & Trapnell 2016).

1.7.2 Methodology

One of the major challenges overcome by single-cell methodologies was how to isolate viable single cells. Ideally, the method would be of high throughput, low labour, low cost and timely. Unfortunately, no single technique has been able to encompass all these points, and as a result, each methodology must compromise on one or more of the optimal properties (Moignard & Göttgens 2016; Poulin *et al.* 2016).

One of the simplest approaches to isolate single cells is manual picking (Figure 1.4). Although this method is by far the most time consuming and low throughput of all the techniques, there are advantages to this approach (Poulin *et al.* 2016). Manual picking is of low cost, with no additional expense for specialised capture machinery. Moreover, as the cell capture is set by the experimenter, a selection criterion can be easily introduced.

Higher throughput and less labour-intensive techniques include the microfluidic and microdroplet systems, Fluidigm C1 and Chromium 10x Genomics, respectively (Poulin *et al.* 2016) (Figure 1.4). The 10x Genomics chip works by step-by-step mixing of the single-cell suspension, barcoded micro-beads and oil to form single-cell tagged beads isolated within oil droplets (gel beads in emulsion, GEM) (Zheng *et al.* 2017). Due to the much higher ratio of gel beads loaded onto the chip relative to single cells, the system should tag a single cell per GEM, uniquely barcoding the cell for later processing (Zheng *et al.* 2017). However, this single cell capture per GEM cannot be visually checked, thus making optimisation difficult. Whereas, Fluidigm C1 microfluidic chip, termed the integrated fluidic circuit (IFC), contains isolated capture sites (butterflies) that can be viewed before and after cell loading (Pollen *et al.* 2014). The cells are directed into empty butterflies using a pressured valve system (smallest IFC, 96 butterflies to largest IFC, 800 butterflies). Thus, the chip design results in Fluidigm C1 system having a lower throughput and higher cost per cell relative to other techniques (Poulin *et al.* 2016). Nevertheless, the ability to visually checked capture efficiency and state of the captured cells (e.g. viable, meets any selection criteria) before lysing is an important quality control check (Pollen *et al.* 2014). This manual check holds value, because first, this prevents any erroneous conclusions being made from poor quality/doublet cDNA, and second, reduces downstream costs as no further reagents are wasted on unusable/undesired samples (Poulin *et al.* 2016).

Once the single-cells have been isolated there is then the choice of reaction chemistry, which includes (but not limited to): SMART-seq, CEL-seq and Tang-seq (Tang *et al.* 2009; Hashimshony *et al.* 2012; Ramsköld *et al.* 2012; Ziegenhain *et al.* 2017). Along with differences in cost, these reaction chemistries differ in the method of reverse transcription (template-switching or second-strand synthesis) and preamplification (in vitro transcription and polymerase chain reaction, PCR) (Tang *et al.* 2009; Hashimshony *et al.* 2012; Ramsköld *et al.* 2012). Ultimately this results in the reaction chemistries differing in the coverage of the mRNA, being full-length or 3′ end only, with full-length coverage providing greater sensitivity and higher read depth option (Poulin *et al.* 2016). Reaction chemistries performing full-length mRNA coverage (SMART-seq) are therefore particularly desirable for detecting low-abundance transcripts (Ziegenhain *et al.* 2017).

Figure 1.4

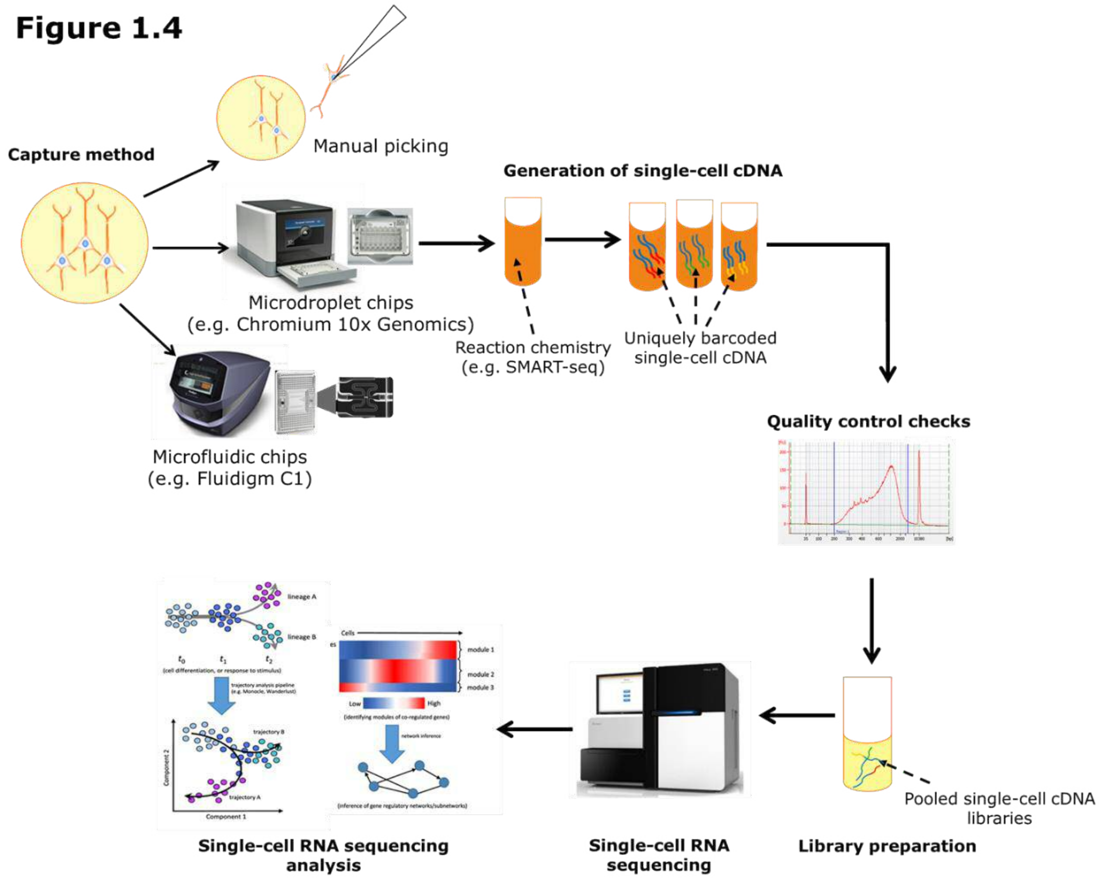

Capture method
Manual picking
Microdroplet chips
(e.g. Chromium 10x Genomics)
Microfluidic chips
(e.g. Fluidigm C1)
Generation of single-cell cDNA
Reaction chemistry
(e.g. SMART-seq)
Uniquely barcoded
single-cell cDNA
Quality control checks
Library preparation
Pooled single-cell cDNA
libraries
Single-cell RNA
sequencing
lineage A
lineage B
t0
t1
t2
(cell differentiation, or response to stimulus)
trajectory analysis pipeline
(e.g. Monocle, Wanderlust)
trajectory B
trajectory A
Component 2
Component 1
Cells
module 1
module 2
module 3
Low
High
(identifying modules of co-regulated genes)
network inference
(inference of gene regulatory networks/subnetworks)
Single-cell RNA sequencing
analysis

Figure 1.4 Step-by-step guide to single-cell RNA sequencing methodology.

Although there are distinctions between single-cell RNA sequencing methodologies, this can mainly be distilled down to two main differences: capture method and reaction chemistry. Otherwise, the basic steps of single-cell RNA sequencing are common across protocols. These steps required the generation of single-cell cDNA that is barcoded to allow the cDNA corresponding to each cell to be uniquely identified. The single-cell cDNA is then carried through certain quality control checks to determine whether the quality is high enough for single cell RNA sequencing. These steps include Bioanalyzer and Qubit quantification, along with single-cell QPCR analysis. On passing these quality control checks the cDNA can be carried forward for library preparation, and then sequencing of the pooled single-cell libraries on the Illumina platform. For more details see Chapter 2. The single-cell RNA sequencing data obtained provides an in-depth, unbiased characterisation of the genes expressed by each cell, which led to an understanding of the heterogeneity within/across populations and the genes responsible. Figures adapted from Liu & Trapnell 2016.

1.7.3 Applications

Single-cell RNA sequencing analysis has provided valuable information concerning interneuron subtype diversity in the adult brain in both humans and mice (Darmanis *et al.* 2015; Lake *et al.* 2016; Boldog *et al.* 2018). The cross-comparisons between species has revealed species-dependent differences in interneuron subtype number and identity (Darmanis *et al.* 2015; Bakken *et al.* 2016; Sousa *et al.* 2017). A classic example is the differences in the predominant interneuron subtype of a region, with cortical PVALB and CR interneurons being the most abundant subtype in mouse and human cortex respectively (Darmanis *et al.* 2015; Bakken *et al.* 2016; Sousa *et al.* 2017). Moreover, human-specific, novel interneuron subtypes have been detected using single-cell RNA sequencing of adult human tissue (Boldog *et al.* 2018). One of these subtypes included the irregular spiking cholecystokinin (CCK) subtype (co-expressing *COUPTF II,* complexin 3 *CPLX3, SV2C*), dubbed "rosehip" cortical interneurons. This specific subtype was rare in the human cortex (making up less than 6% of COUPTF II positive interneuron population) and restricted to layer I of the cerebral cortex, which is in agreement to previous work characterising COUPTF II positive interneuron population in human cortex (Varga *et al.* 2015; Boldog *et al.* 2018).

Moreover, longitudinal single-cell transcriptome analysis of the mouse MGE and CGE has identified age-dependent differentially expressed genes, which has provided novel insight into the developmental trajectory of interneurons (Chen *et al.* 2017; Close *et al.* 2017; Mi *et al.* 2018). Of significance was the ability to detect novel candidate lineage determining transcription factors (Mayer *et al.* 2018; Mi *et al.* 2018). However, such analysis is still awaited in humans.

Finally, single-cell transcriptomic analysis can identify co-expression patterns of genes that represent putative gene regulatory networks/nodes that could be responsible for a particular phenotype or directing progenitors along a specific lineage pathway (Nowakowski *et al.* 2017; Tripathy *et al.* 2017; Mi *et al.* 2018; Mayer *et al.* 2019). Such analysis can be useful in understanding disease mechanisms, by identifying disrupted gene regulatory networks and thereby naming candidate therapeutic targets and biomarkers (Ofengeim *et al.* 2017; Zhang *et al.* 2019).

1.8 The connectivity of neurons

The architecture of neuronal circuits is finely tuned in an activity-dependent manner. Therefore, the connectivity of a neuron, in terms of the identity and the number of synaptic partners, reflects the function of the neuron. However, mapping neuronal connectivity is a challenging feat. Since the 1970s, neuronal tracing tools have been developed, becoming more sophisticated over time, both in terms of sensitivity and versatility (Lanciego & Wouterlood 2011).

1.8.1 How to measure neuronal connectivity

Early neuronal tracers were chemicals or proteins capable of selectively marking neuronal circuits either directly (tracer with a fluorescent tag) or indirectly (immunohistochemistry staining for the tracer). These early tracers included Fluoro-Gold and proteins, such as cholera toxin subunit B and leucoagglutinin, which showed low toxicity, high stability and were selectively taken up by neurons (Lanciego & Wouterlood 2011). However, early tracers often suffered from weak signals, especially over long distance synaptic transmission, and lacked the ability to be re-defined to the experimenter's requirements (Lanciego & Wouterlood 2011; Ginger *et al.* 2013). This led to the generation of viral tracers, a highly versatile tool to measure neuronal connectivity (Wickersham *et al.* 2007; Osakada *et al.* 2011; Osakada & Callaway 2013).

Over the last decade, viral tracers have been used to map neuronal connectivity both in vivo and in vitro (Wickersham *et al.* 2007; Ginger *et al.* 2013; Wall *et al.* 2013; Grealish *et al.* 2015; Wertz *et al.* 2015; Tian *et al.* 2016; Doerr *et al.* 2017; Tornero *et al.* 2017; Sarkar *et al.* 2018). The system exploits the ability of specific viruses to selectively infect neurons, and spread through their synaptic connections. A mechanism termed polysynaptic transmission (Ginger *et al.* 2013). The native wild-type rabies virus is an example of such a virus, which is transmitted through retrograde polysynaptic transmission (Osakada *et al.* 2011). In addition to inducing the expression of viral genes required for self-replication, the viral tracer enforces the expression of reporters (e.g. mCherry), and in some cases additional genes to modify/monitor neuron function (e.g. calcium sensors or light-activated ion channels/proton pumps) (Wall *et al.* 2013; Grealish *et al.* 2015; Wertz *et al.* 2015; Tian *et al.* 2016; Doerr *et al.* 2017; Tornero *et al.* 2017; Sarkar *et al.* 2018). As a result, synaptic partners can be identified as well as manipulated.

1.8.2 Principles of the monosynaptic rabies viral tracing system

Refinements to the rabies viral tracing system have been made over time, leading to the generation of monosynaptic rabies viral tracers that provide a more discrete tracing system (Wickersham *et al.* 2007; Osakada & Callaway 2013). This was achieved by generating a non-self replicating rabies viral tracer through the deletion of one of the five genes needed for rabies virus replication: the rabies glycoprotein (Wickersham *et al.* 2007; Osakada & Callaway 2013). The recombinant glycoprotein deficient (ΔG) rabies tracer must obtain the glycoprotein from an exogenous source, the host neuron, in order to make new viable ΔG rabies tracers, and amplify the tracer signal. This process has been named glycoprotein trans-complementation (Osakada & Callaway 2013).

Moreover, the viral tracer uptake can be limited to the neuronal class of choice by changing the tropism of the viral tracer (Osakada & Callaway 2013). This has been achieved through a step called pseudotyping (Osakada & Callaway 2013). Natively the rabies virus has the ability to enter any mammalian neuron using an envelope protein that is complementary to proteins on the cell surface membrane of mammalian neurons. However, by changing the endogenous envelope protein to that of a non-mammalian targeting virus, such as EnvA of the avian sarcoma leukosis virus, the viral tracer can no longer enter mammalian cells (Osakada & Callaway 2013). The neuronal class of choice is then genetically modified to express the complementary protein to EnvA, the avian cell surface protein, TVA (Osakada & Callaway 2013). This restricts uptake of the viral tracer to this TVA expressing neuronal class, thus limiting analysis to which synaptic connections are made on to this specific neuronal population.

In summary, by combining both the glycoprotein deficiency and pseudotyping of the viral tracer, to generate EnvA pseudotyped ΔG rabies tracer, the entry and amplification of the tracer can be restricted to TVA and rabies glycoprotein expressing neurons, which are termed starter neurons (Figure 1.5). To distinguish between starter neurons, rabies infected starter neurons and their presynaptic partners (termed traced neurons), the starter neurons are also modified to express a reporter distinct to the viral tracer signal (e.g. starter neuron, histone-green fluorescent protein (GFP) positive and traced neurons, mCherry positive, infected starter neuron, both GFP/mCherry positive) (Figure 1.5). Overall, the EnvA pseudotyped ΔG rabies tracer simplifies the analysis of neuronal connectivity down to which neurons have direct (monosynaptic) inputs on the starter neurons (Figure 1.5). Unlike polysynaptic tracers, this provides a connectivity pattern which is both easier and less labour intensive to interpret.

1.8.3 Applications

The mapping of neuronal connectivity is of value for a number of reasons. First, the influence a neuron has over signal integration/coding is determined by the number and type of synaptic partners (Tian *et al.* 2016). Interneurons are a classic example of a neuronal class which typically has a large number of synaptic outputs, to allow interneurons to regulate the firing pattern of multiple postsynaptic partners (e.g. cortical pyramidal neurons) in synchrony (Hirano *et al.* 2008; Sohal *et al.* 2009). This is critical for creating rhythmic patterns of signalling in the brain (Hirano *et al.* 2008; Sohal *et al.* 2009). Second, evaluating connectivity in vivo has allowed identification of the native synaptic partners of neurons, with a prime example being the neuronal inputs onto the striatal projection neuron MSN (Wall *et al.* 2013). This study identified the synaptic inputs onto D1 and D2 MSN subtypes to understand the shared and differential connectivity of each subtype (Wall *et al.* 2013). This has provided clarity on circuit integration across brain regions, and in this case, gave a refined and enhanced understanding of the basal ganglia circuitry (Wall *et al.* 2013). Third, a cross-comparison of neuronal connectivity in control and disease models have implicated aberrant connectivity of neurons in many neurological disorders, including schizophrenia, intellectual disability and Rett syndrome (Brennand *et al.* 2011; Marín 2012; Sun *et al.* 2019). This has been supported by many of the known risk genes, for neuropsychiatric disorders, encoding for synaptic proteins, including cell-adhesion molecules (neuroligins *NLGN*, neuregulin *NRG,* Erb-B2 receptor tyrosine kinase 4, *ErbB4*) and scaffold proteins (SH3 and multiple ankyrin repeat domains gene family *SHANK,* disrupted in schizophrenia 1 *DISC1*) (Südhof 2008). Together, this has enhanced our understanding of disease mechanism, and may enable novel, and more efficacious therapeutic targets to be generated. Finally, cell replacement therapy is a promising therapeutic approach, which relies on the successful and appropriate integration of the neuronal graft in atrophied brain regions to repair and replace the degenerated circuits (Aubry *et al.* 2008; Cunningham *et al.* 2014; Grealish *et al.* 2015). In rodent models, a measure of the outputs/inputs onto the neuronal graft would allow an accurate evaluation of the efficacy of the graft and could guide improvements necessary to progress the development of this therapeutic strategy (Grealish *et al.* 2015; Tornero *et al.* 2017).

Figure 1.5

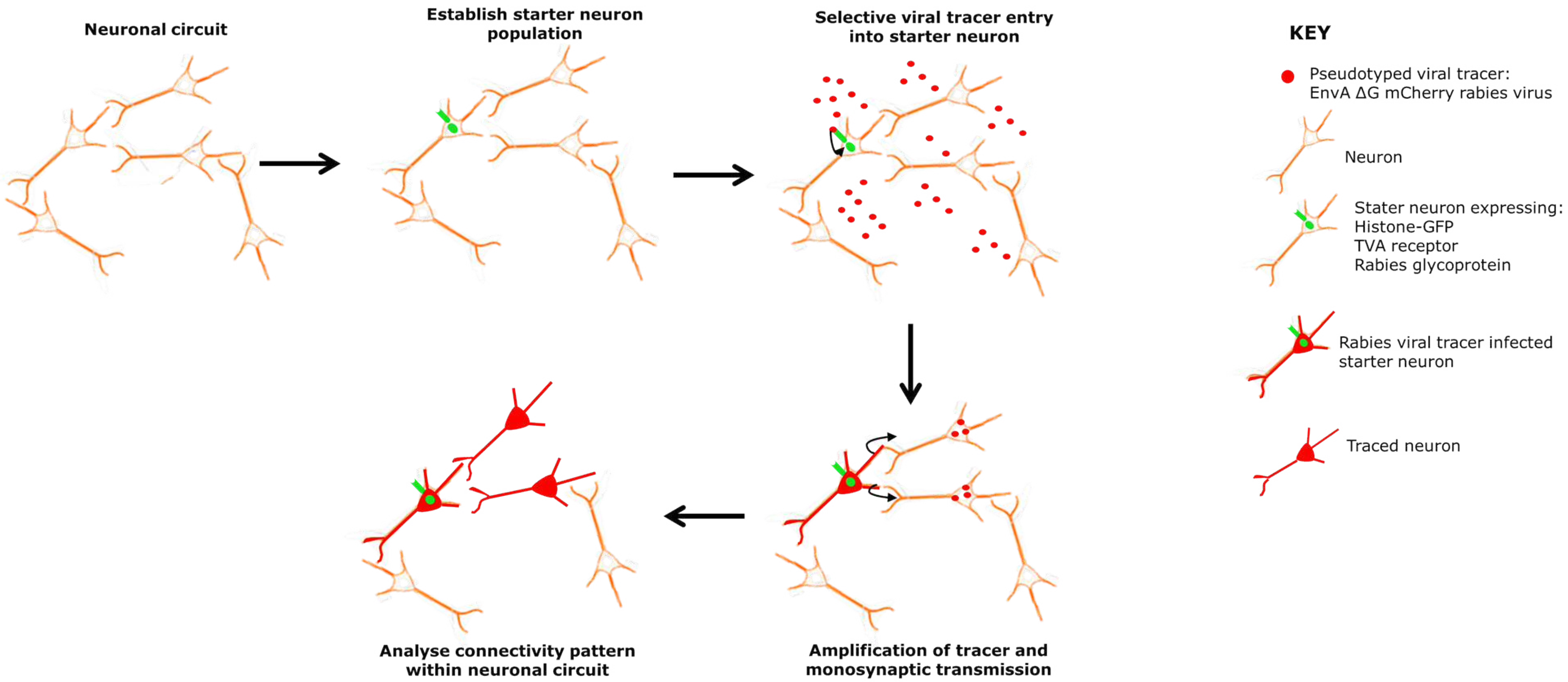

Neuronal circuit
Establish starter neuron population
Selective viral tracer entry into starter neuron
Amplification of tracer and monosynaptic transmission
Analyse connectivity pattern within neuronal circuit
KEY
Pseudotyped viral tracer: EnvA ΔG mCherry rabies virus
Neuron
Stater neuron expressing: Histone-GFP TVA receptor Rabies glycoprotein
Rabies viral tracer infected starter neuron
Traced neuron

Figure 1.5 Schematic of the retrograde monosynaptic viral tracer system: step-by-step guide.

EnvA pseudotyped ΔG rabies viral tracer acts to selectively map discrete neuronal circuits made up of the starter neuron and the monosynaptic presynaptic partners of this starter neuron. The viral tracer system exploits the innate and selective ability of the wild-type rabies virus to spread trans-synaptically from one neuron into the neurons connected presynaptically onto this starter neuron (retrograde transmission). By making the following complementary modifications to the rabies virus and the selected starter neuron population, the monosynaptic viral tracer system can be established. First, restrict entry of the viral tracer through pseudotyping the virus with non-mammalian envelope protein (e.g. EnvA) and enforce the expression of the complementary receptor (e.g. TVA) in starter neurons selectively. Second, restrict viral tracer replication, and therefore the transmission of the tracer signal, by deleting a critical rabies gene from the rabies tracer genetic material (e.g. the glycoprotein), and only provide starter neurons with the ability to express the missing rabies glycoprotein. Therefore, once the pseudotyped ΔG rabies viral tracer has been trans-synaptically transmitted, the tracer signal is terminated (as no glycoprotein provided to allow signal to be amplified). Finally, to allow independent identification of the starter neurons and traced neurons, have starter neurons innately expressing one reporter (e.g. histone-GFP) and the rabies tracer enforce the expression of a different reporter (e.g. mCherry). Any viral tracer infected starter neurons are therefore nuclear GFP and mCherry double positive. Whereas, neurons that are either post-synaptic to starter neuron or not monosynaptically connected to the starter neuron, do not take up EnvA pseudotyped ΔG rabies viral tracer (remain untraced, GPF and mCherry negative). Cartoon illustrations of neurons were adapted from figures taken from https://www.motifolio.com/neuroscience-all.html (Motifolio 2018).

Glycoprotein-deficient, ΔG. GFP, green fluorescent protein.

1.9 Research models

Rodent models have been a valuable resource in research, due to the ability to directly link genetics, output of defined neuronal circuits, and behaviour. However, the use of rodent models as a means to understand human interneuron development and function in the brain also raises many caveats. These limitations mainly relate to between-species differences in interneuron numbers, spectrum of subtype diversity and subtype composition in each brain region (Wu & Parent 2000; Bakken *et al.* 2016; Sousa *et al.* 2017; Boldog *et al.* 2018). The apparent inter-species differences have highlighted the importance of conducting experiments directly in humans. This work has included molecular and functional characterisation of primary human tissue, both post-mortem adult and foetal tissues (Mo *et al.* 2007; Moore *et al.* 2009; Boldog *et al.* 2018; Garas *et al.* 2018; Zhong *et al.* 2018). Unfortunately, primary human tissue use presents many challenges and restrictions, including, but not limited to: the scarcity of tissue, the inconsistency of tissue quality, and legal age restrictions of human foetal tissue usage. To overcome these limitations, several independent groups have established robust means to generate interneuron-like cells from human pluripotent stem cell (hPSC) models (Goulburn *et al.* 2011; Cambray *et al.* 2012; Maroof *et al.* 2013; Nicholas *et al.* 2013; Kim *et al.* 2014; Colasante *et al.* 2015; Noakes *et al.* 2019).

1.9.1 Principles of neuron differentiation of hPSC

Over the last decade, PSC-derived neurons have emerged as a promising tool in regenerative medicine and neuroscience research. The common aim of differentiation paradigms is to recapitulate the neuronal differentiation steps that occur natively during development in order to efficiently convert PSC into the neuronal lineage of choice. The earliest success stories were obtained from mESC, with hPSC-derived neuronal models subsequently following (Bain *et al.* 1995; Okabe *et al.* 1996; Zhang *et al.* 2001; Chambers *et al.* 2009). Although the precise details of the differentiation paradigms differ, common principles are shared across neuronal differentiation paradigms (Figure 1.6).

Broadly speaking in vitro neuronal differentiation paradigms have been categorised into two different methodologies depending on the approach taken to induce neuronal fate commitment: intrinsic or extrinsic cues. Further variations within each method relate to differences in how the cells are cultured, either as 2D-monolayers adhered to the surface or free-floating 3D-spheres/embryoid bodies (and then into organoids) (Figure 1.6). Although embryoid body and organoid cultures better mimic the native form of the brain, these 3D-cultures are often confounded by issues of poor delivery of small molecule factors, nutrients and O_2 to inner layer cells. This often results in greater inconsistencies in the efficiency of the paradigm, along with higher levels of atrophy and expression of stress markers in the surviving cells (Camp *et al.* 2015; Xiang *et al.* 2017).

The basic principle of the extrinsic cue approach is to use small molecule factors to induce the expression of key transcription factors that will guide the PSC along the neuronal lineage of choice. Stage one involves directing PSC to take the first step towards a neural restricted lineage, in a process termed neural induction. A highly efficient and simple means to generate neuroepithelia was achieved using two chemical factors (SB-431542 and noggin/noggin analogue LDN-193189) in a process termed dual SMAD inhibition (Chambers *et al.* 2009; Maroof *et al.* 2013). The factors act to repress PSC taking alternative fates by inhibiting transforming growth factor-beta (TGF-β) signalling pathway. Stage two uses a region of origin dependent set of small molecular factors to restrict the fate of the neuroepithelial cells to a lineage of choice by mimicking the native environment in the culture (Figure 1.6). Alternatively, the intrinsic cue approach enforces the expression of lineage determining transcription factors and neurogenic microRNAs in hPSC using lentiviral transduction (or less efficient, electroporation) to obtain neuron-like progenitors belonging to the desired lineage (Figure 1.6) (Colasante *et al.* 2015; Sun *et al.* 2016; Yang *et al.* 2017; Yuan *et al.* 2018). The final stage of the neuron differentiation paradigm is common across methodologies, and describes the transition of progenitors to a post-mitotic state that marks the terminal differentiation of the neuron-like cell (Figure 1.6). This can be driven by neurotrophic/neurogenic factors (e.g. brain/glial derived neurotrophic factor BDNF/GDNF) (Maroof *et al.* 2013; Close *et al.* 2017; Noakes *et al.* 2019). Moreover, the application of retinol has been

Figure 1.6

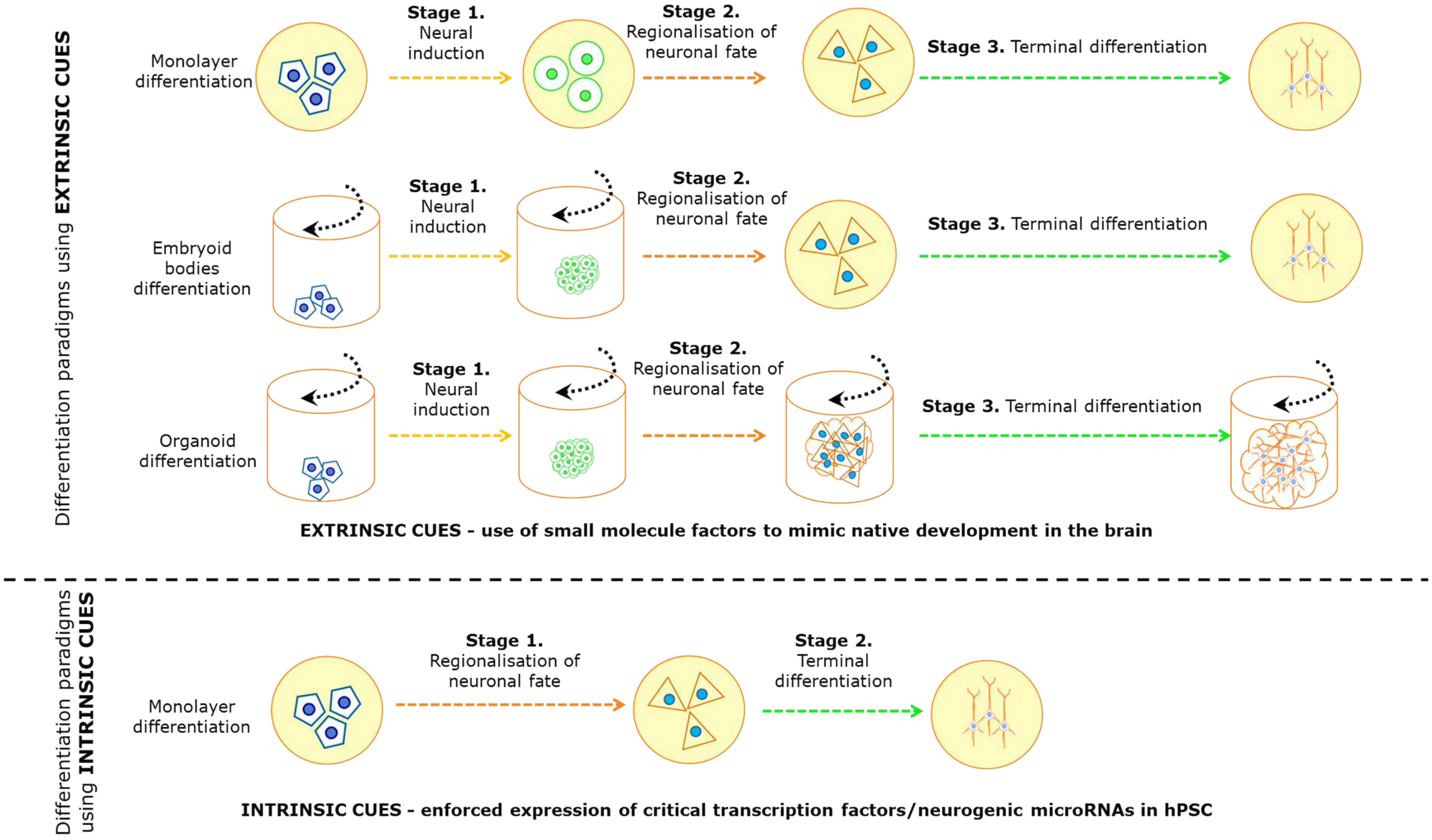

Differentiation paradigms using EXTRINSIC CUES
Monolayer differentiation
Stage 1. Neural induction
Stage 2. Regionalisation of neuronal fate
Stage 3. Terminal differentiation
Embryoid bodies differentiation
Stage 1. Neural induction
Stage 2. Regionalisation of neuronal fate
Stage 3. Terminal differentiation
Organoid differentiation
Stage 1. Neural induction
Stage 2. Regionalisation of neuronal fate
Stage 3. Terminal differentiation
EXTRINSIC CUES - use of small molecule factors to mimic native development in the brain
Differentiation paradigms using INTRINSIC CUES
Monolayer differentiation
Stage 1. Regionalisation of neuronal fate
Stage 2. Terminal differentiation
INTRINSIC CUES - enforced expression of critical transcription factors/neurogenic microRNAs in hPSC

Figure 1.6 Methodologies used to generate hPSC-derived neurons.

Although hPSC-derived neuronal differentiation paradigms have specific requirements to induce a certain neuronal lineage, there are common principles across paradigms. Broadly speaking there are two methods to direct hPSC to take on a particular neuronal fate. **Method one**, the use of extrinsic cues to mimic the cues in the native region were the neurons are born. In the case of interneurons, the MGE/CGE environmental cues are recapitulated in vitro to promote interneuron fate commitment using SHH signalling. Further variations exist within the extrinsically driven differentiation paradigms (summarised by approaches 1 to 3, top to bottom). Approach 1 describes monolayer differentiation protocols, where at all stages cells are adhered to a coated surface of the plate and grown as a 2D-monolayer. Approaches 2 and 3 both begin with generating embryoid bodies, with the culture dish spun to maintain the culture as free-floating 3D neuroepithelial spheres. On sphere formation (end of stage 1), approach 2 seeded the neuroepithelial cells and continued maturation as a 2D-monolayer. Whereas, approach 3 continued to expand the neuronal committed cells as a 3D-culture, which undergoes terminal differentiation to form an organoid. **Method two**, enforced expression of lineage determining transcription factors and neurogenic microRNAs known to be critical for directing hPSC towards the desired neuronal lineage. Classically, interneuron differentiation paradigms using intrinsic cues use lentiviral vectors to drive the expression of one or more of these genes: *LHX6, MASH1, DLX5, miR-124* and *miR-9/9*.* A final difference between extrinsic and intrinsic driven approaches, is the typically extended duration of differentiation paradigms using extrinsic cues (method one).

hPSC, human pluripotent stem cells. LHX6, LIM Homeobox 6. MASH1, achaete-scute family BHLH transcription factor 1. DLX5, Distal-less homeobox 5. miR, microRNA.

demonstrated to further promote neuronal maturation and survival in vitro (Cambray *et al.* 2012; Maroof *et al.* 2013; Noakes *et al.* 2019). Differentiation paradigms using intrinsic cues typically complete terminal differentiation over a shorter timeframe, and often have greater consistency between independently generated cultures (Colasante *et al.* 2015; Sun *et al.* 2016; Yang *et al.* 2017; Yuan *et al.* 2018). However, this method relies on well-established knowledge of the genetics underpinning the neuronal lineage. Such knowledge is not always available for a neuronal class. Moreover, direct induction of the neuronal fate poorly recapitulates the native steps in neurodevelopment, arguably reducing the value as an in vitro neurodevelopmental model.

1.9.2 Interneuron differentiation of hPSC

Differentiation paradigms to generate interneuron-like cells from PSC have been developed using both extrinsically and intrinsically driven methodologies. To efficiently generate MGE-like progenitors both SHH signalling agonists (SHH, smoothened agonist (SAG), purmorphamine) and WNT signalling inhibitors (XAV939) are used to ventralise the neuroepithelia (Maroof *et al.* 2013; Meganathan *et al.* 2017; Xiang *et al.* 2017; Noakes *et al.* 2019). As described in section 1.4.1, WNT inhibitors and SHH are critical for the spatial patterning of cells across the ventral telencephalon and have the strongest signalling within the MGE. Thus, PSC-derived differentiation paradigms with variations on the time and duration of SHH agonist application, and SHH agonist concentration exist for generating LGE, CGE and MGE-derived neurons in vitro (Danjo *et al.* 2011). Alternatively, activin signalling has been successfully used to generate both CGE-derived interneuron subtypes and LGE-derived MSN in vitro (Cambray *et al.* 2012; Arber *et al.* 2015). This paradigm relies on the activation of TGF-β signalling, and acts independently of SHH signalling (Arber *et al.* 2015). Whereas, regionalisation step in differentiation paradigms using intrinsic cues to induce MGE-fate requires ectopic expression of critical transcription factors (*LHX6*, achaete-scute family BHLH transcription factor 1 *MASH1* and distal-less homeobox *DLX 2/5*) and neurogenic microRNAs (miR-124 and miR-9/9*) (Colasante *et al.* 2015; Sun *et al.* 2016; Yang *et al.* 2017; Yuan *et al.* 2018).

Irrespective of the protocol, the ability to consistently and efficiently generate defined interneuron subtypes from PSC is lacking. Moreover, generating interneuron-like cells of the same molecular and functional standard as their native counterparts remains a major challenge within the field.

1.9.3 Application of hPSC-derived neuronal models

Although human in vitro models are limited by the inability to directly test the link of molecular/circuit-level changes to behaviour, hPSC-derived neuronal models have many advantages and applications (Kelava & Lancaster 2016). First, hPSC-derived neuronal models hold great promise as they are not confounded by species differences, a major drawback of rodent models, and are therefore arguably a more translatable model (Bakken *et al.* 2016; Garas *et al.* 2018). Second, hPSC models provide a simpler, cheaper and faster-turnover platform to investigate how specific pathways and neuronal classes can be affected by pharmacological and genetic manipulations (Anton *et al.* 2018; Shao *et al.* 2019). Hence, the malleable nature of hPSC-derived neurons has led to the wide application of in vitro models in neurodevelopment and disease modelling, and drug discovery (Kelava & Lancaster 2016). Third, hPSC-derived neuronal progenitors hold potential as a regenerative medicine strategy. This was marked by the first human clinical trials using hPSC-derived dopaminergic neuronal progenitors as the graft source in Parkinson's disease launching in 2018 (Barker *et al.* 2017). The results are eagerly awaited. Although previous findings in neurodegenerative rodent models would add caution, with limited and inconsistent success reported for the use of hPSC-derived neuronal progenitor grafts as a viable alternative to primary foetal tissue in cell replacement therapies (Delli Carri *et al.* 2013; Cunningham *et al.* 2014; Arber *et al.* 2015). These reports have demonstrated the need to optimise and validate in vitro models, to ensure that the neuron-like cultures closely resemble their authentic derivatives (in primary human tissue), both molecularly and functionally (Camp *et al.* 2015; Yao *et al.* 2016; Chen *et al.* 2017; Close *et al.* 2017; Xiang *et al.* 2017).

1.10 General aims

The series of experiments presented in this thesis aimed to evaluate an in vitro model of interneuron differentiation using several innovative approaches. Particular focus was given to determining how well the differentiation of human embryonic stem cells (hESC), under an extrinsic cue based paradigm, could recapitulate the native steps of interneuron differentiation. With the limited understanding of interneuron subtype diversity that can be achieved in vitro, the culture was assessed in terms of the expression of molecular markers, physiology and neuronal connectivity. In order to investigate the heterogeneity within the culture, both the longitudinal molecular and functional characterisation was performed at single-cell resolution, using single-cell RNA sequencing and whole-cell patch clamp recordings respectively. Finally, a novel cross-comparison was performed between hESC-derived interneurons and human MGE to directly examine the authenticity of the neurodevelopmental model.

Chapter 2 - Material and Methods

2.1 Stem cell and tissue culture work

2.1.1 General cell and primary tissue culture practice

All cell/tissue culture work was carried out under sterile conditions (surfaces and equipment treated with Chemgene, STARLAB and 70% denatured ethanol, VWR), in a laminar flow hood (Maxisafe 2020, Thermo Scientific). For the majority of cultures, cultures were maintained in humidified incubators at 37°C, 5% CO_2 and 21% O_2. The only exceptions were during viral transfection and transduction (altered settings: 35°C, 3% CO_2 and 21% O_2). All viral and human tissue culture work was carried out in specified rooms isolated from other cell culture areas, with additional health and safety measures. These safety measures included: viral-suite specific lab coat, double-gloving, 24hr Chemgene soaking of any solid and liquid wastes, and autoclaving solid wastes within 24hrs of generation.

2.1.2 Cell lines

hESC line used for interneuron differentiation was H7 (female) and included: H7 parent line and H7 *LHX6::mCherry; PVALB::GFP* (dual reporter line gifted by Maria Cruz, Li lab). Other human cell lines used were human embryonic kidney cells (HEK) 293T, and HEK 293T constitutively expressing TVA (HEK TVA). Additional cell lines included the baby hamster kidney (BHK) cells constitutively expressing EnvA (BHK-EnvA cells). Lastly, BHK cells constitutively expressing (under CMV promoter) T7 RNA polymerase, rabies glycoprotein, histone-GFP, termed B7GG cells. The HEK and BHK derived cell lines were gifted from Parmar lab (Prof Malin Parmar, Lund University, Sweden).

2.1.3 Cell maintenance for hESC lines

hESC were maintained on DMEM/F-12 diluted matrigel (matrigel, Corning and DMEM/F-12, Gibco) coated wells, and cultured in E8 basal media (Gibco). Once the confluency of the cells reached approximately 80%, the cells were detached from their original well using EDTA treatment (EDTA 0.02%, Sigma). The EDTA treatment required initial washing of the cells with Ca^{2+} -free, Mg^{2+} -free Dulbecco's phosphate-buffered saline (DPBS, Gibco) to remove debris, followed by approximately 2min incubation in EDTA at 37°C. Immediately after this incubation, EDTA was removed and replaced with E8 basal media, then the cells were manually scratched from the well surface (using 2ml stripette tip guided in a zig-zag pattern across the well). The detached cells were then gently resuspended in the appropriate volume of E8 basal media. If the cells were still required, the resuspended hESC were seeded onto a freshly DPBS washed DMEM/F-12 diluted matrigel wells (typically at 1:3 split ratio, with the original well now being seeded into 3 wells). The plate was gently shaken immediately after placing in the incubator to ensure even distribution of the cells within the well. This process, termed splitting, was repeated to maintain the hESC over time.

2.1.4 Cell maintenance for HEK and BHK lines

HEK and HEK TVA cells were maintained in tissue culture flasks and cultured in "HEK media". HEK media consisted of DMEM/F-12 media with 10% foetal bovine serum (FBS, Biosera), 1% non-essential amino acids (Gibco), 1% L-glutamine (Gibco), 0.2% mycozap (Lonza). All BHK cells were also maintained in HEK media. Typically cells were seeded onto tissue culture flasks directly (no prior coating required). However, to reduce detachment during the transfection protocol (see Section 2.7.1), HEK cells were seeded onto 0.1% gelatin-DPBS coated flasks (gelatin solution, Sigma). To maintain HEK cell lines typical split ratio of 1:10 (1 flask worth of dissociated HEK cells sufficient to seed 10 flasks) was used every 2-3 days. Whereas, BHK cell lines were split at 1:15 ratio every 2-3 days. HEK and BHK cells were split using trypsin treatment (TrypLE Express, Life Technologies). The trypsin treatments began with HEK/BHK cells being washed with DPBS to remove any debris, twice. The HEK/BHK cells were then treated with diluted TrypLE (1:10 diluted in DPBS) for approximately 5min at 37°C (until cells had begun to detach on tapping the flask). The detached cells were collected from the flask, and HEK media was added to terminate trypsinisation (diluting trypsin 1:3, e.g. 5ml TrypLE, add 10ml HEK media). The cell suspension was gently centrifuged (300g, 5min, room temp) to pellet the cells. The supernatant was removed, and cells resuspended in fresh HEK media. HEK and BHK cells grew best when dissociated into single-cell/small clusters, so cells were thoroughly dissociated into the HEK media by trituration before seeding. The flask was also gently shaken once placed in incubator to ensure even distribution of cells. Lower split ratios (1:3 to 1:6) were used to obtain high confluency (80-90%) within 24hrs, a step required for transfection protocols (refer to Section 2.7.1).

2.1.5 Cell maintenance for human foetal tissue

Human foetal tissue was donated as part of the South Wales initiative for transplantation (SWIFT) programme, in accordance with the Polkinghorne and department of health guidelines. The human foetal brain was dissected out to provide MGE, LGE and CGE (dissection carried out by Rachel Hills and Sophie Rowlands, Rosser lab). The crown-rump length (CRL) was measured on the day of dissection, and the conversion of CRL to gestational age was obtained from Loughna *et al.*, and Koornstra *et al.* for CRL larger than 80mm (Koornstra *et al.* 1990; Loughna *et al.* 2009). The dissected foetal tissue was kept at 4°C in Hibernate E (Invitrogen) until collected within 24hrs. If the foetal tissue was to be maintained for a short period ex vivo (e.g. plated for immunohistochemistry), the foetal tissue was seeded onto poly-D-lycine (1:1000 diluted PDL, Sigma) coated plates, which had been repeatedly (3 times) washed with sterile water, and air dried under UV light. The foetal tissue suspension was then seeded onto the dried plates in the following media: DMEM/F-12 with 1% FBS, 2% B27 with retinol (Invitrogen) and 0.2% mycozap. This media was referred to as "foetal tissue media". The culture was then maintained for 2-3 days. After this period cells were fixed with 4% paraformaldehyde (PFA, paraformaldehyde powder dissolved in PBS, Sigma) ready to perform immunohistochemistry (see Section 2.2).

The seeding of the foetal tissue required a single cell suspension to be made. To obtain a single-cell suspension of foetal tissue, TrypLE Select CTS (Life Technologies) treatment was used. The steps were as follows. First, Hibernate E was removed from the foetal tissue, and replaced with prewarmed (37°C) DMEM/F-12 and 20µl/ml Donase-α (also known as Pulmozyme), with the fresh media changed several times to remove debris. Second, DPBS was used to remove traces of DMEM/F-12/Donase-α before adding TrypLE Select CTS with Donase-ɑ (20µl/ml). The foetal tissue was incubated at 37°C for 10min (the fresh tube gently tapped at the 5min interval) in TrypLE Select CTS/Donase-ɑ mix. Once the incubation was complete, the trypsinisation was terminated by adding prewarmed foetal tissue media. The tissue was then manually dissociated into single cells through manual trituration in the foetal tissue media. The cell concentration (cells/ml) was calculated by manual cell counting of viable cells with Neubauer Chamber and Trypan Blue assay as a cell viability measure (Trypan Blue, Gibco). Tissues with viability of <80% were discarded. Approximately 50,000 cells were seeded into each well of a 24-well plate in the form of a single drop of 30µl cell suspension. The well was then flooded with 500µl foetal tissue media between 4-6hrs after seeding.

2.1.6 Interneuron differentiation of hESC

Li lab has an established protocol to generate hPSC derived interneuron-like cells using extrinsic cues (Noakes *et al.* 2019, and Appendix folder 8). The differentiation paradigm consists of three main stages: neural induction, commitment to interneuron fate, and terminal differentiation. An illustration of these steps has been provided in Chapter 3 and 6 (without and with SCM1 factors, respectively).

Stage one began with seeding the hESC onto freshly DPBS washed DMEM/F-12 diluted reduced growth factor matrigel plates (diluted 1:15, Corning). Once cell confluency reached >90% the E8 basal media was replaced with “neural induction media”, which consisted of a basal media and the following small molecule factors: SB431542 (10µM, Tocris), LDN193189 (100nM, StemGent) and WNT inhibitor XAV939 (2µM, Tocris). The basal media, termed N2B27 (retinol-free) media, consisted of: 75% DMEM/F-12, 25% Neurobasal (Gibco), 0.7% N2 supplement (Invitrogen), 0.7% B27 retinol free supplement (Invitrogen), 1% L-Glutamine (Gibco), 0.2% Mycozap and 0.1% β-mercaptoethanol (Gibco). This change in media marked day 0 of the differentiation protocol, and cells were maintained in neural induction media for approximately 9 days with half-volume media changes every 2 days.

Once stage one was completed, the ventralised neuroepithelial cells were split using EDTA treatment as described above (refer to Section 2.1.3), but with an additional step of having the cells exposed to 10μM ROCK inhibitor at least 1hr before splitting. The cells were seeded to obtain 90% confluency onto freshly DPBS washed fibronectin (15 μg/ml, Millipore) coated wells. This marked the start of stage two of the interneuron differentiation paradigm, when SHH signalling is used to generate interneuron-like progenitors. The SHH treatment began on day 10 and typically lasted until day 20, and consisted of 200 ng/ml human recombinant SHH (SHH, R&D) and 1 μM Smo agonist purmorphamine (Calbiochem CAS 483367-10-8, Millipore) added to N2B27 (retinol-free) media. This is referred to as the "interneuron differentiation media".

Once the interneuron-like progenitors had become overly confluent, at approximately day 20, a second split was performed using accutase treatment onto PDL/laminin plates (laminin, Sigma). To gently detach the progenitors, cells were treated with accutase (Gibco) for a 10min incubation at 37°C. To isolate the detached progenitors from the accutase, accutase-cell suspension was diluted (1:5) with N2B27 (retinol-free) media with 10μM ROCK inhibitor, and centrifuged (300g, 5min, room temp) to pellet the detached cells. The supernatant was then aspirated and the cells resuspended in N2B27 (retinol-free) media with 10μM ROCK inhibitor and 10ng/ml BDNF (PeproTech). Once the appropriate cell density was reached (typically 300,000 cells/ml), 0.5ml of cell suspension media was seeded into each well of the 24-well plate. The plates were gently shaken on placing into the incubator to ensure even distribution. This marked the start of stage three, terminal differentiation, which lasted until day 65 day in vitro. Gradually, through half-volume media changes every 3-4 days, the N2B27 (retinol-free) media was replaced with N2B27 with retinol (substitution of B27 retinol free to B27 with retinol, Invitrogen), and BDNF (10ng/ml). This is described as the "neuronal maturation media". To promote functional maturation of the interneuron-like progenitors (see Chapter 6), additional factors were added for approximately 5 to 7 days shortly after the second split (typically from day 22 to day 29). These factors were described as SCM1 factors, and included: BDNF (10ng/ml, PeproTech), DAPT (10uM, Sigma), ascorbic acid

(200 µM, Sigma), GABA (300 µM, Tocris), 0.5mM calcium chloride (Sigma), CHIR99021 (3 µM, Tocris), LM22A4 (1 µM, Tocris), Forskolin (10 µM, Tocris), PD0332991 (2 µM, Selleckchem) (Telezhkin *et al.* 2016). After the 5-7 day treatment with SCM1 factors, the added factors were reduced to BDNF (10ng/ml) and ascorbic acid (200 µM). In any experiments were SCM1 factors were used in the differentiation protocol, this has been clearly stated (see Chapter 6).

2.1.7 Sequential serial expansion of hESC-derived interneuron progenitors

A modified design of the differentiation paradigm was implemented when performing sequential serial expansions of hESC-derived interneuron progenitors (Chapter 3). Stage one (neural induction) and stage two (interneuron fate commitment) were completed as previously described in Section 2.1.6. On completing stage two (at approximately day 20) a proportion of the hESC-derived interneuron progenitors were carried forward into stage three, for terminal differentiation. The remaining day 20 progenitors were re-seeded into stage two, under SHH treatment for another 10 days (until day 30). The seeding was carried out to achieve high density (approximately 1.5 million cells per well on 12-well plate, 90% confluency), to encourage the maintenance of the proliferative progenitors, and the expansion of the progenitor population. This process was repeated 5 times to generate 6 sequentially derived cultures made from serial expansions, with the last culture seeded at day 70 (serial expansion 6, SE6). Although the final fix date was the same across sequentially generated cultures (day 110), the duration each culture had in the neuronal maturation media progressively decreased in each subsequent culture from 90 days in culture 1 (SE1 + 90 days) to 40 days in culture 6 (SE6 + 40 days). Samples were collected for immunohistochemistry (Section 2.2), either 1 day after serial expansion (SE + 1 days) onto fibronectin plates in SHH treatment (progenitor plate immunostaining) or at day 110 (terminally differentiated plate immunostaining). An illustration of the steps described above has been provided in Chapter 3.

2.2 Immunohistochemistry

Immunohistochemistry was carried out on cells fixed using a PFA treatment. The following protocol was used to fix cells. First, cells adhered to the surface of a 24-well plate were washed twice with chilled (4 °C) DPBS. Second, cells were left in chilled 4% PFA for 15mins at room temperature. The PFA was then collected, and disposed of appropriately, and replaced with chilled DPBS, and stored at 4°C. The fixed cells were stored for only a short period (no longer than 2 weeks) before immunostaining. To stain the cells for selected protein markers, the cells were washed with PBS (sterile water dissolved PBS tablets, Sigma), then treated with block solution. Block solution consisted of PBS, 0.02g/ml Bovine Serum Albumin (BSA, Sigma) and 5% donkey serum (Sigma). If nuclear markers needed to be identified, an additional methanol wash step, which involved sequential washes in methanol with dilutions gradually increasing in concentrations (to 100% methanol) and then gradually decreasing back to methanol-free wash, before the blocking step. Moreover, PBST (PBS with 0.1% Triton-X 100, Sigma) was used instead of PBS in wash steps. Once the blocking step was completed (approximately 1hr blocking, room temp, plate on shaker), primary antibodies diluted in block solution replaced the block solution (for list of primary antibodies see Appendix folder 1 table 1). Primary antibodies were left overnight, with the plate on a gentle shaker at 4°C. The next day, primary antibody-block solution was removed and cells washed with PBS (or PBST) for 20min intervals, repeated 3 times (room temp). The secondary antibodies were prepared in PBS (or PBST) at 1:1000 dilution, and added to the cells for approximately 1hr at room temperature while on a gentle shaker (for list of secondary antibodies see Appendix folder 1 table 2). From this point onwards, the plates were kept in the dark using tin foil covering. Once incubation was completed, cells were exposed to nuclei marker DAPI (Molecular Probes) for 5min (DAPI stock 1mg/ml, 1:3000 dilution in PBS). Three 10min washes in PBS then followed to remove unbound antibodies and reduce background staining. Finally, coverslips were placed over the stained cells using fluorescence mounting media (Dako) to mount the coverslip. If cells were cultured on coverslips (applicable for electrophysiology, Section 2.11), the coverslips were removed from the plate and moved onto a microscope glass slide with coverslip facing cell side down, and adhered with Dako. Plates and glass slides remained at 4°C, in tin foil, for preservation.

2.3 Stages prior to quantitative polymerase chain reaction

Throughout all steps, benches and pipettes used were sprayed first with 70% ethanol and then RNase Zap (RNase Decontamination Solution, Invitrogen). Moreover, sterile filtered tips and RNase-free sterile tubes were used.

2.3.1 RNA extraction

TRIzol method was used to extract RNA from tissue/cell cultures. Details of the protocol in full can be found in the manufacture's guide: TRI Reagent Protocol (Sigma-Aldrich 2019). The samples were first washed twice with DPBS to remove any debris, then TRIzol (Invitrogen) was applied to generate a lysed, homogeneous TRIzoled sample. The TRIzoled samples were stored immediately at -80°C. On the day of RNA extraction, the samples were removed from -80°C stores and thawed on ice, and then left to reach room temp for approximately 5min. The appropriate volume of chloroform was added to the TRIzol samples, then mixed by vigorously shaking, and left to separate into three layers, which was then further promoted by centrifugation (12,000g, 15min, 4 °C). The colourless upper aqueous phase was collected into a fresh tube, and treated with the appropriate volume of isopropanol and shaken (isopropanol treatment 10min, room temp). The samples were centrifuged to pellet the RNA (RNA precipitation 12,000g, 10min, 4 °C). The supernatant was then removed, and the RNA was washed with freshly-made cold 75% absolute ethanol. All traces of ethanol was removed by centrifuging the sample at 12,000g, 5min, 4 °C. A second short spin (<10s) was then repeated once the ethanol supernatant had been removed to collect any remaining traces of ethanol. Once all traces of the ethanol was removed, RNA pellet was left to air dry for approximately 20min (until RNA pellet appeared matt). Finally, RNA pellet was resuspended in approximately 20-30µl PCR-grade water. To help promote RNA resuspending in water, samples were held on a heat block at 65°C for 15min. Using NanoDrop 2000 Spectrophotometer (Thermo Scientific), the A260/280 and concentration of the RNA samples were determined (pre-DNase treatment measure).

2.3.2 DNase treatment of RNA

In order to remove traces of genomic DNA from the extracted RNA sample made in Section 2.3.1, DNase treatment was carried out (1 unit of DNase per µg of RNA). The DNase was added to the RNA samples for 30min incubation at 37 °C with DNase with $MgCl_2$ and 10X reaction buffer mix. The RNA samples were gently mixed at the 15min interval. To terminate the DNase treatment 50mM EDTA was added (1:10 dilution) to the samples and incubated at 65°C for 10min (gentle mixing of samples at 5min interval). Post-DNase treatment measure of the A260/280 and concentration of the RNA samples were performed on NanoDrop 2000 Spectrophotometer. RNA A260/280 should approximately be 2.0. Good quality RNA was stored at -80 °C for long term storage, the rest were discarded. Details of the protocol in full can be found in the manufacture's guide: Product Information, DNase I, RNase free (ThermoScientific 2016).

2.3.3 Reverse transcription

Once the RNA samples generated in through the steps described in Section 2.3.1- 2.3.2 had been confirmed as good quality, 1µg of RNA from each sample was carried forward for 20µl reverse transcription (RT) reaction to convert RNA to cDNA. The RT reaction used reagents from the EvoScript Universal cDNA master kit (Roche Molecular Systems). The 20µl RT reaction mix consisted of: 4µl reaction buffer mix (consisting of magnesium acetate $Mg(OAc)_2$, random primers, anchored oligo$(dT)_{18}$ and deoxynucleoside triphosphates) 2µl enzyme mix (with reverse transcriptase and protector RNase inhibitor), appropriate volume of RNA sample for 1 µg, and then PCR-grade water to bring the total final volume to 20µl. For the reverse transcriptase free (-RT) negative controls, the 2µl enzyme mix was replaced with PCR-grade water. After gentle mixing, the 20µl RT reaction mix samples were taken through RT protocol on the thermocycler. The resulting cDNA samples were stored at -20°C. Full instructions and reagent details can be found in the manufacture's guide: EvoScript Universal cDNA Master, Instructions for Use (Roche Molecular Systems 2017).

2.4 Quantitative polymerase chain reaction

Through the steps described in Section 2.3, good quality cDNA, and the corresponding -RT negative controls, have been generated ready for use in the quantitative polymerase chain reaction (QPCR) experiments. The QPCR reaction consisted of a 20µl final reaction volume made of 15µl master mix consisting of: 10µl MESAgreen (Eurogentec), 1µl 10 µM forward and reverse primers, and 4µl PCR-grade water. The remaining 5µl came from diluted cDNA stock (1:12). Two controls were included in all QPCR runs, first -RT negative control (determine genomic DNA contamination) and cDNA negative control (5µl PCR-grade water, to determine primer dimerisation contamination). The QPCR reaction protocol was then carried out on Bio-Rad CFX real-time PCR detection system with Bio-Rad CFX manager 3.1 software. Stages of the QPCR reaction protocol were as follows. Initial incubation for 4 min, 94°C, followed by the 45 times repeated cycle of 30s at 94°C, 15s at 60°C, 30s at 72°C. The final stage was the transition from 65 °C to 95 °C in 0.2 °C increments (over 0.01s). Note that the primers used were typically chosen from published scientific articles or designed through Primer-BLAST (National Center for Biotechnology Information 2019). For the full list of primers see Appendix folder 1 table 3.

Bio-Rad CFX manager 3.1 software provided the Ct value for each QPCR reaction corresponding to a particular gene, which was in turn used to calculate the delta Ct (relative to beta-actin). The relative gene expression of a gene (RE) was calculated as RE= $2^{-(\text{double delta Ct})}$ in reference to day 10 hESC-derived interneuron samples (Chapter 3).

2.5 Molecular characterisation using single-cell RNA sequencing

For full details of the individual reagents and volumes used please refer to SMART-Seq v4 Ultra Low Input RNA Kit for the Fluidigm C1 System, IFCs User Manual (Clontech Laboratories 2018).

2.5.1 Fluidigm C1 – single-cell capture and generation of single-cell cDNA

Single-cell transcriptome analysis, in the form of single-cell QPCR and RNA sequencing, was conducted on foetal tissue derived from human MGE, along with hESC-derived interneurons (Chapter 4-6). The age of foetal tissue or time point along the interneuron differentiation protocol has been specified in any results presented (Chapter 4-6). Fluidigm C1 IFC was selected as the single-cell capture methodology. IFC size was medium (optimised to capture cells sized 10-17µm) to match the approximate size of neurons and neuronal progenitors. Each C1 run had the ability to capture up to 96 cells, with capture efficiency typically between 70-80%. The reaction chemistry was SMART-Seq v4 Ultra Low Input RNA kit chemistry (Clontech). RNA spikes were added to the reaction chemistry as a positive control. On completion of the C1 run, the single-cell pre-amplified cDNA was immediately harvested for quality control checks to generate the "Harvest Plate".

To achieve a high capture efficiency required the generation of a single-cell suspension at the appropriate concentration. The single-cell suspension was made through enzymic dissociation of the cells into small clumps, followed by manual dissociation into single cells. Accutase treatment was used for hESC-derived interneuron (Section 2.1.6). Whereas, TrypLE Select CTS (Life Technologies) treatment was used to dissociate foetal tissue (Section 2.1.5). Once the single-cell suspension was ready, cold Suspension Reagent was added to the cell sample and gently mixed (1:1 ratio). The cell suspension mix was loaded onto the IFC. Once cell loading was completed, capture efficiency was determined by inspecting the IFC butterflies by eye under the inverted microscope. Any capture sites not meeting the criteria of a single viable cell were excluded from downstream applications.

2.5.2 Quality control checks

2.5.2.1 Bioanalyzer

Quality control checks began with a proportion of each single-cell cDNA sample being tested on the Bioanalyzer (Agilent Bioanalyzer with the Agilent High Sensitivity DNA Kit). The BioAnalyzer run generated electropherograms, and from these traces an approximation of quantity and average size of the cDNA fragments were determined for the raw single-cell sample (see Appendix folder 2 and 3 figures). The protocol can be found in full at: Agilent High Sensitivity DNA Kit Guide (Agilent Technologies 2017).

2.5.2.2 Picogreen assay

For downstream use in library preparation steps (Section 2.5.3), all single-cell cDNA samples obtained in Section 2.5.1 must meet the criterion of being between 0.1 – 0.3ng/µl. An accurate measure of single-cell cDNA concentration from the Harvest Plate was determined using the Picogreen assay. The Picogreen assay was run as a single 96 well microplate (flat bottom, charcoal black). The lambda standards were run in each independent assay as triplicates. As a back-ground reference, 1xTE buffer, Harvest reagent and C1 DNA dilution reagent were also run as triplicates on each plate. The rest of the plate had single-cell cDNA samples added directly from the Harvest Plate. The Picogreen master mix was added to each well. Once sufficiently mixed, the plate was run as part of the Picogreen programme on Tecan Infinite 200 Pro microplate fluorometer. The measured absorbance for each of the lambda standards was averaged across the replicates and used to generate the lambda standard concentration curve. From this standard curve, the concentration of each single cell sample was calculated. cDNA was then diluted with C1 DNA dilution reagent to achieve a final concentration within 0.1 – 0.3ng/µl range. Diluted cDNA samples were generated on a new 96 well plate described as the "Dilution Plate". Step-by-step instructions for Picogreen assay can be found at: Quant-iT PicoGreen double-stranded DNA Reagent and Kits (Invitrogen 2008).

2.5.2.3 Fluidigm Biomark - single-cell QPCR

Before any single-cell cDNA samples were carried forward for library preparation, a final quality control check was conducted. This involved using a proportion of the single-cell cDNA samples from the Dilution Plate (Section 2.5.2.2) in single-cell QPCR. The QPCR reaction was run using the Fluidigm Biomark HD and the Biomark chip (96.96 IFC). For full details please refer to the protocol titled: gene expression with the 96.96 IFC using delta gene assays on preamplified samples (Fluidigm, 2015). In brief, 96 independent single-cell cDNA samples were carried forward for use in each single-cell QPCR run, and the relative expression of 31 different genes, along with RNA spikes were determined for each sample (please refer to Appendix folder 1 table 3 for full list of primers). An assay mix was made for each primer set, along with a sample mix for each single-cell sample. Each of the 96 sample mixes were placed into each inlet on one half of the Biomark 96.96 IFC. In the remaining 96 inlets, the assay mixes were loaded as triplets for each of the 31 sets of primers, with the remaining three inlets containing assay mixes against the RNA spikes (spikes 1, 4 and 7). For single-cell QPCR analysis steps see Section 2.12.3.

2.5.3 Library preparation

2.5.3.1 Tagmentation and indexing

On successful completion of the quality control checks (Section 2.5.2), the final steps of single-cell RNA sequencing protocol were performed. The first stage began with tagmentation. A small proportion of each single-cell cDNA sample (1.25 µl) was transferred from the Dilution Plate into a fresh 96 well plate, and mixed with the tagmentation reaction chemistry. The new 96 well plate was named the "Library Prep Plate". The Library Prep Plate was incubated at 55°C for 10min, then cooled to 10°C. Immediately on reaching 10°C, NT buffer was added to each sample to neutralise the tagmentation reaction. Once tagmentation was completed, each sample on the Library Prep Plate had Nextera PCR Master Mix added, along with an unique dual combination of indexes (N7xx and S5xx, Nextera XT DNA Library Prep Index Kit, Illumina). The indexed fragments for each cell were then amplified through PCR carried out on thermocycler.

2.5.3.2 Bead clean up

With each single-cell cDNA sample now uniquely tagged and amplified (Section 2.5.3.1), the indexed libraries for cells from the same foetal tissue source/time point along differentiation were pooled together (typically 60-70 cells per pooled library). The pooled library was cleaned up using AMPure XP beads (Beckman Coulter) to remove any short fragments (likely dimerised indexes) or too long fragments (which would not sequence well on Illumina platform). The bead clean up involved first incubating the pooled library with AMPure XP beads to allow attachment (5min incubation, room temp). The 1.5ml tubes containing bead-bound library were then placed on a magnetic separation device. As the beads were magnetic, this caused the bead-bound library to become fixed to the tube surface instead of free-floating in the supernatant. The supernatant was then removed without disrupting the beads, and the bead-bound library was then washed with freshly-made 70% absolute ethanol to remove too long/short cDNA fragments, followed by air drying until the bead-bound library appeared matt (5-10min). The tubes were removed from the magnetic stand, and the bead-bound library was eluted using C1 DNA dilution reagent. The bead clean up was repeated twice, and the final bead-purified library was eluted in 50μl C1 DNA dilution reagent.

2.5.3.3 Qubit assay

For the bead-purified library to be run on the Illumina platform the concentration of the library must be 4nM. In order to determine the concentration of the post-clean up pooled libraries, the following information was required: Qubit measure of bead-purified library concentration (given in ng/μl) and average library size (bp). The post-clean up average library size was determined by running the library in triplicates on Bioanalyzer (see Section 2.5.2.1). Qubit dsDNA HS Assay Kits (Invitrogen) was used to accurately determine the bead-purified library concentration in ng/μl, full instructions to complete the Qubit assay can be found within the Qubit double-stranded DNA HS assay kit protocol (Invitrogen 2015). In brief, Qubit dsDNA HS assay kits standards were used to calibrate the fluorometer before starting each assay set. These HS standards were diluted 10-fold in Qubit working solution. Each of the pooled libraries were diluted 100-fold in 200μl of Qubit working solution. Once the stock concentration and average library size were known, calculations were carried out to determine stock concentration of the pooled library in nM. Dilutions of stock pooled libraries were made to get the library to 4nM (within +/- 10%). All sequencing was carried out on HiSeq 4000 Illumina platform with the following parameters: paired-end reads, read length 2 x 75bp, and read depth of approximately 1million reads per cell. The sequencing was carried out within the Wales Gene Park Genomic Facility (under Dr. Sarah Edkins and Dr Karen Reed). In Appendix folder 2 table 4 and folder 3 table 10, there is an overview of samples ran on Illumina HiSeq 4000.

For single-cell RNA sequencing analysis steps see Section 2.12.3.

2.6 Bacterial work leading to transfections

The following steps were carried out using bacteria as a vehicle to amplify the DNA plasmids required for transfection. The plasmids were used to generate the two components of the monosynaptic viral tracer system (Chapter 6). First, the lentivirus to enforce the expression of TVA, histone-GFP, rabies glycoprotein (under the human synapsin promoter) in hESC-derived interneurons, thereby generating the starter neuron population. Second, the monosynaptic rabies viral tracer. Details of all the plasmids used are provided in the Appendix folder 5 tables.

2.6.1 General practice in bacterial work

Steps described below for bacterial work were all carried out within a sterile (Chemgene and 70% ethanol cleaned) horizontal laminar flow hood. The shaker (300rpm) and non-shaker incubators used for bacterial culture were at 37°C, and were reserved for bacterial work only.

2.6.2 Bacterial transformation

To amplify the quantity of DNA plasmid stocks, bacterial transformation was carried out. Bacterial transformation was induced using the heat-shock method. Tubes of 50µl NEB 5-alpha competent E. *coli* cells were thawed on ice, and to each tube 100ng DNA plasmid was added (mixed by the gentle flicking of the tube). The tubes were left on ice for 30min. The tubes were then placed in a water bath at exactly 42°C for precisely 30s. Once heated, the tube was immediately placed back on ice for 5min. SOC medium was added to each tube to bring the mixture to 1ml (950µl SOC medium added). The tubes were incubated in the shaker incubator for 1hr to allow bacterial growth.

2.6.3 Colony selection

All plasmids, described in detail in Appendix folder 5 tables, conferred ampicillin resistance to the transformed E. *coli* bacteria, allowing colony selection based on this property. 1ml bacterial culture generated in Section 2.6.2 was streaked onto at least 10 agar- ampicillin plates (ampicillin

100µg/ml, Gibco) and left to either grow (transformed) or die (lacking plasmid) overnight (14-16hr, non-shaker incubator). The agar-ampicillin plates were made by warming Luria Bertani (LB)-agar to liquid state. Once the LB-agar was fully in molten form, the required volume of LB-agar was collected and once cooled sufficiently (just cool enough to hold LB-agar bottle), ampicillin was added (final concentration 100µg/mL). The agar was then poured carefully into sterile 10-cm dishes (approximately 5ml agar-ampicillin per dish). The agar was left to set within the dish for approximately 30min at room temp within a horizontal laminar flow hood, and then used.

Individual colonies were then picked, by scrapping off a single bacterial colony using a sterile pipette tip, and grown overnight in 3ml LB medium with ampicillin (100µg/mL) (shaker incubator 300rpm, 37°C). All well-grown bacterial cultures were carried forward for DNA plasmid MINIprep (spin microcentrifuge method) for an initial determination of whether the bacterial colonies carried the appropriate plasmid. Only 1ml of the bacterial culture was required for MINIpreps. If the bacterial culture was confirmed to be carrying the appropriate plasmid (using restriction enzyme digestion and gel electrophoresis, Section 2.6.5 and Section 2.6.6 respectively), 1.5ml of the remaining bacterial culture was frozen down in L-glycerol aliquots (stored at -80°C). The rest of the bacterial culture was added to 150ml LB medium with ampicillin (100µg/mL), and grown overnight (shaker incubator). The large flask of bacterial culture was then used in the QIAGEN MAXIprep protocol to obtain a high yield of the plasmid (0.5 to 1mg/ml).

2.6.4 DNA plasmid extraction

The well-grown flasks of 150ml bacterial culture, generated through the steps described in Section 2.6.3, were collected from the shaker incubator ready to be used in the MAXIprep protocol. The bacteria in each flask had been transformed with a different plasmid, therefore although MAXIpreps from different flasks may be performed in parallel, all tubes, stripettes and tips were kept separate to prevent contamination. The MAXIprep steps for a single 150ml culture have been provided as an example. Full details of the QIAGEN protocol can be found at: QIAGEN plasmid purification handbook (QIAGEN 2012).

First, the 150ml bacterial culture was equally split between three 50ml sterile centrifuge tubes, and centrifuged gently to pellet the bacteria (6000g, 15min, 4°C). During this time P1 buffer was prepared by adding RNase A (3:1000). On collecting the tubes, the supernatant was removed and for one of the three tubes, the P1 buffer-RNase A mix was added. The resulting bacterial suspension was then added in turn to the remaining two tubes of pelleted bacteria. This led to all the bacteria from the same original flask being combined. The bacterial suspension was then transferred into a single high-speed 50ml tube, which had pre-warmed P2 buffer (37 °C) added. The mixed solution was then left to incubate for 5min at room temp. On exactly 5min, pre-chilled (4 °C) P3 buffer was added to the tube, and the tube inverted until cloudy white debris was seen. To help precipitation of the debris, the tube was incubated on ice for 30min. The debris was then pelleted by centrifuging the high-speed centrifuge tube at 15,000g for 30min at 4 °C. The supernatant was then transferred to a fresh 50ml sterile high-speed centrifuge tube, with care taken not to transfer any of the white debris. The clear supernatant was then passed through the pre-prepared QIAGEN-tip 500, thus trapping the DNA plasmid within the QIAGEN-tip. To remove any remaining contaminants, 30 ml QC buffer wash was performed twice. To elute the plasmid DNA pre-warmed (65°C) QF buffer was passed through the QIAGEN-tip. To precipitate the DNA out of the QF buffer, 15ml isopropanol was added, and the tube was left at room temp for 15min. The plasmid DNA was pelleted by centrifugation (15,000g for 30min at 4 °C). The supernatant was then carefully removed, and pellet washed with chilled 70% absolute ethanol. To remove all traces of ethanol, the tube was spun again. Lastly, the pellet was left to air dry for 15-20min at room temp. The DNA plasmid was resuspended in 500µl PCR-grade water. The concentration and A260/280 of the DNA plasmid stock were determined using NanoDrop 2000 Spectrophotometer. Good quality DNA should have A260/280 of 1.8.

2.6.5 Restriction enzyme digestion

To ensure that the DNA obtained through MINIpreps or MAXIpreps (see Section 2.6.4) corresponded to the appropriate plasmid, restriction enzyme digestion of each plasmid was performed, followed by gel electrophoresis to visualise the DNA plasmid and digestion products (Section 2.6.6). Plasmid maps obtained from Addgene were viewed using SnapGene DNA programme, which provided details of the restriction enzyme cut sites in the DNA plasmid, and the expected digestion product sizes (GSL Biotech 2019). For each plasmid the following reactions were carried out: control with no restriction enzyme added, single-cutter only digestion, and two independent single-cutters (double digestion) or multi-cutters (2 or more products). Each 30µl restriction enzyme digestion reaction consisted of: 1µg of plasmid, 3µl of the appropriate NEB 10x buffer (either NEB 3.1 Buffer or NEB 2.1 Buffer or Cutsmart), 10 units of each restriction enzyme(s), and PCR-grade water added to bring the total volume to 30µl. If no restriction enzyme was added, the volume was replaced with PCR-grade water. The restriction enzyme digestion was carried out at 37°C for 5-7hrs using a thermocycler. The digestion products are shown in Appendix folder 5 figures and tables.

2.6.6 Gel electrophoresis

Once incubation for the restriction enzyme digestion was completed (Section 2.6.5), the samples were collected and ran on 1% agarose gel to determine the approximate size of DNA digestion products present. The 1% agarose gel was made by adding 2g of agarose powder (UltraPure Agarose, Invitrogen) to 200ml of 1x TAE buffer. The solution was warmed to dissolve the agarose fully. Once the molten agarose was just cool enough to hold, 20µl of Safeview (Safeview 1:10,000 dilution, NBS Biological Limited) was added, and thoroughly mixed. The agarose-Safeview liquid mix was poured into a gel tray and left to set for 30min at room temp with the 16-well comb (that forms the lanes) in place. During this time, the digestion samples (30µl) had 6x loading dye (NEB) added and mixed, and kept on ice until needed. Once set, the gel was removed from gel tray and placed within the gel electrophoresis box, and submerged in 1xTAE buffer. The digestion samples, with loading dye, were pipetted into each lane (30µl per lane). Typically 8µl of the DNA ladder (Hyperladder 1kb, Bioline or Smartladder 1kb, Eurogentec) was added into free lanes for each run (at least two lanes per run). The electrodes were connected to electrophoresis battery pack with the following settings: 90V, 300mA, 2hr. The gel was checked after first 15min to check for the appearance of DNA bands (ladder and samples) under ultraviolet (UV)-light using the UV Mini Transilluminator (Bio-Rad), and then left to run until the DNA ladder bands were clearly separated and samples had run approximately three-quarters of the gel length. The DNA bands were imaged using Molecular Imager Gel Doc XR System (Bio-Rad). The approximate size of each digestion product was checked against the expected size determined from the plasmid map on SnapGene (see Appendix folder 5 figures and tables).

2.7 Lentiviral vector: pBOB-synP-HTB

2.7.1 Transfection of HEK cells and viral harvest

To generate the non-self replicating lentivirus carrying pBOB-synP-HTB (encodes for TVA receptor, rabies glycoprotein, histone-GFP under the human synapsin promoter), this plasmid along with those encoding for critical lentiviral structural proteins (packaging plasmid psPAX2, and envelop plasmid pMD2.G) were introduced into HEK cells using lipofectamine 3000 transfection reagent (Thermo Fisher Scientific). Note that HEK cells had passage <12 on transfection. On the day preceding transfection, HEK cells were split at a low split ratio (1:3) using trypsin treatment on to 0.1% gelatin-coated plates (refer to Section 2.1.4). Only plates within the optimal confluency range (80-90%) were carried forward for lipofectamine based transfection. Between 2-4hr before transfection, the selected plates were washed with DPBS and given 10ml fresh HEK media. The Lipofectamine 3000 reagent protocol provided in the manufacturer's guide (Invitrogen 2014) was used with a few adjustments. First, the following quantity of plasmids were used to transfect a single high confluency 10-cm plate: 14µg pBOB-HTB, 5µg pMD2.G and 9µg psPAX2. The appropriate volumes of each plasmid were added to the P3000 reagent, and OptiMEM reagent mix. A second reagent mix of lipofectamine 3000 and OptiMEM was also generated. The two OptiMEM reagent mixes were added together (1:1) to generate the lipofectamine-DNA mix, and then left for 5min (at room temp). During this time, HEK media was removed from each plate and replaced with 6ml OptiMEM. The lipofectamine-DNA mix was added dropwise to each plate to bring the final volume for each 10-cm plate to 9ml. Plates were placed at

37°C, 5% CO_2 in viral suite incubator overnight. The media was replaced the following day (12-14hrs post-transfection) with 15ml fresh HEK media, and plates moved to 35°C, 3% CO_2 incubator until viral harvest. When the media began to turn orange, typically after 2-3 days, HEK media from each plate was collected, and replaced with fresh 10ml HEK media. The first collected viral harvest was centrifuged at 800g, 5min, 4 °C to pellet cell debris. The supernatant was then filtered through 0.45μm filter. The supernatant was then frozen down as crude viral harvest (stored at -80 °C) or concentrated through ultracentrifugation (Optima XPN 80 Ultracentrifuge, Thermofisher, with SW 32 Ti rotor, at 70,000g, 2hrs, 16°C). On completing ultracentrifugation, the supernatant was carefully decanted from the tubes, and viral pellet was resuspended in ice-cold DPBS (150-200μl). To aid resuspension, the tubes were gently vortexed (short 5s vortex, 10 times) and left for 1hr on ice before aliquoting and storing at -80°C. This process of viral harvest collection, centrifugation and filtering, and ultracentrifugation was repeated for the second viral harvest media.

2.7.2 Titering of lentivirus in hESC-derived interneurons

The crude and concentrated lentivirus was titered in hESC-derived interneurons (see Chapter 6 and Appendix folder 5 figures). The lentivirus was added to the neuronal maturation media at a final concentration 1:250, 1:500 or 1:750. The administration was either given as a single or a double hit (same concentration repeated 7 days apart). For all transductions, negative controls given viral-free media were used. After 5 days post-transduction, GFP positive cells were seen by eye using the fluorescent microscope. The GFP positive (transduced) population was quantified using flow cytometry (Section 2.10 and Appendix folder 5 figures) and immunostaining against GFP (Section 2.2). In all future lentiviral transductions of hESC-derived interneurons, single hit 1:250 dilution was used.

2.8 Pseudotyped ΔG ChR2-mCherry rabies viral tracer

The full details of the protocol to obtain pseudotyped glycoprotein deficient (ΔG) channelrhodopsin 2 (ChR2)-mCherry rabies virus (the viral tracer) can be found from Osakada & Callaway published protocol (Osakada & Callaway 2013).

2.8.1 Transfection of B7GG cells

Lipofectamine 2000 transfection reagent was used to transfect the B7GG packaging cells, which in turn generated and enabled amplification of the ΔG ChR2-mCherry rabies virus (the viral tracer). Note that B7GG cells used were at passage <10 at transfection. The transfection steps were similar to those described for lipofectamine 3000 protocol in Section 2.7.1, but with the following differences specified below and in the manufacturer's guide (Invitrogen 2013). The following quantity of each plasmid was used for each 10-cm plate transfection: 5µg pcDNA-SADB19G, 7.5µg pcDNA-SADB19L, 7.5µg pcDNA-SADB19P 15µg pcDNA-SADB19N, 30µg pSADdeltaG-ChR2-mCherry. The volume of OptiMEM media added to 10-cm plate of highly confluent B7GG cells was 5.5ml, and 2.5ml of DNA-lipofectamine 2000 mix was added dropwise to 10-cm plate. The B7GG cells were left for 6hr at 37 °C, 5% CO_2 and then the media was replaced with 15ml fresh HEK media. The following day the B7GG plates were transferred to the incubator with the adjusted conditions of 35 °C, 3% CO_2. Once the B7GG cells reached maximum confluency and media was orange, the B7GG cells were split (1:10) onto new T75 flasks using trypsin treatment (Section 2.1.4), and maintained to amplify the viral harvest.

2.8.2 Viral harvest of unpseudotyped ΔG ChR2-mCherry viral tracer

As transfected B7GG cells, generated in Section 2.8.1, have the ability to generate new ΔG ChR2-mCherry rabies virus, by providing the missing rabies glycoprotein, the transfected B7GG were maintained over a 2 week period to increase the concentration of the ΔG ChR2-mCherry rabies viral harvest. The mCherry positive cells were monitored by eye under the fluorescent microscope, and once a high proportion of high intensity mCherry positive cells were observed, the first viral harvest was collected (typically 1-2 weeks post-transfection). The viral harvest was performed in the same manner described in Section 2.7.1, except with Optima XPN 80 Ultracentrifuge, SW 32 Ti rotor, set to 70,000g, 2hrs, 4°C. Also, the majority of the ΔG ChR2-

mCherry rabies viral harvest was kept as crude media, ready to be pseudotyped.

2.8.3 Generating EnvA pseudotyped ΔG ChR2-mCherry viral tracer

Crude unpseudotyped viral harvest obtained from B7GG cells (Section 2.8.2) was pseudotyped using BHK-EnvA cells. BHK-EnvA cells were grown to 70% confluency in T75 flasks and transduced with the crude viral harvest media (0.45μm filtered). Within 12hr, fresh HEK media was added (1:2 dilution). After a further 24hrs, BHK-EnvA cells were split using trypsin treatment (Section 2.1.4), and resuspended into fresh HEK media. BHK-EnvA cells were washed thoroughly, twice, with DPBS on splitting. On the day following splitting, the media was replaced with fresh HEK media. These wash steps and repeated change of HEK media ensured that any unpseudotyped ΔG ChR2-mCherry rabies virus was removed, thereby preventing contamination of the pseudotyped viral harvest media. The expression of mCherry by the BHK-EnvA cells was checked under the fluorescent microscope every day. The mCherry positive cell count, by eye, appeared at approximately day 5 post-transduction, peak at day 7 and then plateau past day 7. At day 7-10 post-transduction, the pseudotyped viral harvest media was collected and processed as described for the lentivirus in Section 2.8.2.

2.8.4 Titering of pseudotyped and unpseudotyped ΔG ChR2-mCherry viral tracer

Titering of pseudotyped and unpseudotyped ΔG ChR2-mCherry rabies virus was conducted using HEK and HEK TVA cells (see Chapter 6 and Appendix folder 5 figures). Pseudotyping had been used to change the tropism of the ΔG ChR2-mCherry rabies virus, so that EnvA pseudotyped ΔG ChR2-mCherry rabies virus can only enter and infect TVA expressing cells. To determine if restricted entry had been achieved, transduction using the concentrated pseudotyped and unpseudotyped ΔG ChR2-mCherry rabies viral stocks (Section 2.8.2 -2.8.3) were carried out in HEK and HEK TVA cells in parallel. The appropriate viral stock was added to the HEK media at a final concentration 1:100, 1:250, 1:500 or 1:750. This marked day 0 of transduction. For all transductions, negative controls given viral-free media were used. After 5 days post-transduction, mCherry positive cells were seen by eye using the fluorescent microscope (although not in all culture conditions). The mCherry positive (transduced) population was quantified using flow cytometry (Appendix folder 5 figures) and immunostaining against GFP (Section 2.2). In all future EnvA pseudotyped ΔG ChR2-mCherry rabies viral transductions, 1:100 dilution was used.

2.9 Application of the completed viral tracing system

Lentiviral transductions were performed on day 30-35 hESC-derived interneurons at 1:250 final concentration. This led to the generation of 40-50% starter neuron (GFP positive) population in the culture. At day 50, this culture was then infected with the EnvA pseudotyped ΔG ChR2-mCherry rabies tracer (final concentration 1:100). The culture was then left 7 days before downstream applications began. These applications included immunostaining for GFP and mCherry (Section 2.2), and electrophysiology and optogenetics (Section 2.11). The complete viral tracer system was used up to 14 days post-infection with rabies tracer.

2.10 Flow cytometry

To prepare cells for flow cytometry, a single well of 12-well plate was used for each viral dilution, along with the negative (non-viral treated) control. Cells were washed twice with DPBS, and then detached from the well using enzymic treatment (trypsin for HEK293T/HEK TVA cells, accutase for hESC-derived interneurons) (see Section 2.1.4 and 2.1.6, respectively). The cell suspension was diluted 10-fold in DPBS and manually dissociated by trituration. A repeated cycle of centrifuging the suspension (300g, 5min, 4°C) to pellet the cells, removing the supernatant, and then resuspending the pelleted cells in chilled DPBS was performed three times. The washed single-cell suspension in DPBS was then passed through a cell strainer (35μm nylon mesh). The strained single-cell suspension was used on the flow cytometer (BD LSR Fortessa cell analyser, BD Biosciences). Flow cytometer was used to detect hESC-derived interneurons expressing GFP ("FITC" excitation laser 488nm, excitation wavelength 493nm, emission max 525nm) and HEK/HEK TVA cells that were mCherry positive ("PE-Texas Red" excitation peaks 496nm and 565nm, excitation laser 488nm, emission max

613nm). The BD FACSDiva 8.0.1 software was used to generate the histograms and scatterplots (Appendix folder 5 figures).

2.11 Functional characterisation of hESC-derived interneurons

2.11.1 Electrophysiology solutions for whole-cell patch clamp

External solution was prepared fresh weekly. The external solution consisted of fresh 1L sterile water with: sodium chloride (NaCl) 7.89g, potassium chloride (KCl) 0.373g, magnesium chloride ($MgCl_2$) 0.244g, 2mM stock calcium chloride ($CaCl_2$) 0.625ml, D-glucose 1.8g, HEPES 1.192g. Once the solution was thoroughly mixed, 1L bottle remained on the magnetic spinner and sodium hydroxide (NaOH) was added drop-wise to achieve a final pH of 7.4 (measured using pH meter). The internal solution (also known as the pipette solution) was made in bulk and stored at -20°C in 500µl aliquots. The internal solution used was potassium-gluconate based. The internal solution consisted of 1L sterile water with: 27.4g K-gluconate, 0.584g NaCl, 2.62g HEPES, 1.1g Na_2-ATP, 1.05g Na-GTP, 0.306g Na_2-phosphocreatine, 0.41g $MgCl_2$, 2mM stock $CaCl_2$ 0.5ml, and 4.184g EGTA. In a similar manner described for the external solution, the pH of the internal solution was adjusted to pH 7.2 using KOH.

2.11.2 Immediate steps prior to patch-clamp recordings

For all electrophysiology experiments, progenitors were seeded on UV-light sterilised PDL/laminin coated coverslips to allow removal of the cells and transfer to electrophysiology rig. On collection of the coverslip (on day of the experiment), cells were gradually exposed to the external solution. The internal solution aliquot was thawed and filtered through polypropylene centrifuge tube filter (0.22µm filter) using a short spin, to separate internal solution from any contaminants. The filtered internal solution was stored on ice until needed. Glass patch pipettes were prepared fresh each day using a programmable mechanical Puller (Flaming/Brown Sutter micropipette puller P1000). Glass capillaries (Sutter) were heated and pulled to generate a patch pipette with resistance between 4-8 MΩ.

2.11.3 Whole-cell patch clamp

Electrophysiology was carried out on hESC-derived interneurons, the day in vitro that the culture was collected for whole-cell patch clamping has been specified in the data shown (Chapter 6). On immediately breaking into the neuron, the resting membrane potential was noted. To determine additional intrinsic and evoked parameters of the patched neuron, three distinct whole-cell patch clamp paradigms (in current-clamp mode) were performed sequentially on the patched neuron (holding potential -60mV). For illustrative examples of each paradigm please see Appendix folder 6 figures. In paradigm one (Appendix figure S) the patched neuron was exposed to a current injection drop of 10pA for 1s, repeated 10 times in succession. The input resistance and membrane time constant (tau) were determined from the averaged voltage deflection traces during the final 0.1s of -10pA current injection. The capacitance could then be calculated from these values (capacitance = tau/input resistance). In paradigm two (Appendix folder 6 figure T), the patched neuron was injected with current progressing from −70pA to maximum of +110 pA in 20pA steps (10 steps in total, each lasting 1s). The ability of the neuron to generate an action potential or train of action potentials was then determined for each current step. In paradigm three (Appendix folder 6 figure U) an alternative strategy was used to evoke an action potential response, namely the short 5ms pulse of a very large current injected (+200pA). This pulse of current was repeated 5 times in succession. In paradigm two and three, an action potential was only counted as successful if the spike reached above 0mV. While holding the patched neuron at -60mV spontaneous activity (mini excitatory postsynaptic potentials and action potentials) were also recorded. Analysis of traces was carried out on Clampfit 10.6 (Molecular Devices 2019).

2.11.4 Optogenetics

The complete monosynaptic viral tracer system described in Section 2.9 was analysed using whole-cell patch clamp recordings and optogenetics. The EnvA pseudotyped ΔG ChR2-mCherry rabies tracer enforces the expression of ChR2-mCherry fusion protein. ChR2 is a light-gated ion channel stimulated by blue light (490nm). To test this component of the viral tracer system, mCherry positive traced neurons were identified in the culture and patched.

The patched neurons were stimulated with both current injections (paradigms two and three described in Section 2.11.3) and blue light pulses (Chapter 6) to determine if ChR2 was efficiently expressed and functional. The photostimulation of the traced neuron ranged from 0.5ms exposure to up to 10s exposure.

2.12 Data processing and statistics

2.12.1 Figures

Illustrative figures were created in Powerpoint (Microsoft, 2016) and when required, edited using InkScape (Inkscape 0.92.4). Fluorescent images were then taken using Leica DMi 6000B inverted or Leica DM 6000B upright (for glass slides) microscopes at 20x magnification (within 2 weeks of staining, Section 2.2). Where indicated, deconvolution images were taken, which involved collecting a series of stacked images through the z plane, and then merging into a single image using post-imaging processing steps on Leica AF Hardware and ImageJ (Rasband 1997-2018). Cell counting was performed using ImageJ and Cell Profiler (Carpenter *et al.* 2006). Graphs were created using Graphpad Prism 5 (GraphPad Software 1995-2017).

2.12.2 Statistics

Statistical tests were also conducted in Graphpad Prism 5, and included: nonparametric, Mann-Whitney test, when comparing two groups, and Kruskal-Wallis test with Dunn’s multiple comparisons post-hoc test when comparing 3 or more groups.

2.12.3 Single-cell transcriptome analysis

Single-cell QPCR data, in the form of a matrix of Ct values for each sample-assay reaction, obtained from the Biomark HD 96.96 IFC run was initially processed on the Fluidigm real-time PCR analysis software (Fluidigm 2018a). The data was processed to calculate delta Ct values (relative to house-keeping gene, beta-actin). Any poor quality data (e.g. those that failed to express house-keeping genes) were identified and excluded from further analysis. The exported data sheet of delta Ct values was then ready for analysis on R programme (R Core Team 2013). Single-cell RNA sequencing data was also pre-processed for use in the Fluidigm Singular analysis toolset (Fluidigm 2018b). The initial pre-processing steps were performed by Dr Peter Giles (Cardiff institute of medical genetics), with the remaining analysis using the Singular analysis package performed by myself. The pre-processing steps began with trimming the reads to remove sequencing adapter and poor quality bases using Trim Galore. Trimmed reads were then mapped against the GRCh38 no_alt_plus_hs38d1_analysis_set (GENCODE GRCh38.p10) using STAR aligner tool (version 2.5.0b) (Dobin *et al.* 2013). Finally, the expression counts were converted to transcripts per million ([number of reads mapped to a selected gene / total number of mapped reads of a sample]*10^6) unit using RSEM. Quality control checks of the reads led to removal of low quality reads prior to analysis. Once in the correct format, the expression data set was exported as a matrix that could be uploaded into R programme for analysis. Singular analysis package was used to generate a list of differentially expressed genes between time points/ages/independent samples by performing ANOVA pair-wise analysis, with genes only included if the p-value <0.05 and the relative fold change >4 (in either direction). Moreover, the Singular analysis package was used to generate the gene expression heatmaps and tSNE plots (Chapter 4 and 5, and Appendix folders 2-4). All analysis was conducted using data trimmed to the top 400 differentially expressed genes. The script required to convert Ensembl gene IDs to gene names was created by Dr Daniel Cabezas de la Fuente (Li lab).

Chapter 3 - Interneuron differentiation of hPSC

3.1 Introduction

As described in detail within the General Introduction (Chapter 1), inhibitory interneurons represent a classic example of lineage diversification, which to this date remains to be fully understood (Flames & Marín 2005; Kessaris *et al.* 2014). The use of hPSC-derived interneurons could, therefore, provide novel insight into interneuron lineage commitment and differentiation trajectories that are not easy obtained from the primary source, due to the scarcity and legal age restrictions of human foetal tissue usage.

3.1.1 Extrinsic cue based interneuron differentiation paradigms

Differentiation paradigms using extrinsic cues aim to recapitulate the native environment in vitro, and therefore provide a more authentic neurodevelopmental model. In general, knowledge of which extrinsic cues has been guided from rodent work (Backman *et al.* 2005; Xu *et al.* 2010). These in vivo studies have led to the understanding that critical morphogens regulate neuronal induction and spatial patterning within the brain (refer to Chapter 1). A decade ago, a simple and highly efficient method to indirectly induce neural fate commitment in hPSC (by suppressing alternative fates) was described using two chemical factors that inhibit SMAD signalling pathways (dual SMAD inhibition) (Chambers *et al.* 2009). The introduction of additional morphogens to direct a specific neuronal lineage further enhanced the value of hPSC-derived neuronal models, by allowing enrichment of the neuronal class of interest. In the case of interneuron differentiation, SHH and/or SMO receptor agonists (SAG and purmorphamine) were applied (Maroof *et al.* 2013; Nicholas *et al.* 2013; Anderson *et al.* 2016).

3.1.2 Shortcomings of in vitro interneuronal models

However, obtaining defined interneuron subtypes through in vitro interneuron differentiation paradigms remains a critical goal and challenge within the field. Due to the vulnerability of specific subtypes in different neuropathologies, the inability to fine-tune the in vitro interneuron subtype composition has limited the scope of PSC-derived interneurons as an in vitro neurodevelopmental and disease model. For example, there has been a common focus on generating MGE-derived cortical PVALB interneurons in vitro, which stems from the implication of cortical PVALB interneurons in numerous neuropathologies (schizophrenia, ASD, epilepsies) (Powell *et al.* 2003; Kataoka *et al.* 2010; Marín 2012). However, current paradigms typically report either low abundance or absence of PVALB interneurons in the terminally differentiated culture (Liu *et al.* 2013; Nicholas *et al.* 2013; Kim *et al.* 2014; Sun *et al.* 2016; Yang *et al.* 2017; Noakes *et al.* 2019). Although when hPSC interneuron differentiation has been completed in vivo, in the rodent neocortex, a greater abundance of PVALB interneurons are generated from the graft (Chen *et al.* 2013; Nicholas *et al.* 2013). This has demonstrated that, provided the optimal environment has been provided, PVALB interneurons can be efficiently generated from hPSC-derived interneuron progenitors. This has encouraged research aiming to explore the full extent and potential of PSC-derived interneuron heterogeneity, and whether refinements to current differentiation paradigms could enrich the culture in desired subtypes (Tyson *et al.* 2015; Close *et al.* 2017; Xiang *et al.* 2017).

3.1.3 Refining interneuron differentiation paradigms that use extrinsic cues

3.1.3.1 Sonic hedgehog signalling

Through attempts to improve the efficiency of interneuron differentiation protocols, both the effect of different timings and duration of SHH treatment on interneuron fate commitment have been investigated (Danjo *et al.* 2011; Maroof *et al.* 2013; Tyson *et al.* 2015). This led to an understanding that an optimal window exists during differentiation when SHH selectively and efficiently enhances interneuron fate commitment (Danjo *et al.* 2011; Maroof *et al.* 2013). Moreover, the differentiation of mESC-derived interneurons under high and low SHH signalling has been shown to change SST: PVALB commitment bias of mESC-derived interneurons after completion of differentiation in vivo (early postnatal mouse neocortex) (Tyson *et al.* 2015). This nicely mirrored work in rodent models that demonstrated a SHH signalling gradient across the MGE, which caused a SST and PVALB commitment bias of the dorsal and ventral progenitors, respectively (Flames *et al.* 2007; Tyson *et al.* 2015). However, the same result has not been independently achieved by manipulation of SHH concentrations during interneuron differentiation of hPSC (Close *et al.* 2017; Xiang *et al.* 2017). If a change was reported, this was limited to a shift in the proportion of *SST* and *NPY* expressing neurons (elevated in higher SHH treatment culture), but no change in *PVALB* expression between treatments (Xiang *et al.* 2017). This may be due to species differences and/or poor efficiency of interneuron differentiation paradigms in generating PVALB interneurons in vitro (Liu *et al.* 2013; Nicholas *et al.* 2013; Kim *et al.* 2014; Sun *et al.* 2016; Yang *et al.* 2017; Noakes *et al.* 2019).

3.1.3.2 Temporal manipulation

Furthermore, one of the major shortcoming of in vitro differentiation paradigms is the inability to recapitulate the generation of temporally distinct progenitor pools that occur over an extended period in native development (late first to second trimester in humans) within the ventricular zone of the MGE/CGE (Zecevic *et al.* 2011; Reinchisi *et al.* 2012; Onorati *et al.* 2014; Al-Jaberi *et al.* 2015). Whether hPSC-derived interneuron progenitors could even be maintained in their highly proliferative, progenitor state through serial expansions, without loss of their differentiation potential, is currently unknown. If such an unlimited source of interneuron-like progenitors could be achieved from a single parent culture, this would greatly simplify and shorten the duration of any subsequent differentiation, thereby reducing costs and labour. However, evidence of temporal influence on interneuron lineage commitment would suggest that the sequential sister progenitor cultures may not have the same resulting subtype composition or functional maturity (Butt *et al.* 2005; Miyoshi *et al.* 2010; Inan *et al.* 2012; Tyson *et al.* 2015; Arendt *et al.* 2016; Bardy *et al.* 2016). Primarily, this evidence came from utero-transplantation fate-mapping studies in rodents (Butt *et al.* 2005; Inan *et al.* 2012). On transplanting reporter labelled mouse MGE/CGE of different ages (E13.5 vs E15.5) into wild-type embryos of matching age, the proportion and diversity of the reporter labelled interneuron subtypes identified in the adult mouse brain were found to be dependent on the age of graft at transplantation (Butt *et al.* 2005; Inan *et al.* 2012). But why? Molecular characterisation of MGE/CGE primary tissue over an extended period of development in mouse and human is limited, and often left as a qualitative histology staining rather than a quantified measure (Butt *et al.* 2005; Zecevic *et al.* 2011; Hansen *et al.* 2013; Zechel *et al.* 2014). Therefore, markers differentially expressed between early and late-born progenitors of the same region of origin are unknown. Such knowledge would be fruitful for selectively enriching in vitro derived interneuron subtype populations for a specific subtype (e.g. PVALB interneurons).

3.2 Aims

To characterise the developmental trajectory of hPSC-derived interneurons from progenitor to post-mitotic stages of the differentiation paradigm.

To investigate whether sequential serial expansions of hPSC-derived interneuron progenitors can successfully maintain the progenitor population over an extended period.

To determine whether the serial expansions mimic the generation of early and late-born progenitors with differential lineage commitments, as reported during native MGE/CGE development.

3.3 Results

As described in detail in Chapter 2, and illustrated in Figure 3.1A, hESC were differentiated as a monolayer culture under extrinsic cue paradigm to generate interneuron-like progenitors, which completed terminal differentiation in vitro to give rise to post-mitotic interneuron-like cells. During the first stage, hESC undergo forebrain neural induction (using dual SMAD and WNT inhibition), proceeded by treatment with interneuron differentiation media (SHH and purmorphamine) to mimic native conditions that promote MGE/CGE fate commitment in the brain (Figure 3.1A). The progenitors were driven to terminally differentiate under neuronal maturation media, supplemented with neurotrophic factor BDNF (Figure 3.1A).

3.3.1 SHH treatment efficiently generated interneuron-like cells

In order to evaluate the differentiation trajectory of the ventralised neuroepithelial cells, samples were collected at day 10, 20, 35, 50 and 65 for bulk-QPCR analysis (Figure 3.1B-3.1C). The day 10 marked the start of SHH treatment, and was used to determine the relative change induced by exogenous SHH signalling (Figure 3.1B-3.1C). To complement the RNA level analysis, samples were also collected for immunohistochemistry staining at day 20, 35, 50 and 65 (Figure 3.1D-3.1G).

At RNA level, an early MGE marker, *NKX2.1* level was found to be highest at the earliest time point, and then remain at a slightly reduced, but maintained expression level across the differentiation period (Figure 3.1B). The high expression of *NKX2.1* from day 10 likely relates to the ventralising effects of WNT inhibitor XAV939 in the neural induction culture media. The maintained expression of *NKX2.1* may indicate progenitor contamination of later cultures or a striatal (rather than cortical) interneuron identity of the post-mitotic interneuron-like cells. Whereas, late MGE markers *LHX6* and *LHX8* displayed mirrored temporal dynamics in expression levels, each peaking at day 35 and then downregulated in the later stages of differentiation (Figure 3.1B). This coincided with the peak of *SST* at RNA level at day 35 (approx. 45-fold increase relative to day 10 expression level) (Figure 3.1C). This could indicate the insufficiencies of the neuronal maturation media to maintain these subpopulations of post-mitotic interneurons and/or promote high expression of these markers. A dorsal MGE and CGE marker, *COUPTF II* expression was upregulated between day 10 and 20 (5-fold increase), then declined to a similar level of expression as at day 10 in the later time points (Figure 3.1B). Whereas, post-mitotic GABAergic interneuron markers *GAD1* and *CR* progressively increased over time, both of which reached maximal expression at day 65 (Figure 3.1C). Relative to day 10, *PVALB* expression showed a 8-fold increase in expression at day 20 and then continued to decline back to day 10 expression level (Figure 3.1C).

3.3.2 The day 20 population was abundant in MGE-like progenitors

Immunostaining of day 20 revealed that the population was highly abundant in cells expressing a pan-neuroepithelia marker Nestin, the forebrain transcription factor FOXG1 (60%) and NXK2.1 (80%) (Figure 3.1D-3.1E). In agreement with bulk-QPCR analysis, the proportion of NKX2.1 positive cells remained high in the later stages differentiation (30-50%) (Figure 3.1E). However, the day 20 cultures had low abundance of cells expressing a late MGE marker LHX6 (20%), which remained low in later cultures (Figure 3.1E). Immunostaining for COUPTF II, revealed a low and constant population of COUPTF II expressing cells over time (25%) (Figure 3.1E). To further evaluate the dorsal-ventral and anterior-posterior identity of the day 20 cultures, the proportion of cells expressing the dorsal telencephalon marker PAX6 and floor plate marker FOXA2 was examined (Figure 3.1D). Less than 20% of the cells stained positive to these markers ,which likely accounted for the NKX2.1/FOXG1 negative population (Figure 3.1D).Together the QPCR and immunohistochemistry data would indicate that the day 20 cultures were mostly composed of MGE-like progenitors with a small proportion of CGE-like progenitors and nascent postmitotic interneurons.

3.3.3 Terminal differentiation gave rise to diverse interneuron-like subtypes

As the period within neuronal maturation media lengthened, the post-mitotic interneuron-like cells grew in abundance (Figure 3.1F-3.1G). This was shown by the high proportion of cells staining positive for pan-neuronal marker TUJ1 at day 65 (Figure 3.1F). At day 65, the culture was enriched in GABAergic neuronal marker GAD1 positive cells (55%), with a very small proportion of glutamatergic marker vGLUT1 positive cells (<1%) (Figure 3.1F). Hence, the interneuron-like progenitors generated by day 20 successfully differentiated into GABAergic neuron-like cells. In order to evaluate the subtype diversity of the interneuron-like cells, known MGE-derived and CGE-derived subtype markers were checked (Figure 3.1F-3.1G). At day 65, approximately 30% of the population stained positive for dorsal MGE and CGE-derived CR positive interneurons. This subpopulation of interneuron-like cells grew progressively in numbers, reaching and maintaining a peak of 25-30% abundance at day 50 to 65 (Figure 3.1F). The CR interneurons may be derived from the CGE-like progenitors (COUPTF II positive, NKX2.1 negative) at day 20 (Figure 3.1D). Although the high abundance of CR relative to MGE-derived subtype markers SST (10%) and PVALB (<3%) may suggest that NXK2.1 positive/COUPTF II negative progenitors may also give rise to CR interneurons (Figure 3.1F). In a similar manner to CR, the number of SST positive cells progressively increased with time, which did not reflect the sharp fold-change at day 35 reported in QPCR data (Figure 3.1C and 3.1F). Calbindin (CALB1) was also expressed by a large proportion (35%) of the day 65 population, likely due to the post-mitotic interneuron marker not being a mutually exclusive subtype marker, but expressed by multiple interneuron subtypes (e.g. CR, PVALB and SST) (Figure 3.1F). Finally, a proportion of the GABAergic negative population was accounted for by ChAT interneuron-like cells, which made up approximately 20% of the total population at day 65 (Figure 3.1F). ChAT interneurons are also MGE-derived, and most likely accounts for the *LHX8* expressing interneurons precursors (Figure 3.1B and 3.1F).

In summary, molecular characterisation of interneuron-like cells over the course of the differentiation paradigm using bulk QPCR and immunohistochemistry indicated that a predominantly early MGE-like progenitor population was efficiently generated by day 20. This led to an abundant interneuronal population, marked by the high proportion of GAD1, CALB1 and CR positive cells by day 65. Cholinergic interneuron-like cells also added to the heterogeneity of the culture. However, does this characterisation describe the full differentiation potential of the early MGE-like progenitor population?

Figure 3.1A

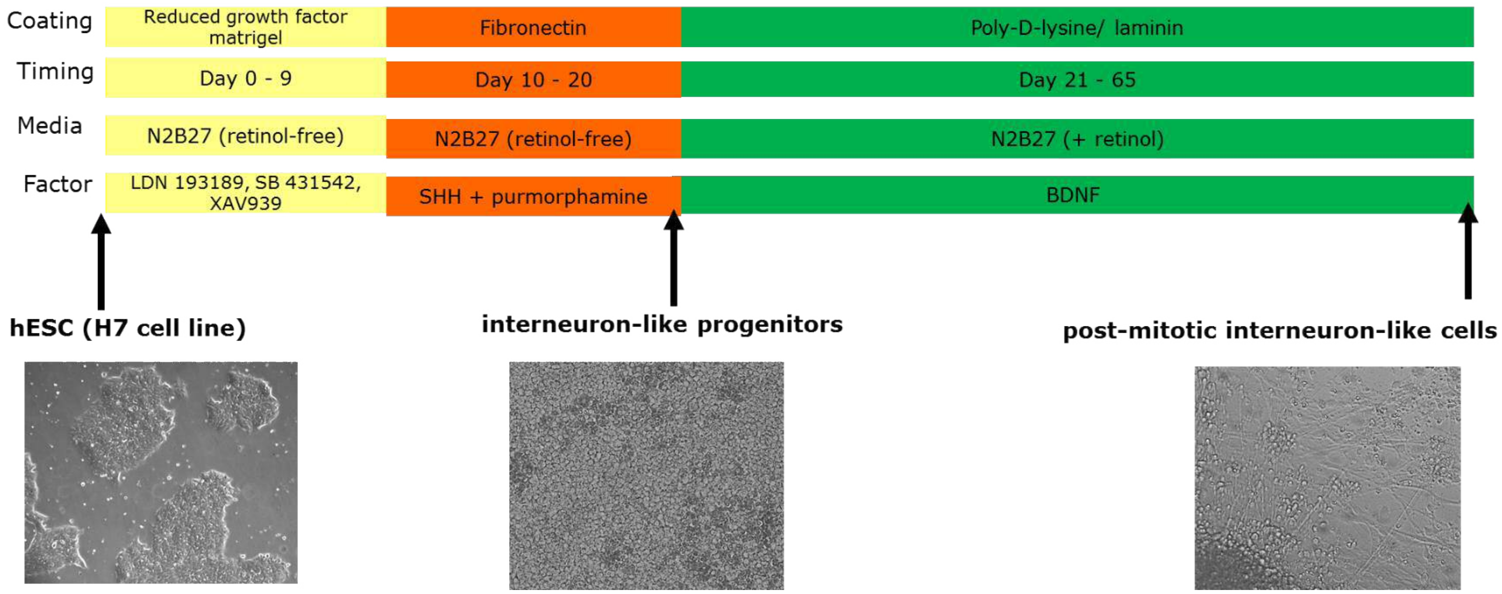

Coating
Reduced growth factor matrigel
Fibronectin
Poly-D-lysine/ laminin
Timing
Day 0 - 9
Day 10 - 20
Day 21 - 65
Media
N2B27 (retinol-free)
N2B27 (retinol-free)
N2B27 (+ retinol)
Factor
LDN 193189, SB 431542, XAV939
SHH + purmorphamine
BDNF
hESC (H7 cell line)
interneuron-like progenitors
post-mitotic interneuron-like cells

Figure 3.1B

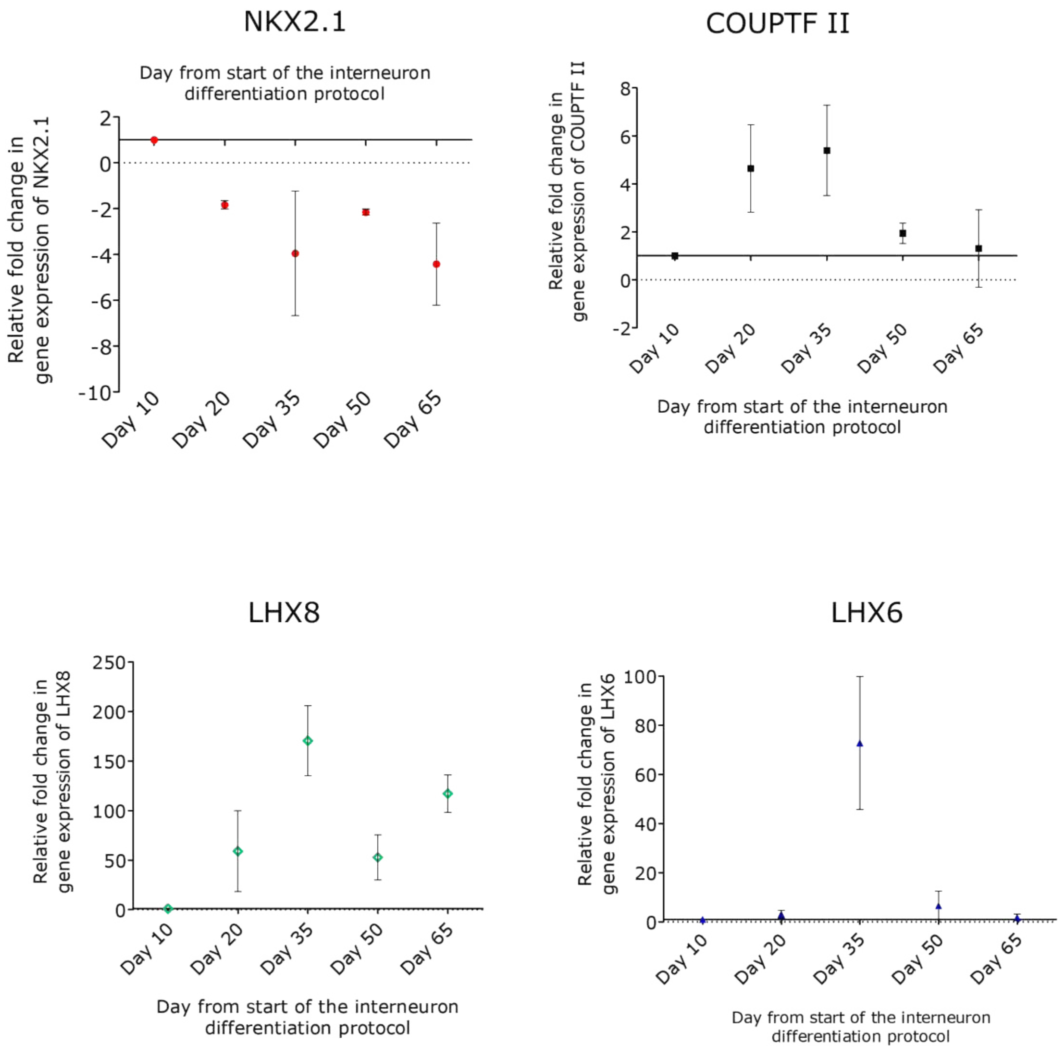

NKX2.1
Day from start of the interneuron differentiation protocol
Relative fold change in gene expression of NKX2.1
2
0
-2
-4
-6
-8
-10
Day 10
Day 20
Day 35
Day 50
Day 65
COUPTF II
Relative fold change in gene expression of COUPTF II
8
6
4
2
0
-2
Day 10
Day 20
Day 35
Day 50
Day 65
Day from start of the interneuron differentiation protocol
LHX8
Relative fold change in gene expression of LHX8
250
200
150
100
50
0
Day 10
Day 20
Day 35
Day 50
Day 65
Day from start of the interneuron differentiation protocol
LHX6
Relative fold change in gene expression of LHX6
100
80
60
40
20
0
Day 10
Day 20
Day 35
Day 50
Day 65
Day from start of the interneuron differentiation protocol

Figure 3.1C

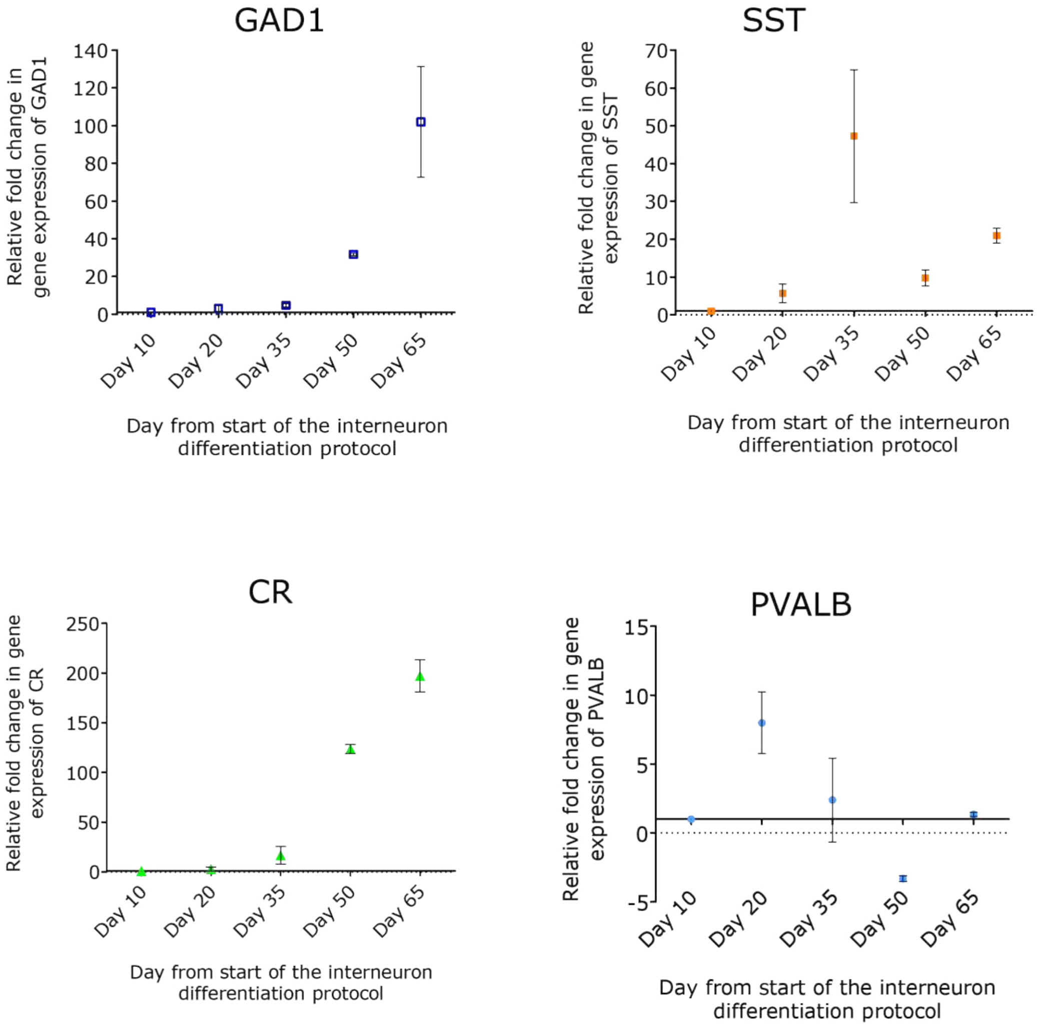

GAD1
Relative fold change in gene expression of GAD1
140
120
100
80
60
40
20
0
Day 10
Day 20
Day 35
Day 50
Day 65
Day from start of the interneuron differentiation protocol
SST
Relative fold change in gene expression of SST
70
60
50
40
30
20
10
0
Day 10
Day 20
Day 35
Day 50
Day 65
Day from start of the interneuron differentiation protocol
CR
Relative fold change in gene expression of CR
250
200
150
100
50
0
Day 10
Day 20
Day 35
Day 50
Day 65
Day from start of the interneuron differentiation protocol
PVALB
Relative fold change in gene expression of PVALB
15
10
5
0
-5
Day 10
Day 20
Day 35
Day 50
Day 65
Day from start of the interneuron differentiation protocol

Figure 3.1D

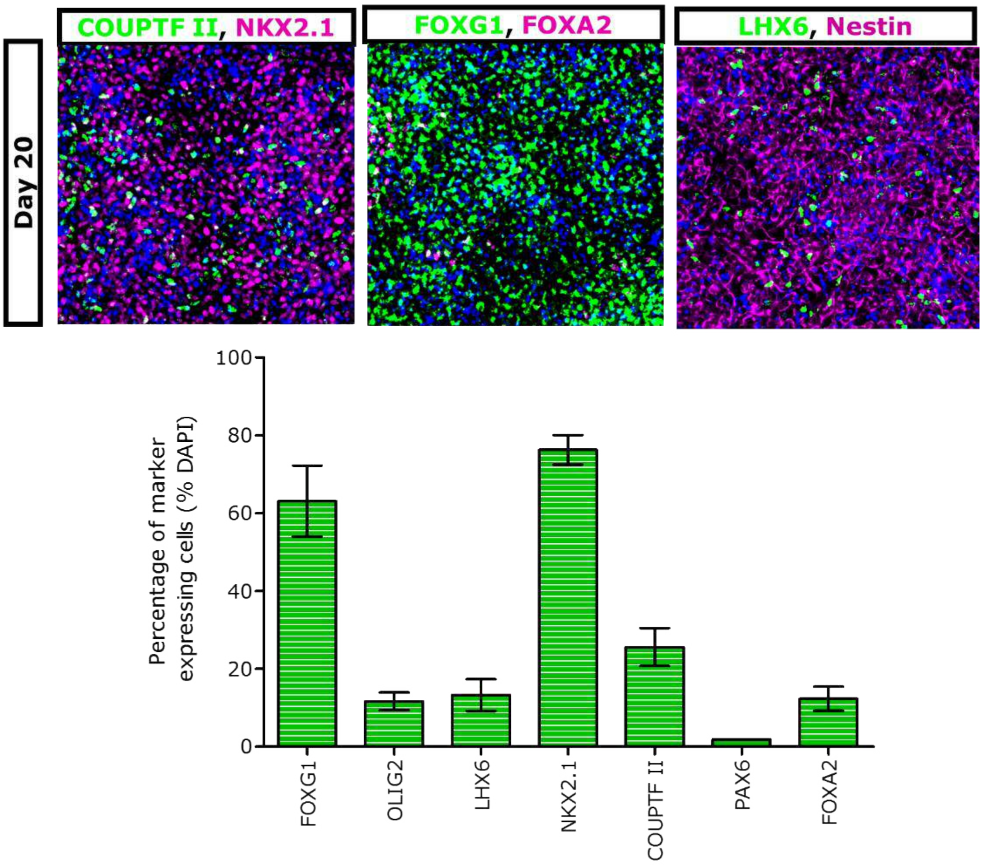


Figure 3.1E

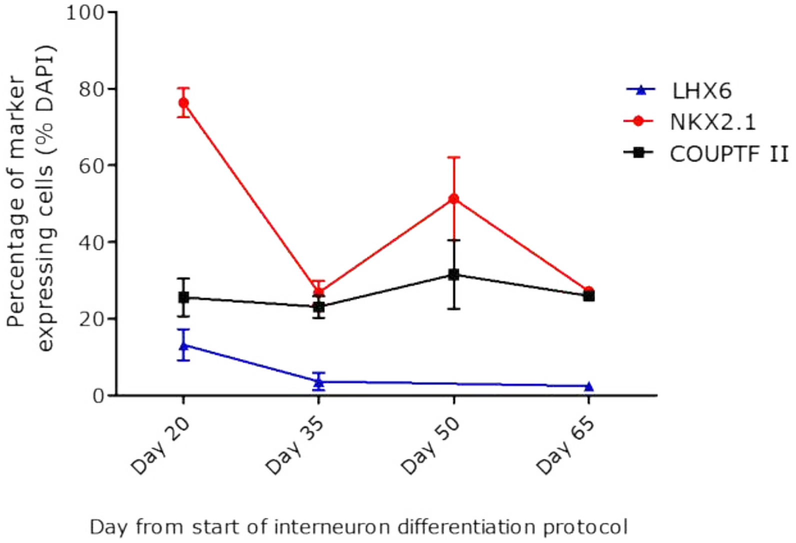


Figure 3.1F

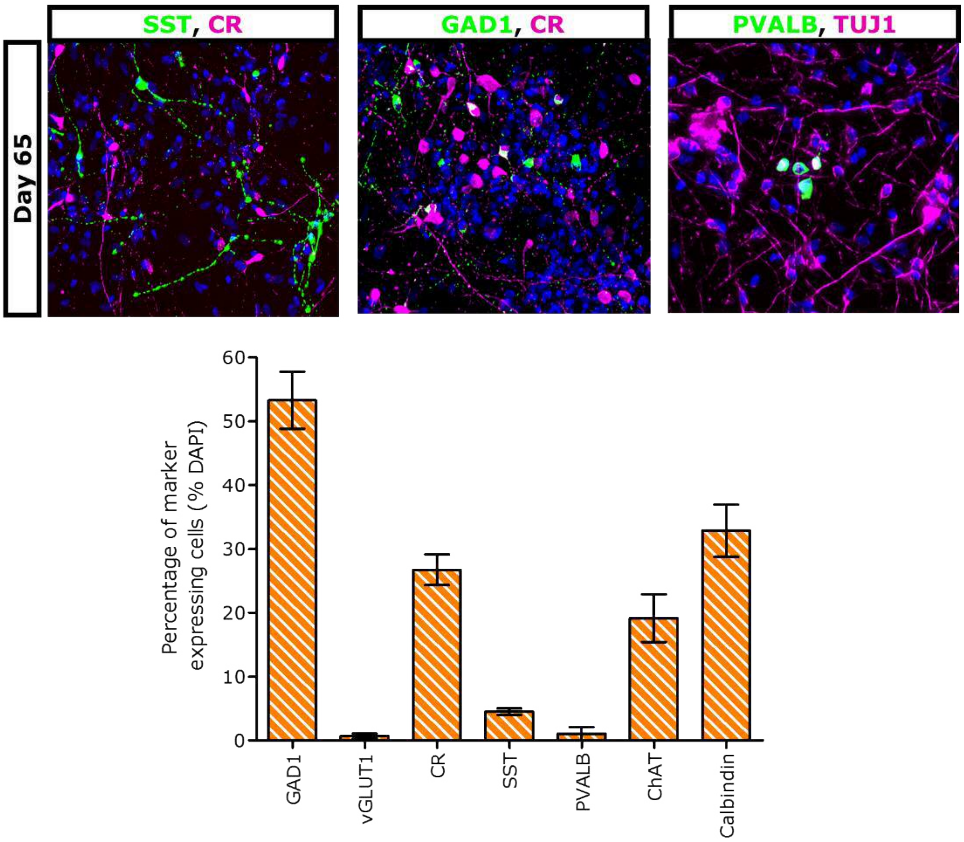


Figure 3.1G

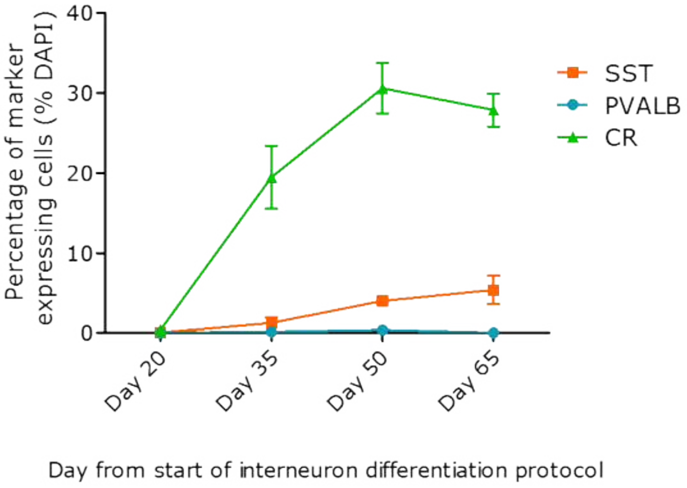


Figure 3.1 Longitudinal molecular characterisation of hESC-derived interneurons.

A) Schematic of the in vitro hPSC interneuron differentiation protocol utilised in this study, with real-time brightfield images of the in vitro culture at different time points. The paradigms has been simply divided into three different stages. Stage 1, neural induction (yellow). Stage 2, interneuron fate commitment (orange). Stage 3, terminal differentiation (green). For more information please refer to Chapter 2.

B-C) Developmental trajectory of hESC-derived interneurons – at RNA level, day 20, 35, 50 and 65 in vitro. The data has been presented in the form of relative gene expression, with the relative fold change made in reference to day 10 of the in vitro interneuron differentiation protocol (all day 10 fold changes =1). The initial delta Ct was calculated in reference to beta-actin as the standard housekeeping gene. The relative gene expression of a gene (RE) was calculated as RE= 2-(double delta Ct), along with +/- SEM. Biological replicates n=3. Technical replicates n=3.

D) Molecular characterisation of hESC-derived interneuron-like progenitors – at protein level, day 20 in vitro. The data has been presented in the form of a panel of immunohistochemistry deconvoluted images, and percentage of marker expressing cells out of the total cell population (total cell count measured using DAPI count). The graph plotted shows the mean percentage of marker expressing cells for each time point, along with +/-SEM. Biological replicates n=4. Technical replicates n=12 per biological replicate. The average total number of cells counted per technical replicate can be found in Appendix folder 7.

E) Longitudinal evaluation of the expression of lineage critical transcription factors – at protein level, day 20, 35, 50 and 65 in vitro. The data has been presented in the form of percentage of marker expressing cells out of the total cell population (total cell count measured using DAPI count). The graph plotted shows the mean percentage of marker expressing cells for each time point, along with +/- SEM. Biological replicates n=3 . Technical replicates n=12 per biological replicate. The average total number of cells counted per technical replicate can be found in Appendix folder 7.

F) Molecular characterisation of hESC-derived interneuron-like cells – at protein level, day 65 in vitro. The data has been presented in the form of a panel of immunohistochemistry deconvoluted images, and percentage of marker expressing cells out of the total cell population (total cell count measured using DAPI count). The graph plotted shows the mean percentage of marker expressing cells for each time point, along with +/- SEM. Biological replicates n=3. Technical replicates n=12 per biological replicate. The average total number of cells counted per technical replicate can be found in Appendix folder 7.

G) Longitudinal evaluation of interneuron subtype markers– at protein level, day 20, 35, 50 and 65 in vitro. The data has been presented in the form of percentage of marker expressing cells out of the total cell population (total cell count measured using DAPI count). The graph plotted shows the mean percentage of marker expressing cells for each time point, along with +/- SEM. Biological replicates n=3 . Technical replicates n=12 per biological replicate. The average total number of cells counted per technical replicate can be found in Appendix folder 7.

SHH, Sonic Hedgehog. BDNF, Brain derived neurotrophic factor. NK2 Homeobox 1, NKX2.1. COUP transcription factor 2, COUPTF II. LHX6, LIM Homeobox 6. LHX8, LIM Homeobox 8. OLIG2, oligodendrocyte transcription factor 2.FOXG1, Forkhead Box G1. PAX6, Paired box protein 6. FOXA2, Forkhead Box A2. SST, somatostatin. PVALB, parvalbumin. CR, calretinin. GAD1, glutamate decarboxylase 1. TUJ1, β-Tubulin 3. ChAT, choline acetyltransferase. vGLUT1, vesicular glutamate transporter 1. Gamma-aminobutyric acid-ergic, GABAergic. hESC, human embryonic stem cell. MGE, medial ganglionic eminence. dMGE, dorsal medial ganglionic eminence. CGE, caudal ganglionic eminence. SEM, standard error of the mean.

3.3.4 Serial expansions generated molecularly distinct interneuron-like progenitor populations

Sequential serial expansions of interneuron-like progenitor populations were performed to investigate two main questions. First, can the generation of newborn interneuron-like progenitors be extended by sequential serial expansions? Second, do progenitors derived from sequential serial expansions give rise to different interneuron subtypes, and is this reflected by changes in the molecular profile of the progenitors?

The experimental design has been summarised in Figure 3.2A. In brief, hESC underwent the three stages of interneuron differentiation paradigm, as described in Figure 3.1A. At day 20, which marked the end of stage 2 (MGE/CGE regionalisation with SHH), a proportion of the cells were allowed to undergo terminal differentiation in the neuronal maturation media until day 110 in vitro (stage 3). In parallel, a proportion of sister progenitors were maintained under stage 2 (SHH treatment) condition for 10 days (Figure 3.2A). This process was then repeated 5 more times to generate serial expansions of interneuron-like progenitors, each round generating a sequential culture to undergo terminal differentiation (named through 1 to 6). The first culture was seeded for terminal differentiation at day 20 and maintained until day 110 (total of 90 days in neuronal maturation media) (Figure 3.2A). The final culture (culture 6) was seeded for terminal differentiation at day 70 and maintained until day 110 (total of 40 days in maturation media) (Figure 3.2A).

To characterise the interneuron-like progenitors generated in each round of serial expansion, a proportion of the cells replated into stage 2 (SHH treatment) condition were collected for immunohistochemistry staining 1 day post replating (serial expansion number X + 1 day) (Figure 3.2B-3.2E). Immunostaining for early MGE progenitor marker NKX2.1, and late MGE interneuron precursor marker LHX6 was performed as an indication of the proportion of earlier-born (LHX6 positive) and newborn (NKX2.1 positive) progenitors in the population (Figure 3.2B-3.2D). Over the first three serial expansion (cultures 1 to 3), the proportion of NKX2.1 and LHX6 positive cells remained unchanged (Figure 3.2C- 3.2D). The abundant NKX2.1 positive population (70-80%), and absence of LHX6 positive cells in sequential cultures 1 to 3 would indicate that these first few serial expansions efficiently maintained proliferating interneuron-like progenitors, which gave rise to newborn progenitors (NKX2.1 positive). However, serial expansions of later derived cultures (cultures 4 to 6) led to the NKX2.1 positive population significantly decreasing to approximately 20% (***p<0.001, Figure 3.2C). Whereas, LHX6 positive population significantly increased, although to only 8% (***p<0.001, Figure 3.2D). Together these findings would suggest that MGE-like progenitors were transitioning out of the cell cycle by the later serial expansions, thereby failing to give rise to newborn NKX2.1 positive progenitors. This was supported by visual observations of reduced proliferation in the culture (confluency did not increase over the 10 days in SHH treatment). As a result, later derived cultures replated into stage 2 condition simply maintained interneuron precursors (NKX2.1 negative, LHX6 positive) for longer and longer durations in SHH before seeding for maturation. However, as the proportion of LHX6 positive cells was low, and the proportion of other late ganglionic eminence markers OLIG2 and GSX2, did not significantly change, this explanation cannot account for the entire loss of NKX2.1 positive cells (Figure 3.2B). Alternatively, one could argue that prolonged expansions of interneuron-like progenitors had resulted in the downregulation of NKX2.1 expression and subsequent loss of MGE-like identity. Although, to what alternative lineage is unclear, but mid-brain dopaminergic neuronal lineage marker FOXA2 and CGE marker COUPTF II remained unchanged (Figure 3.2B).

Overall, molecular characterisation of the sequential serial expansions of hESC-derived progenitors using immunohistochemistry indicated that proliferative interneuron-like progenitors were only maintained over a restricted, and small number of serial expansions.

Figure 3.2A

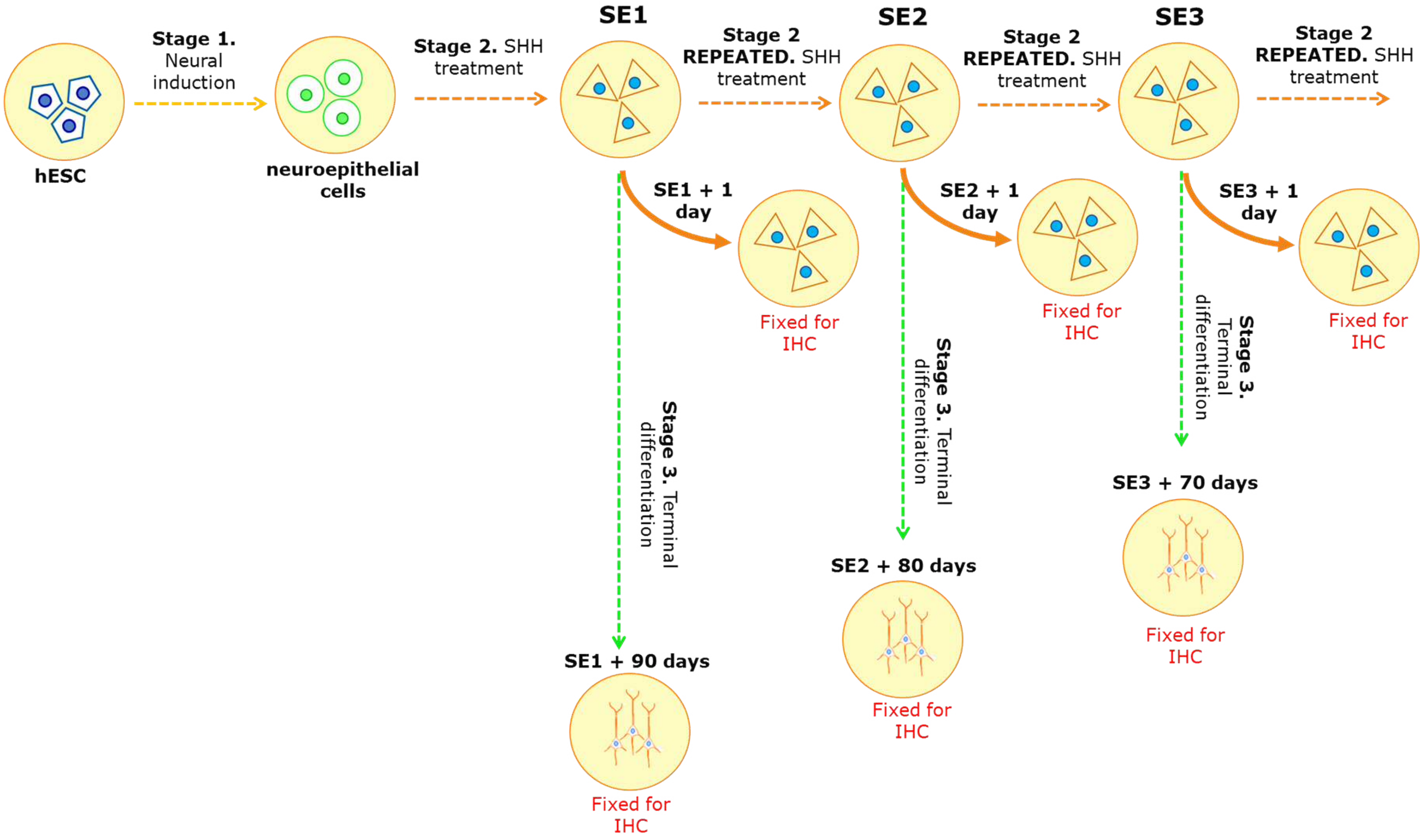

hESC
Stage 1. Neural induction
neuroepithelial cells
Stage 2. SHH treatment
SE1
Stage 2 REPEATED. SHH treatment
SE2
Stage 2 REPEATED. SHH treatment
SE3
Stage 2 REPEATED. SHH treatment
SE1 + 1 day
Fixed for IHC
SE2 + 1 day
Fixed for IHC
SE3 + 1 day
Fixed for IHC
Stage 3. Terminal differentiation
SE1 + 90 days
Fixed for IHC
Stage 3. Terminal differentiation
SE2 + 80 days
Fixed for IHC
Stage 3. Terminal differentiation
SE3 + 70 days
Fixed for IHC

Figure 3.2B

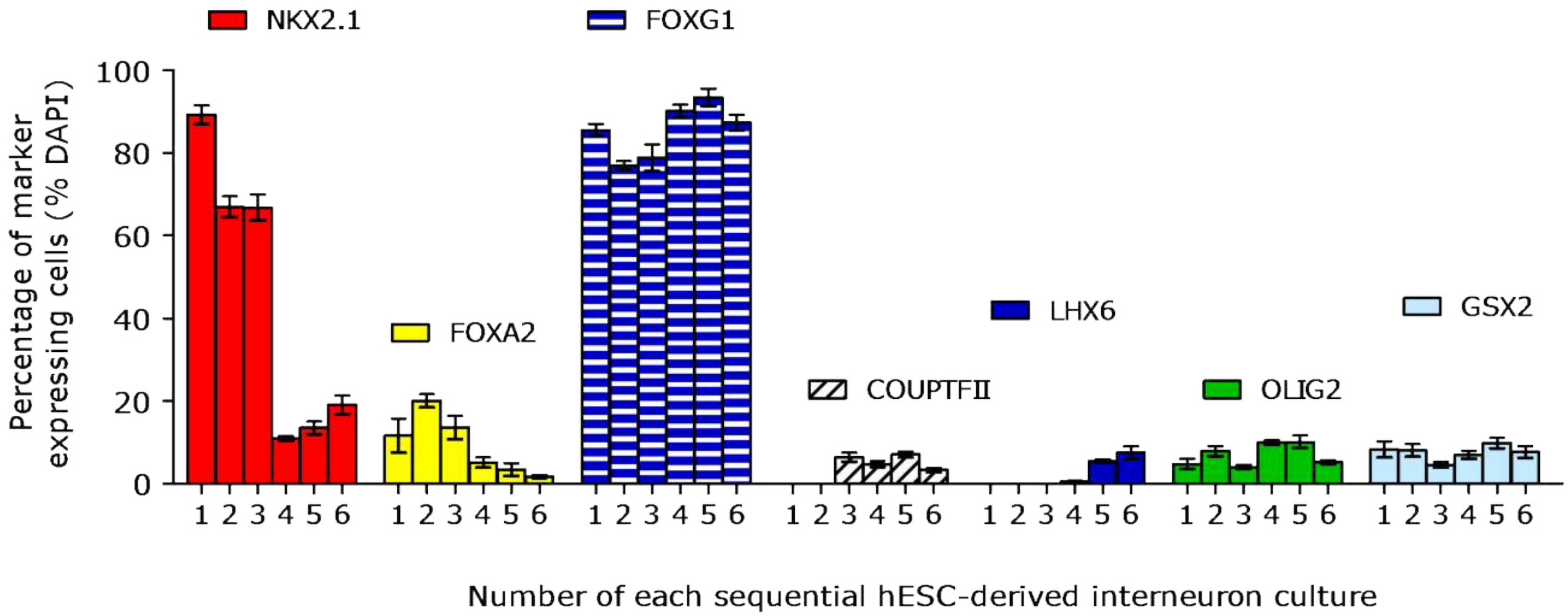

NKX2.1
FOXG1
FOXA2
COUPTFII
LHX6
OLIG2
GSX2
Percentage of marker expressing cells (% DAPI)
0
20
40
60
80
100
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
1 2 3 4 5 6
Number of each sequential hESC-derived interneuron culture

Figure 3.2C

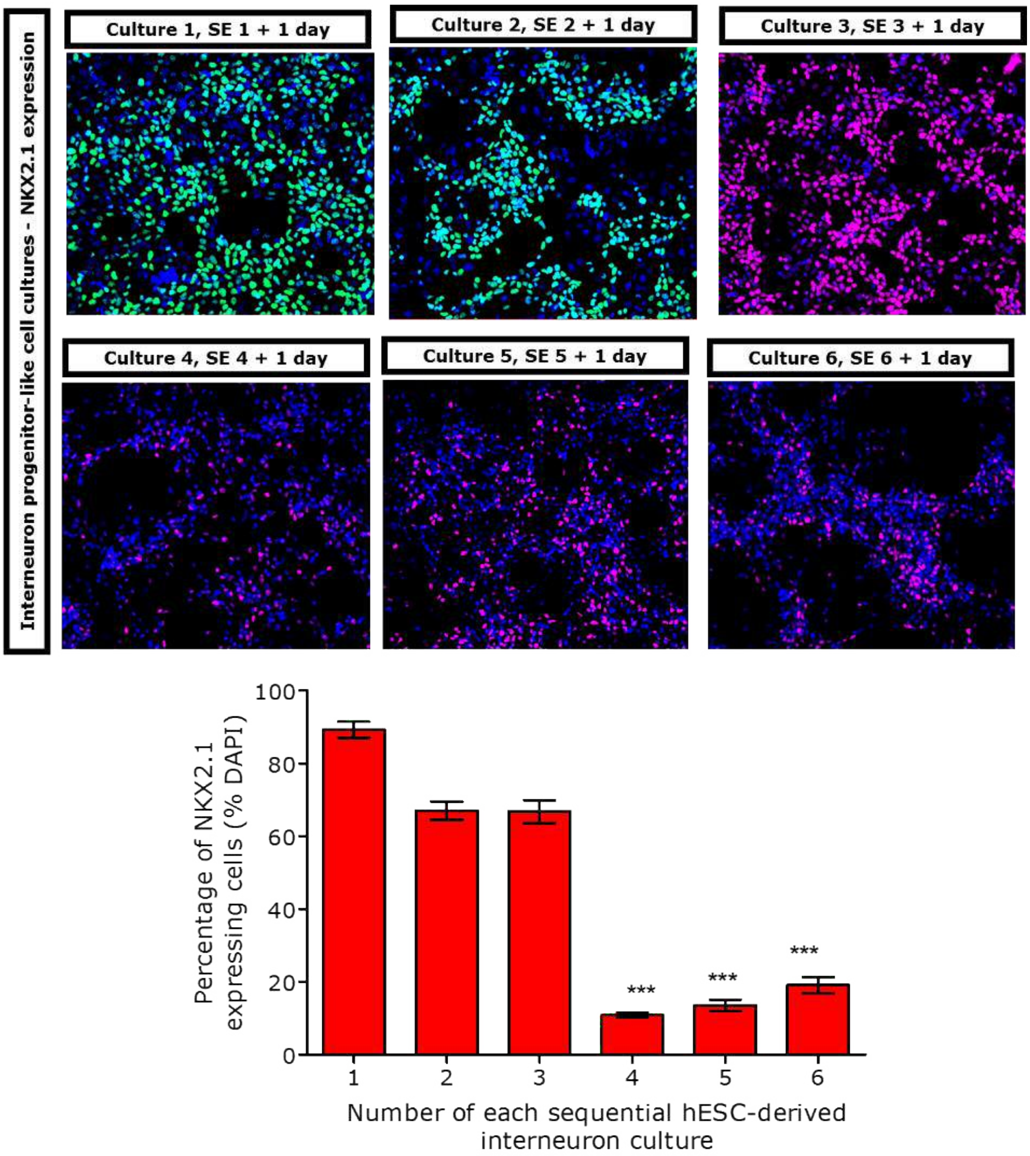

Interneuron progenitor-like cell cultures - NKX2.1 expression
Culture 1, SE 1 + 1 day
Culture 2, SE 2 + 1 day
Culture 3, SE 3 + 1 day
Culture 4, SE 4 + 1 day
Culture 5, SE 5 + 1 day
Culture 6, SE 6 + 1 day
Percentage of NKX2.1 expressing cells (% DAPI)
100
80
60
40
20
0
***
***
***
1
2
3
4
5
6
Number of each sequential hESC-derived interneuron culture

Figure 3.2D

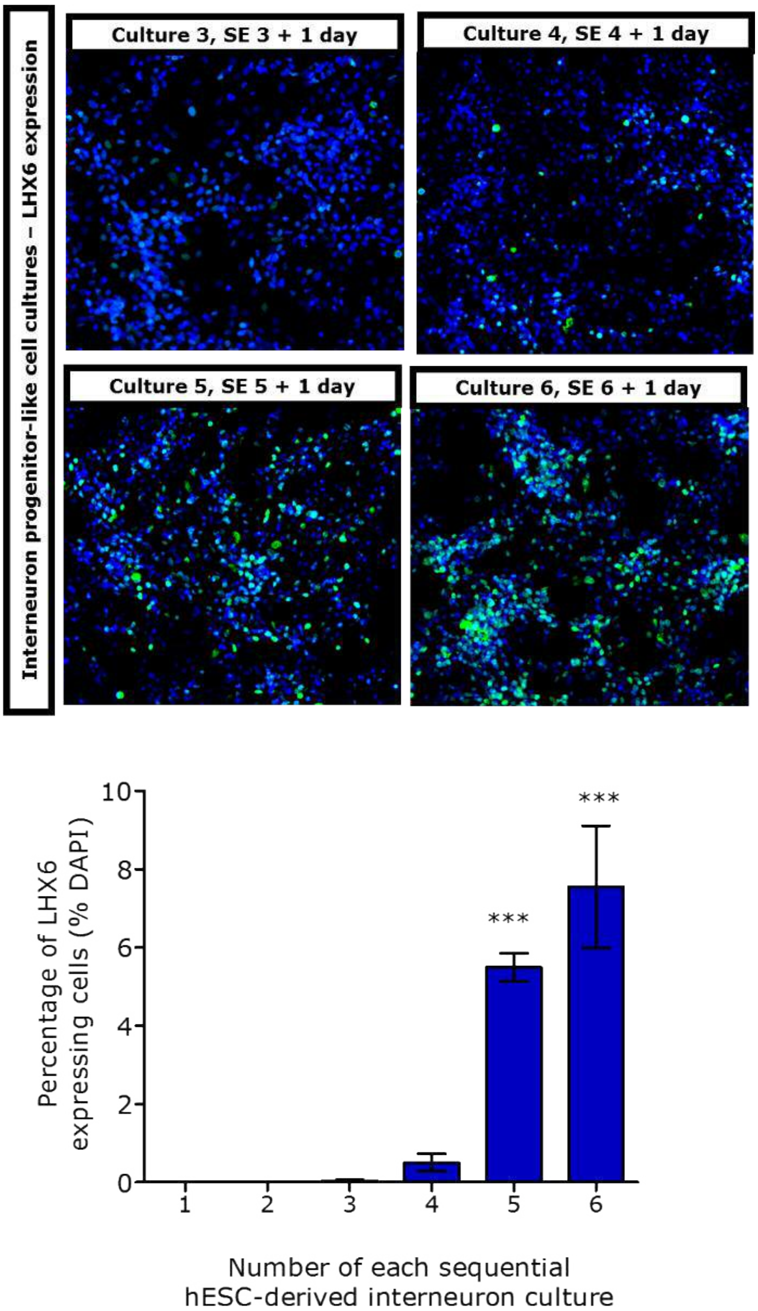

Interneuron progenitor-like cell cultures – LHX6 expression
Culture 3, SE 3 + 1 day
Culture 4, SE 4 + 1 day
Culture 5, SE 5 + 1 day
Culture 6, SE 6 + 1 day
Percentage of LHX6 expressing cells (% DAPI)
10
8
6
4
2
0
***
***
1
2
3
4
5
6
Number of each sequential hESC-derived interneuron culture

Figure 3.2E

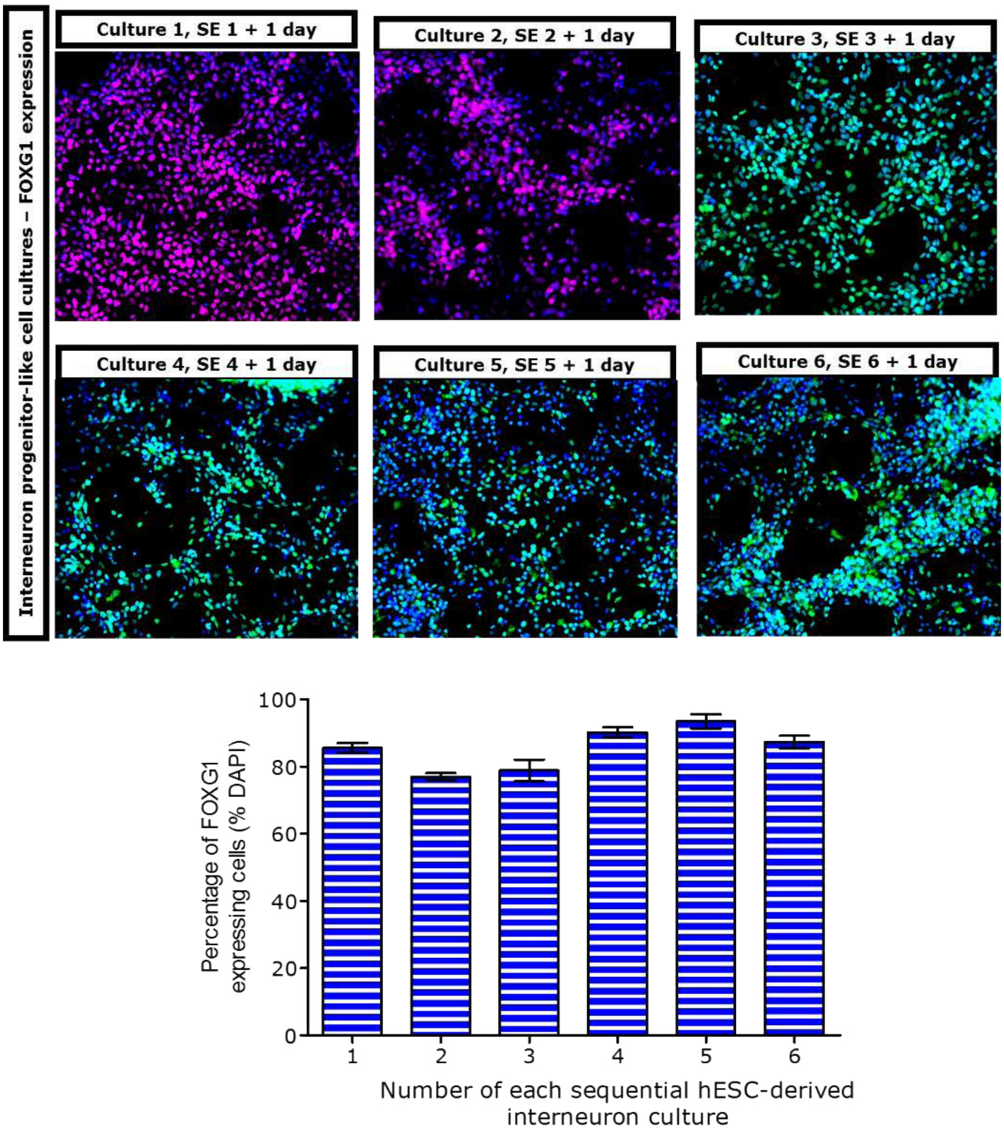

Interneuron progenitor-like cell cultures – FOXG1 expression
Culture 1, SE 1 + 1 day
Culture 2, SE 2 + 1 day
Culture 3, SE 3 + 1 day
Culture 4, SE 4 + 1 day
Culture 5, SE 5 + 1 day
Culture 6, SE 6 + 1 day
Percentage of FOXG1 expressing cells (% DAPI)
100
80
60
40
20
0
1
2
3
4
5
6
Number of each sequential hESC-derived interneuron culture

Figure 3.2 Effect of sequential serial expansions of hESC-derived interneuron progenitors on molecular profile of progenitors.

A) A schematic of the modifications made to hESC interneuron differentiation protocol in order to generate sequential hESC-derived interneuron cultures (1 to 6) through serial expansions into interneuron differentiation media (10 day SHH treatment). The passaging of the interneuron progenitors was repeated every 10 days to generate 6 temporally distinct cultures. Molecular characterisation of the progenitor plates was performed 1 day post serial expansion back into SHH treatment (SE + 1). The day immunohistochemistry staining was performed on terminally differentiated plates was day 110.

B) Molecular characterisation of each of the 6 hESC-derived interneuron progenitor cultures–protein level, 1 day post each serial expansion. The immunohistochemistry data has been presented as a percentage of marker expressing cells out of the total cell population (total cell count measured using DAPI count). Average total cell count per technical replicate can be found in Appendix folder 7. The graph in figure 3.2 shows the mean percentage of marker expressing cells for each sequentially derived culture (1 to 6), along with +/- SEM. The actual day in vitro each culture was sampled was day 21 culture 1, day 31 culture 2, day 41 culture 3, day 51 culture 4, day 61 culture 5 and day 71 culture 6. Only data from the final 4 cultures (cultures 3 to 6) were available for COUPTF II and LHX6 markers due to antibody issues. Biological replicates n=1. Minimum technical replicates n=6.

C-E) Molecular characterisation of each of the hESC-derived interneuron progenitor cultures–protein level, 1 day post each serial expansion. Only data from final 4 cultures (culture 3 to 6) were available for LHX6 marker. The immunohistochemistry data has been presented in the form of a panel of immunohistochemistry, deconvoluted images taken from each sequential culture of NKX2.1 (3.2C), LHX6 (3.2D) and FOXG1 (3.2E) expressing cells. The graph plotted shows the mean percentage of NKX2.1 (3.2C), LHX6 (3.2D) and FOXG1 (3.2E) expressing cells out of total cell population (total cell count measured using DAPI count, blue) for each time point, along with +/- SEM.). Average total cell count per technical replicate can be found in Appendix folder 7. Biological replicates n=1. Minimum technical replicates n=6.

Statistics performed to determine significant differences in percentage of marker expressing cells, for each marker individually over time relative to culture 1 (standard paradigm) using nonparametric, Kruskal-Wallis test and Dunn's Multiple Comparison Test (*$p<0.05$, **$p<0.01$, ***$p<0.001$). NK2 Homeobox 1, NKX2.1. FOXA2, Forkhead Box A2. FOXG1, Forkhead Box G1. COUP transcription factor 2, COUPTF II. LHX6, LIM Homeobox 6. OLIG2, oligodendrocyte transcription factor 2. GSX2, GS Homeobox 2. hESC, human embryonic stem cell. SHH, Sonic Hedgehog. SE, serial expansion. SEM, standard error of the mean.

3.3.5 Sequentially generated cultures showed progressively elevated SST: CR on terminal differentiation

The serial expansions into stage 2 (SHH treatment) were terminated after seeding culture 6 for terminal differentiation on day 70, due to the limited proliferation observed. The resulting 6 cultures generated by sequential serial expansions were then terminally differentiated until day 110 in vitro. Although all cultures had the same total duration in vitro, the time in stage 3 (neuronal maturation media) progressively decreased for each sequential culture from a maximum of 90 days for culture 1 to a minimum of 40 days for culture 6 (Figure 3.2A).

The proportion of post-mitotic interneuron markers, GAD1, LHX6 and PVALB remained unchanged between sequentially generated cultures 1 to 6 (Figure 3.3A). The high abundance of GAD1 positive cells (>75%) after terminal differentiation would suggest that serial expansion did not affect the efficiency of the progenitors to give rise to GABAergic neurons (Figure 3.3A). Although the later derived cultures did not show a bias towards PVALB lineage, with PVALB positive cells rare in all cultures (<5%), other subtype markers were significantly differentially expressed across cultures (Figure 3.3B-3.3E). Firstly, the proportion of post-mitotic interneuron marker CALB1 positive cells significantly declined in cultures generated from later serial expansions (culture 3 onwards) (Figure 3.3C). A similar reduction was observed for CR positive interneurons across later cultures, from 15% to 8% (***$p<0.001$ Figure 3.3D). Whereas, SST positive interneurons significant increased in abundance, although only peaking at 7% in culture 6 population (***$p<0.001$ Figure 3.3E). Although the rise in SST interneurons was not matched by an increase in LHX6 positive cells in the terminal differentiated cultures, the cultures with significantly higher proportion of SST expressing cells did have significantly greater LHX6 and lower NKX2.1 positive population when seeded for stage 3 (***$p<0.001$ Figure 3.2C-3.2D and 3.3E). This would support the conclusion that the progressively elevated SST:CR in the cultures was driven by changes in the lineage commitment of the progenitors from CR to SST fate. However, these experiments cannot rule out the effect of subtype-selective survival bias, which could have been driven by the serial expansions and/or differential time in neuronal maturation media (Figure 3.2A). In either case, sequential serial expansions gave rise to a more defined population, enriched in SST interneurons.

Figure 3.3A

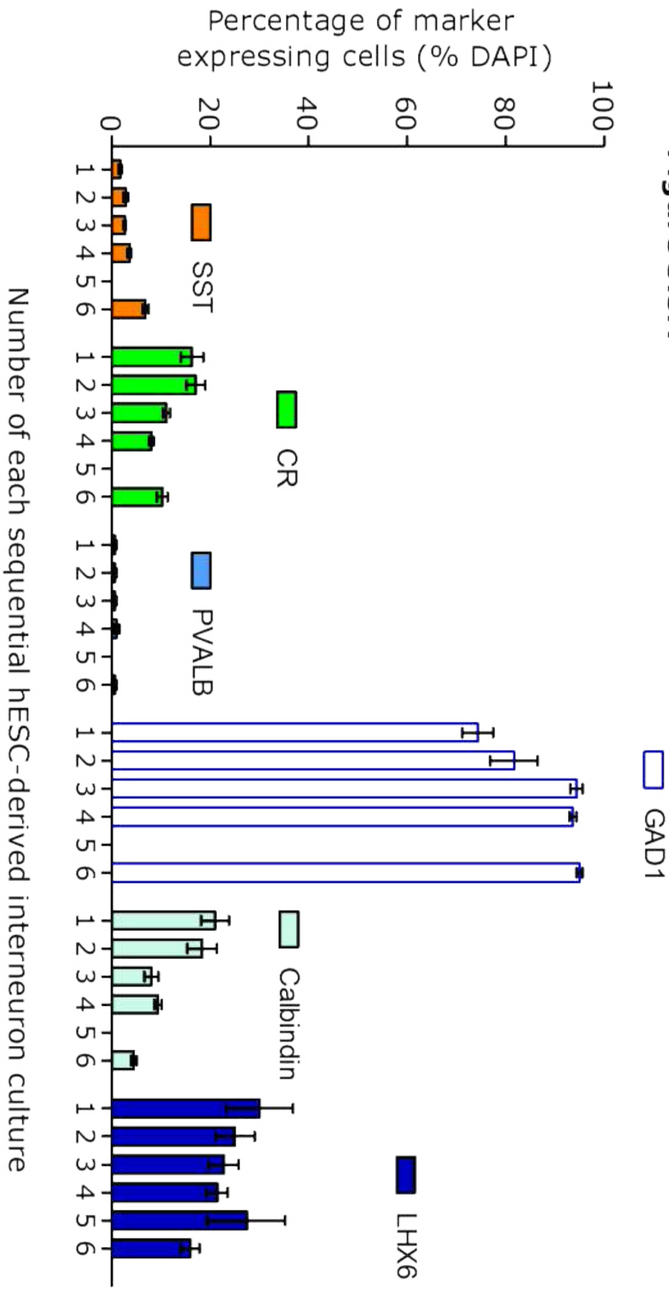


Figure 3.3B

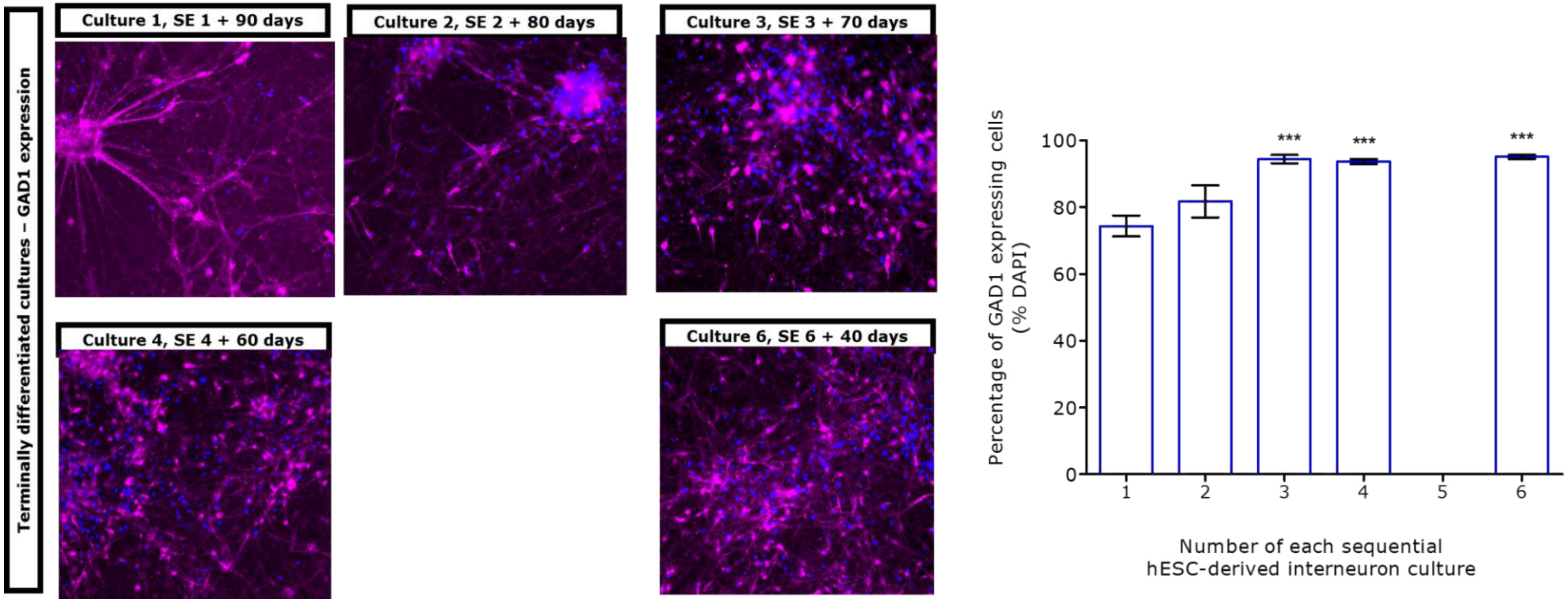

Terminally differentiated cultures – GAD1 expression
Culture 1, SE 1 + 90 days
Culture 2, SE 2 + 80 days
Culture 3, SE 3 + 70 days
Culture 4, SE 4 + 60 days
Culture 6, SE 6 + 40 days
Percentage of GAD1 expressing cells
(% DAPI)
100
80
60
40
20
0
***
***
***
1
2
3
4
5
6
Number of each sequential
hESC-derived interneuron culture

Figure 3.3C

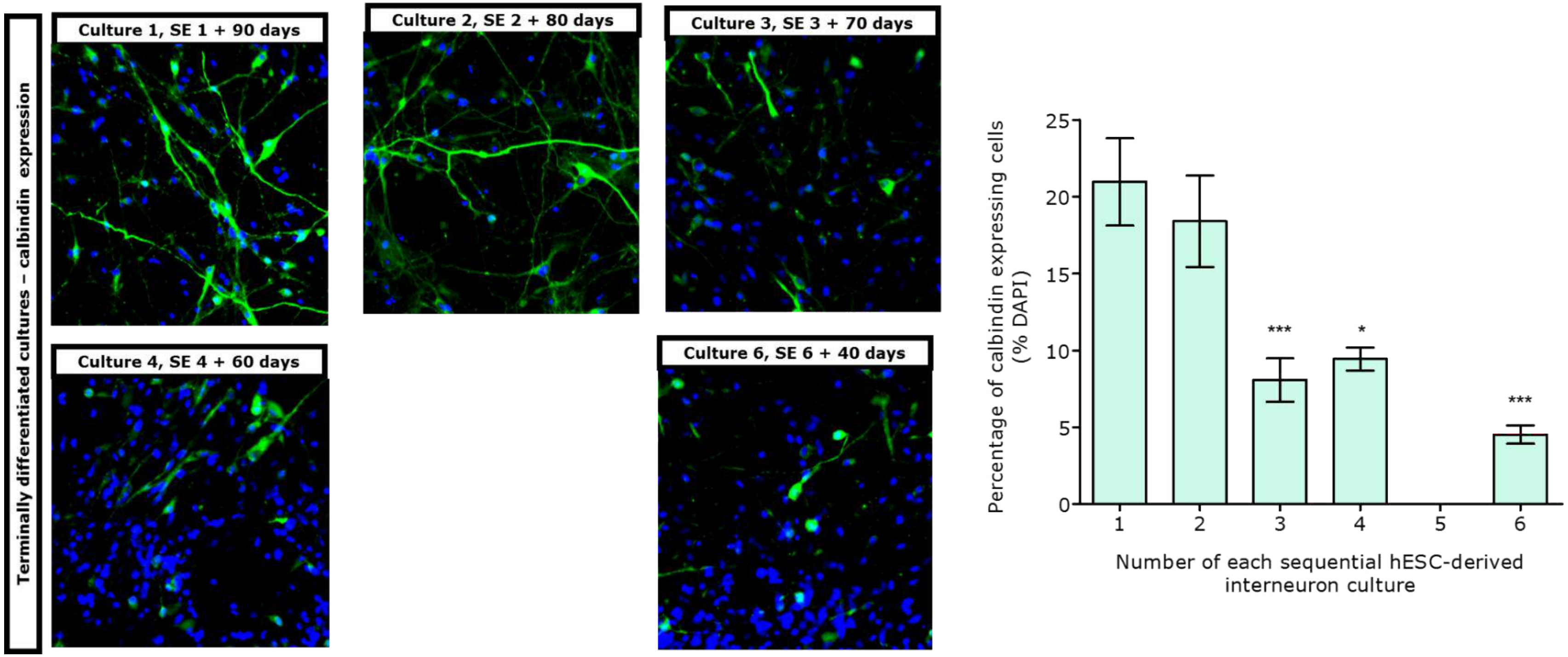

Terminally differentiated cultures – calbindin expression
Culture 1, SE 1 + 90 days
Culture 2, SE 2 + 80 days
Culture 3, SE 3 + 70 days
Culture 4, SE 4 + 60 days
Culture 6, SE 6 + 40 days
Percentage of calbindin expressing cells (% DAPI)
0
5
10
15
20
25
***
*
***
1
2
3
4
5
6
Number of each sequential hESC-derived interneuron culture

Figure 3.3D

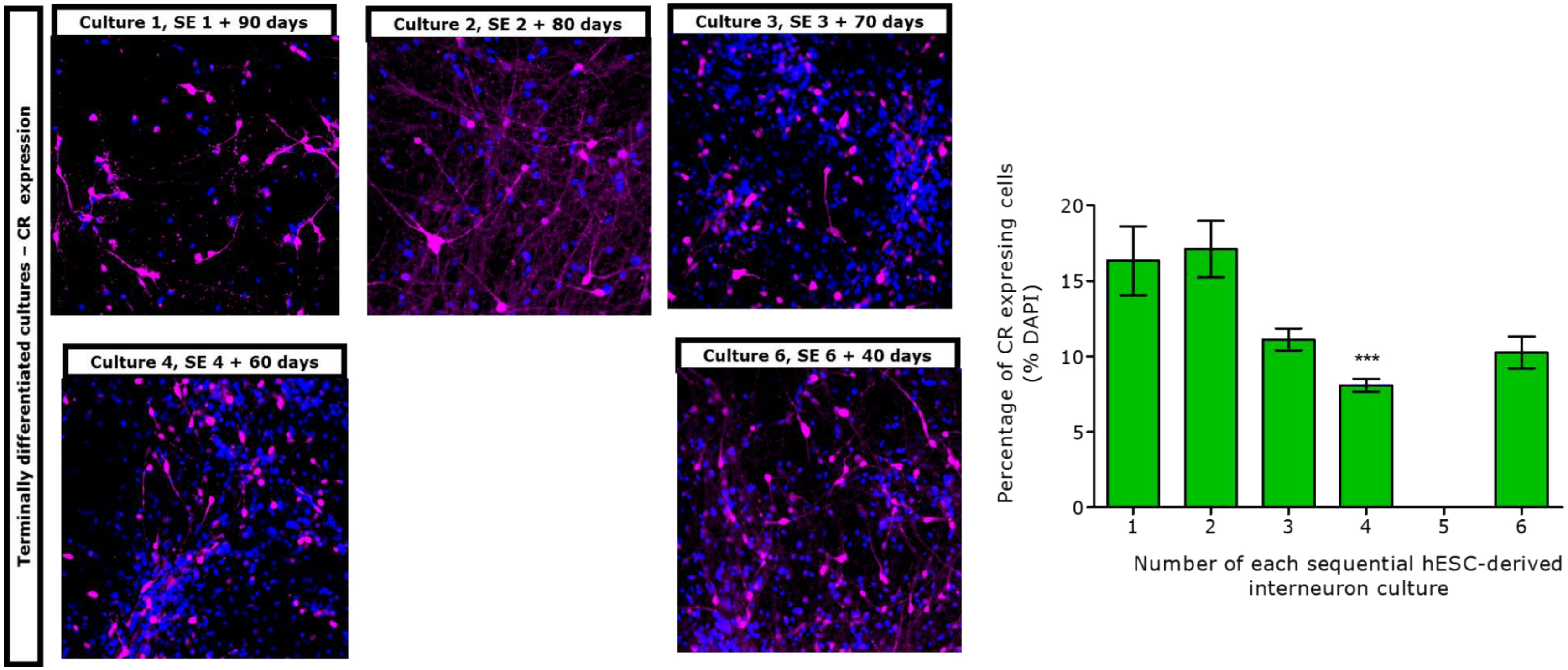

Terminally differentiated cultures – CR expression
Culture 1, SE 1 + 90 days
Culture 2, SE 2 + 80 days
Culture 3, SE 3 + 70 days
Culture 4, SE 4 + 60 days
Culture 6, SE 6 + 40 days
Percentage of CR expresing cells
(% DAPI)
0
5
10
15
20
1
2
3
4
5
6
***
Number of each sequential hESC-derived
interneuron culture

Figure 3.3E

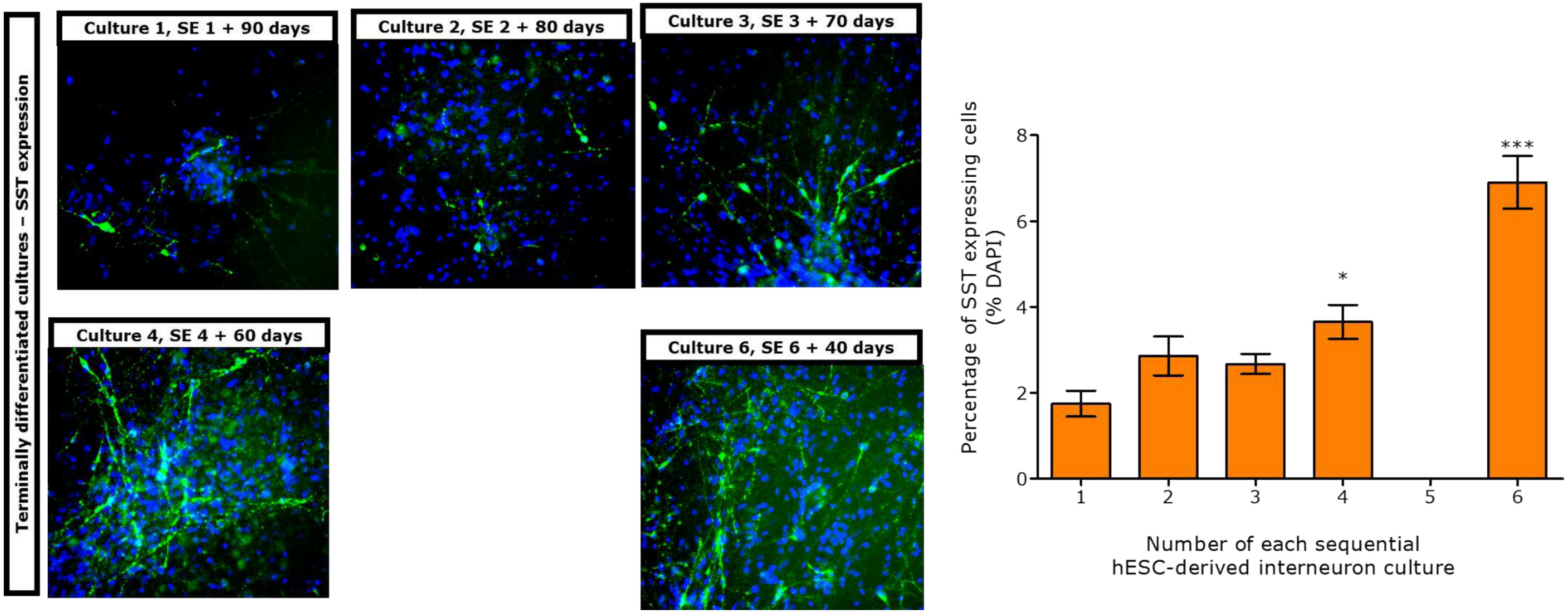

Terminally differentiated cultures – SST expression
Culture 1, SE 1 + 90 days
Culture 2, SE 2 + 80 days
Culture 3, SE 3 + 70 days
Culture 4, SE 4 + 60 days
Culture 6, SE 6 + 40 days
Percentage of SST expressing cells (% DAPI)
0
2
4
6
8
*
***
1
2
3
4
5
6
Number of each sequential hESC-derived interneuron culture

Figure 3.3 Effect of sequential serial expansions of hESC-derived interneuron progenitors on molecular profile of the post-mitotic interneuronal population.

A) Molecular characterisation of each of the 6 hESC-derived interneuron cultures generated through serial sequential expansions of the progenitors–protein level, at total of 110 days in vitro. The immunohistochemistry data has been presented as percentage of marker expressing cells out of the total cell population (total cell count measured using DAPI count). The graph plotted shows the mean percentage of marker expressing cells for each time point, along with +/- SEM. Culture 5 (SE5 +50 days) results are missing due to detachment of culture on fixing for all markers except LHX6. Average total cell count per technical replicate can be found in Appendix folder 7. Biological replicates n=1. Minimum technical replicates n=6.

B-E) Molecular characterisation of each of the hESC-derived interneuron cultures–protein level, at 110 days in vitro. Only LHX6 marker expression data is available for culture 5 (SE5 seeded for terminal differentiation on day 60). The immunohistochemistry data has been presented in the form of a panel of immunohistochemistry, deconvoluted images taken from each sequential culture of GAD1 (3.3B), calbindin (3.3C), CR (3.3D) and SST (3.3E) expressing cells. The graph plotted shows the mean percentage of GAD1 (3.3B), Calbindin (3.3C), CR (3.3D) and SST (3.3E) expressing cells out of total cell population (total cell count measured using DAPI count, blue) for each time point, along with +/- SEM. All stainings were carried out on day 110 cultures. Average total cell count per technical replicate can be found in Appendix folder 7. Biological replicates n=1. Minimum technical replicates n=6.

Statistics performed to determine significant differences in percentage of marker expressing cells, for each marker individually over time relative to culture 1 (standard paradigm) using nonparametric, Kruskal-Wallis test and Dunn's Multiple Comparison Test (*$p<0.05$, **$p<0.01$, ***$p<0.001$). LHX6, LIM Homeobox 6. SST, somatostatin. PVALB, parvalbumin. CR, calretinin. GAD1, glutamate decarboxylase 1. hESC, human embryonic stem cell. SEM, standard error of the mean.

3.4 Discussion

3.4.1 Cross-comparison of hPSC-interneuron differentiation paradigms

Through the series of experiments described in this chapter evidence of the efficient generation of interneuron-like cells has been provided, both at the RNA and protein level. Yet how does the quality of this hPSC-interneuron differentiation protocol fair to others? Within the literature, there are several independent reports that characterise interneuron-like cells derived from hPSC (Cambray *et al.* 2012; Maroof *et al.* 2013; Colasante *et al.* 2015; Meganathan *et al.* 2017). These paradigms differ on multiple factors, including, but not limited to: cell line used, co-culture conditions, extrinsic or intrinsic cue methodology, and in vivo vs in vitro terminal differentiation (Cambray *et al.* 2012; Maroof *et al.* 2013; Colasante *et al.* 2015; Meganathan *et al.* 2017). In this study the interneuron differentiation used extrinsic cues to generate an abundant MGE-like progenitor population, with a small proportion of CGE-like progenitors (and even smaller proportion of alternative lineages). Whereas, others have used similar interneuron differentiation media, but selectively enriched the culture for MGE-like progenitors using MGE specific reporter cell lines (Maroof *et al.* 2013; Tyson *et al.* 2015; Xiang *et al.* 2017). This focus would account for a greater proportion of MGE-derived SST and PVALB interneurons reported in these studies, and why CGE-derived subtypes were left unreported (Maroof *et al.* 2013; Tyson *et al.* 2015; Xiang *et al.* 2017). Here a wider diversity of subtypes (ChAT, SST, PVALB and CR) were detected through bulk-QPCR and immunostaining. Two of the most abundant subtypes generated in vitro were CR and ChAT interneuron-like cells. Both subtypes are enriched in the human striatum, which would account for the proportion of NKX2.1 positive cells in the post-mitotic population (Marín *et al.* 2000; Wu & Parent 2000; Batista-Brito *et al.* 2008; Nóbrega-Pereira *et al.* 2008). The maintained NKX2.1 expression marks MGE-derived interneuron precursors for striatal rather than cortical destination in vivo (Marín *et al.* 2000; Wu & Parent 2000; Batista-Brito *et al.* 2008; Nóbrega-Pereira *et al.* 2008). Moreover, the bias towards CR and ChAT lineage commitment was, in part, supported by the molecular profile of the early and late interneuron precursors, which had high expression of *LHX8* and *COUPT II*, and low expression of SST/PVALB interneuron precursor marker *LHX6* (Tyson *et al.* 2015; Hu *et al.* 2017a; Elbert *et al.* 2019). Whereas, others have achieved higher *LHX6* expression which has been followed by higher abundance in SST interneuron-like cells in the post-mitotic culture (Maroof *et al.* 2013; Close *et al.* 2017; Yuan *et al.* 2018). In fact, paradigms combining the use of SHH signalling and overexpression of *LHX6* selectively enhanced the efficiency of generating SST interneurons in vitro (5% using extrinsic cues to 12% with *LHX6* overexpression and extrinsic cues) (Yuan *et al.* 2018). Therefore, the failure of the extrinsic cues to induce *LHX6* expression likely accounts for the low abundance of PVALB and SST interneurons in the culture.

3.4.2 The differentiation trajectory of interneuron-like cells

The longitudinal molecular characterisation of interneuron-like cells across differentiation revealed the temporal dynamics in key transcription factors, as well as subtype markers (*SST, CR, PVALB*). Together these findings would indicate that although early MGE-like progenitors were efficiently generated, the differentiation paradigm failed to maintain and enhance the expression of the late MGE markers (*LHX6* and *LHX8*). Therefore, although a high proportion of interneuron-like cells were generated using this paradigm, one would question the full authenticity of the cells compared to their native counterparts. This is likely due to exogenous SHH signalling alone not being sufficient to recapitulate all extrinsic cues interneurons are subject to across their differentiation trajectory. Therefore, introducing additional extrinsic cues (such as BMP and fibroblast growth factor, FGF signalling) may help drive and maintain the expression of these later markers, by reproducing some of the signals received during migration to the region of destination (Mukhopadhyay *et al.* 2009; Danjo *et al.* 2011). Moreover, the use of additional neurogenic and neurotrophic factors during terminal differentiation may help selectively promote neuronal survival and maturation, and potentially shorten the duration of the terminal differentiation stage (Telezhkin *et al.* 2016).

3.4.3 Extended production of interneuron-like progenitors

During development temporally distinct waves of progenitors are born within the ventricular zone of the MGE/CGE. The ventricular zone is maintained over an extended period of development, only beginning to significantly shrink from 14WPC onwards (Hansen *et al.* 2013). However, the extrinsic cues required to maintain proliferating interneuron progenitors remains unclear. In attempt to extend production of proliferating hESC-derived interneuron progenitors in vitro, sequential serial expansions of cultures maintained in SHH treatment was performed. Although this strategy successfully generated newborn interneuron-like progenitors, this success was restricted to the first three serial expansions. The maintenance of proliferating progenitors (although short-lived) reflects the mitogen activity of SHH, which acts to increase the expression of cell cycle regulatory proteins (including cyclin D1 *CCND1*, E2F Transcription Factor *E2F1/2*) and inhibit differentiation (Fuccillo *et al.* 2006; Elkabetz *et al.* 2008; Cambray *et al.* 2012). However, SHH signalling alone was not sufficient to maintain proliferating interneuron-like progenitors over extended periods. This would encourage investigating whether additional mitogens, potentially FGF-2 and Notch signalling, could work synergistically with SHH to lengthen the duration in vitro that the progenitors could be maintained (Okabe *et al.* 1996; Elkabetz *et al.* 2008).

3.4.4 Early vs late-born interneuron-like progenitors

The depletion of proliferating progenitors in the later cultures altered the goal of the serial expansion paradigm. Instead of generating newborn progenitors, earlier-born MGE-like progenitors transitioning to post-mitotic state (LHX6 positive, NKX2.1 negative) were maintained under SHH signalling before seeding for in vitro maturation. With each expansion the proportion of earlier-born (LHX6 positive) interneuron precursors progressively increased, as did the duration the interneuron precursors were cultured in SHH before terminal differentiation. Although unintended this had generated a situation similar to the one reported by Tyson *et al.* (Tyson *et al.* 2015). However, Tyson *et al.* directly isolated these populations using a dual MGE specific reporter mESC line (*Nkx2.1::mCherry, Lhx6::GFP*) and transplanted each enriched population into early postnatal mouse cortex (Tyson *et al.* 2015). After completing terminal differentiation in vivo, earlier-born (*Lhx6::GFP* positive)

MGE-like progenitor-derived grafts preferentially generated SST interneuron subtypes compared to late-born (*Nkx2.1::mCherry*) MGE-like progenitors (Tyson *et al.* 2015). These findings by Tyson *et al.* were, in part, in agreement to the significantly higher abundance of SST interneurons within later derived cultures, which increased in a graded manner with serial expansions (Tyson *et al.* 2015; Xiang *et al.* 2017). However, no change was observed in PVALB interneuron abundance, which was low across all cultures. The difficulty in modifying PVALB subtype abundance during interneuron differentiation of hPSC through SHH manipulation has been reported previously, and may be due to the low efficiency that interneuron differentiation paradigms have in generating PVALB interneurons in vitro (Close *et al.* 2017; Xiang *et al.* 2017). Instead, only the loss of CR and CALB1 interneuron-like cells was detected.

3.4.5 Mechanisms underpinning the elevated SST: CR interneuron

The elevated SST:CR in the terminally differentiated later derived cultures could be explained by alternative mechanisms. First, the serial expansion of the culture and extended duration under exogenous SHH signalling changed the lineage commitment of interneuron precursors from CR to SST. However, this would require interneuron fate commitment of interneurons derived from the same ancestor (part of the same clonal lineage) to be unrestricted. With conflicting evidence concerning the flexibility of interneuron lineage commitments within the literature, this mechanism remains under scrutiny (Rolando *et al.* 2010; Ishino *et al.* 2017; Quattrocolo *et al.* 2017). Moreover, the shift towards SST lineage in later cultures would not have been expected from mouse CGE and MGE fate-mapping studies (Butt *et al.* 2005; Inan *et al.* 2012). Mouse CGE and MGE grafts obtained at E13.5 and E15.5 showed a differential bias towards CR and PVALB interneuron (in particular PVALB Chandelier subtype) commitment in the older grafts (Butt *et al.* 2005; Inan *et al.* 2012). The inability to reproduce this result in hPSC-derived interneuronal model may be due to between-species differences, or due to terminal differentiation of sequentially generated cultures failing to recapitulate the temporally induced change in lineage commitment of the progenitors. Instead, the serial expansion and shorter duration in the neuronal maturation media, may have resulted in the selective survival bias in favour of SST interneuron precursors, and against CR interneuron precursors. This theory is, in part, supported by an independent study, using a similar hPSC interneuron differentiation paradigm, which has shown that extended durations in the neuronal maturation media (up to 100 days) significantly depleted the culture of GAD1 and SST positive cells (Close *et al.* 2017). Nevertheless, no change was reported for the proportion of CR positive cells (Close *et al.* 2017). Therefore, the extended duration in neuronal maturation media (e.g. 90 days for culture 1) alone cannot account for the higher proportion of CR interneuron-like cells in these cultures.

3.4.6 Conclusions

Overall, serial expansion of interneuron-like progenitors holds promise as a means to simulate the endogenous generation of temporally distinct progenitor pools, and thereby expanding the spectrum of interneuron subtypes derived from a single original parent culture. However, certain optimisations to the design would be required to enhance proliferation, and to enable these populations to be monitored in real-time to better guide manipulation of early vs late-born progenitor ratio (e.g. dual reporter cell line) (Tyson *et al.* 2017). However, previous work within the literature has indicated a potential negative impact of repeated passaging of neuronal progenitors on the quality of neurons generated on terminal differentiation (Bardy *et al.* 2016). With each serial expansion, the neuronal cultures derived were progressively less active and displayed more immature spike patterns (Bardy *et al.* 2016). Therefore, conducting electrophysiology on each temporally distinct culture would determine whether repeated passaging is a feasible and advantageous approach to take to generate interneuron-like cells that are both a realist molecular and functional model of their authentic counterparts.

Chapter 4 - Single-cell transcriptome analysis of hESC-derived interneurons

4.1 Introduction

Single-cell transcriptomics remains a hot topic within science due to the broad utility of this innovative technique, which has started to untangle the puzzling threads of research within multiple fields of study (Drissen *et al.* 2016; Poulin *et al.* 2016; Close *et al.* 2017; Azizi *et al.* 2018). Over the years following the publication of the first single-cell transcriptomic protocol, further refinements have been reported, providing users with an array of choices on reaction chemistry and single-cell capture technique to enhance both the sensitivity and throughput respectively (Tang et al. 2009; Hashimshony et al. 2012; Ramsköld et al. 2012; Poulin *et al.* 2016).

4.1.1 Single-cell transcriptomics – applications

Single-cell transcriptomics has been an invaluable tool to researchers due to the ability to uncover data previously hidden in bulk-level analysis of heterogeneous samples (Wills *et al.* 2013; Liu & Trapnell 2016; Ofengeim *et al.* 2017). In particular, single-cell RNA sequencing analysis describes the entire (within detection limits) expression profile of each cell, from which valuable information can be gained. First, one can identify the cell types within the population, and potentially detect novel cell types and/or novel genes associated with the lineage (Liu & Trapnell 2016; Boldog *et al.* 2018). Second, one can tease apart the molecular heterogeneity within and across samples, thereby identifying the differentially expressed genes responsible for this heterogeneity (Liu & Trapnell 2016). This has been particularly useful when characterising the differentiation trajectories of cells during development, and has led to candidate lineage determining transcription factors being named (Close *et al.* 2017; Xiang *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018). Finally, one can detect gene expression patterns that may represent putative gene regulatory networks/nodes (Camp *et al.* 2015; Liu & Trapnell 2016).

4.1.2 Single-cell transcriptomics – interneurons

Overall these applications of single-cell transcriptomics makes the tool ideally suited to teasing apart the genetic heterogeneity of interneurons (Darmanis *et al.* 2015; Lake *et al.* 2016; Boldog *et al.* 2018). The in-depth profiling capable with single-cell RNA sequencing has expanded the molecular characterisation of interneurons in the brain, leading to novel species-dependent subtypes being detected, along with novel interneuron markers (e.g. prepronociceptin *Pnoc*) (Zeisel *et al.* 2015; Lake *et al.* 2016; Boldog *et al.* 2018). Moreover, between-region analysis has enabled region-specific subtypes to be determined, including the mouse striatal, hippocampal (CA1 region) and cortical (primary somatosensory cortex) restricted interneuron subtypes: PVALB (*Pthlh, Pvalb* co-expressing), RELN (*Reln, Gabrd, Lhx6* co-expressing) and VIP/CR (*Vip, Cr, Penk, Crh* co-expressing), respectively (Zeisel *et al.* 2015). Finally, single-cell RNA sequencing analysis performed on primary mouse cortical tissue has detected genes that are selectively expressed by interneurons in particular cortical layers, so-called layer markers (Darmanis *et al.* 2015; Lake *et al.* 2016; Frazer *et al.* 2017; Nowakowski *et al.* 2017). These markers include tachykinin 3 (*TAC3*) and *SV2C* (upper layer markers), and MEIS homeobox 2 (*MEIS2)* (deep layer and white matter marker) (Darmanis *et al.* 2015; Lake *et al.* 2016; Frazer *et al.* 2017; Nowakowski *et al.* 2017). This has provided evidence of an intrinsic regulation of cortical interneuron lamination, which likely interplays with contact-dependent extrinsic cues from the principal cells of the cortex (Nowakowski *et al.* 2017; Wester *et al.* 2019).

Moreover, longitudinal single-cell transcriptomics has the potential to decipher the genetic basis of interneuron lineage divergence (Darmanis *et al.* 2015; Lake *et al.* 2016; Chen *et al.* 2017; Boldog *et al.* 2018; Mayer *et al.* 2018; Mi *et al.* 2018). Through independent single-cell RNA sequencing study on mouse MGE and CGE, MAF BZIP transcription factor *(Maf)* and myocyte enhancer factor 2C *(Mef2c)* have been named as SST and PVALB interneuron lineage determining transcription factors respectively (Mayer *et al.* 2018; Mi *et al.* 2018). Although these findings in mice are promising, issues concerning between-species differences in interneuron development may impair the translation of these finding to humans (Wu & Parent 2000; Bakken *et al.* 2016; Lake *et al.* 2016; Sousa *et al.* 2017; Boldog *et al.* 2018).

4.1.3 Single-cell transcriptomics – hPSC-derived interneurons

However, due to the scarcity and legal age restrictions on primary human foetal tissue usage, obtaining an unrestricted outlook of interneuron differentiation trajectory from human primary tissue has not been possible. This has promoted single-cell RNA sequencing to be performed on hPSC models of interneuron differentiation (Close *et al.* 2017; Xiang *et al.* 2017). However, our understanding of interneuron progenitor and subtype diversity generated using hPSC interneuron differentiation paradigms is incomplete. At present, hPSC-derived interneurons have only been independently profiled, at single-cell resolution, by two studies (Close *et al.* 2017; Xiang *et al.* 2017). These reports have been valuable as a guide in the work presented here, encouraging an expanded number of time points to be sampled throughout the differentiation protocol, to gain a more detailed understanding of the interneuron-like cell differentiation trajectory. Currently, lineage determining transcription factors have not been named from longitudinal single-cell transcriptomics conducted on hPSC-derived interneurons, although successful findings from primary mouse MGE and CGE would encourage persistence in the search (Close *et al.* 2017; Xiang *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018).

4.2 Aims

To characterise the differentiation trajectory of hPSC-derived interneurons using longitudinal single-cell RNA sequencing.

To analyse the list of differentially expressed genes to identify any transcription factors that may be responsible for the diversification of interneuron-like cells into different subtypes.

To detect genes potentially uniquely co-expressed by SST, PVALB or CR interneuron-like cells or their precursors, and therefore may underlie the unique phenotype of these subtypes.

4.3 Results

Single-cell RNA sequencing was carried out on hESC-derived interneurons at multiple time points along the differentiation protocol to give a longitudinal description of gene expression profiles. The differentiation paradigm and time points used matched those in Chapter 3 to allow cross-technique comparison

(RNA and protein level). As described in detail in Chapter 2, Fluidigm C1 IFC was the selected platform to both isolate single cells and perform reaction chemistry (lysis, reverse transcription, preamplification). Full-length mRNA coverage and high sensitivity reaction chemistry SMART-seq was selected to allow low expressed, rare and potentially novel transcripts to be detected.

4.3.1 hESC-derived interneuron subtypes resembled those originating from dorsal MGE and CGE

From the longitudinal characterisation of hESC-derived interneurons presented in Chapter 3, initial analysis of single-cell RNA sequencing data began with a description of the *NKX2.1* expressing MGE-like progenitor population at day 20 (Figure 4.1A). In agreement with the immunostaining conducted in Chapter 3, the day 20 cultures were abundant in *NKX2.1* expressing cells (90%) (Figure 4.1E). The co-expression patterns of the *NKX2.1* expressing cells revealed a high co-expression with forebrain marker *FOXG1*, and low co-expression with late MGE markers (*LHX6/8* and *DLX1/5*) (Figure 4.1A). This would indicate that the day 20 hESC-derived progenitors were abundant in early MGE-like progenitors. Moreover, 40% of *NKX2.1* expressing cells co-expressed *COUPTF II*. The *NKX2.1/COUPTF II* population is natively restricted to dorsal MGE, suggesting a good representation of the dorsal MGE-like progenitors in the day 20 cultures. By day 65, hESC-derived interneuron culture was abundant in *MAP2* expressing cells, which expressed a diverse range of interneuron subtype markers, the most abundant including *CALB1*, *SST, CR*, and *RELN* (Figure 4.1B). Although rare within the day 65 hESC-derived interneurons (<2%), additional subtype marker expressing cells were also detected, such as CGE-derived interneuron markers *VIP, nNOS* and *5HT-R-3A*, and MGE-derived interneuron markers *SST/CR*, *PVALB* and *PVALB/CR* (Figure 4.1B). However, *CCK, NPY* and *PVALB/SST* expressing neurons were absent from the culture (Figure 4.1B). Overall, the subtype composition most closely resembled one obtained natively from the dorsal MGE and CGE.

4.3.2 hESC-derived interneurons showed subtype-dependent expression patterns

To further explore the differences in the molecular profile of *CR, SST* and *PVALB* expressing interneuron-like cells, genes uniquely co-expressed by each of the subtypes were determined from single-cell RNA sequencing data (Appendix folder 2, tables). Many of these genes included transcription coactivators (SS18 like 2, PVALB specific), cyclin proteins (cyclin I, SST specific), chaperons (heat shock protein 90 alpha family class A member 1, SST specific) and intracellular signalling proteins (G protein nucleolar 2, PVALB specific and GNAS complex locus, CR specific) (see Appendix folder 2). Although transcription factors previously named in mouse CGE/MGE single-cell RNA sequencing analysis (*MAF* and *MEF2C*) were found not to be uniquely associated with a single lineage in this search, genes regulating transcription and protein activity have been identified, and therefore may have an important role in shaping the fate and phenotype of these subtypes. Moreover, the subtype-dependent expression pattern likely accounted for (at least in part) the heterogeneity in the population, as seen in the tSNE plots at each time point (Figure 4.1C).

4.3.3 Known interneuron markers showed temporal dynamics in expression profile

tSNE plots act to visually present the heterogeneity within population, clustering cells with higher degree of similarity in expression profile closer together. An example of any genes highly expressed by a specific cluster has been displayed on each tSNE plot, with the full list of differentially expressed genes between clusters provided in Appendix 2 table 5 (Figure 4.1C). Of potential significance were the differentially expressed transcription factors (cellular retinoic acid binding protein 1 CRABP1, high mobility group box 2 HMGB2) and signalling factors (secreted protein acidic and cysteine rich, SPARC, delta like non-canonical Notch ligand 1 DLK1), which may be driving the divergence into subpopulations (Figure 4.1C). When looking across tSNE plots for each time point, more discretely segregated clusters of cells are present in the day 50 culture (Figure 4.1C). Therefore, although the tSNE plots showed heterogeneity in the population at each time point, the diversity in expression profile appeared to sharply increase between day 20 and 35, and peak at day 50 (Figure 4.1C). By studying the expression trajectory of known interneuron markers in the population, one can begin to understand the dynamics in gene expression within and across populations that may be responsible for this heterogeneity (Figure 4.1D).

The tSNE plot for day 20 showed the least distinct between clusters (Figure 4.1C). This dispersal pattern likely reflects the homogeneity of expression pattern at day 20, with 90% of the population displaying an early MGE-like co-expression pattern (NKX2.1/FOXG1 positive, LHX6/LHX8 negative) (Figure 4.1C-4.1D). However, over time the abundance of NKX2.1 expressing cells decreases to 20%, along with the NKX2.1 expression level in these cells, as shown by the population average transcripts per million (TPM) value progressively decreasing at each time point (Figure 4.1D). However, late MGE markers *LHX8* only showed a transient increase in abundance and expression level (at day 35, in agreement to bulk QPCR data, Chapter 3), but otherwise remained low across differentiation. Similarly, *LHX6* expressing cells were rare in the culture at all time points (<5%) and was only weakly expressed (average TPM value <5 at each time point) (Figure 4.1D). Whereas, the number of cells expressing *COUPTF II* progressively increased in the population until day 50, and then plateaued at 80-90%. This high proportion of *COUPTF II* expressing cells was inconsistent with findings at protein level, which reported a steady proportion of COUPTF II positive cells across day 20 to day 65 (25-30% Chapter 3). This discrepancy may underlie the additional regulatory steps between transcription and translation, and demonstrates the importance of performing analysis at both RNA and protein level. Nonetheless, these findings would suggest that the MGE-like progenitors failed to go on to express late MGE markers, and instead took on a more CGE/dorsal MGE expression pattern.

In addition to the abundant MGE-like progenitor population at day 20, there were also nascent post-mitotic neurons (*MAP2* expressing) present, which may account for the smaller, more discrete clusters on tSNE plot (Figure 4.1C). Within this post-mitotic neuronal populations, at day 20, were both *PVALB* and *CR* expressing cells (Figure 4.1E). Moreover, longitudinal analysis of the *PVALB* expressing population indicated that day 20 had the highest abundance of *PVALB* expressing cells (7%), with <2% *PVALB* expressing neurons detected in later time points (Figure 4.1E). However, PVALB interneurons remained undetected at protein level at day 20 (Chapter 3). This may be due to PVALB expression level being low (66 TMP average for day 20 population) (Figure 4.1E). Nonetheless, the co-expression pattern of PVALB expressing neurons would lend support for these cells at day 20 being true PVALB interneuron-like cells (all co-expressed *ER81, CCND1/2, DLX5, GAD1*, data not shown). Whereas, in agreement to the immunostaining for SST and CR, the proportion of *SST* and *CR* expressing cells gradually increased over the course of in vitro differentiation (SST 12% out of MAP2, CR 15% out of MAP2 Figure 4.1E). The higher expression levels of *SST* and *CR*, compared to *PVALB*, may have enabled these markers to be detected in immunostaining (Chapter 3). However, the abundance of *CR* expressing cells suggested at RNA level mismatched immunostaining for CR (underestimated proportion of CR at RNA level), which could be due to the RNA analysis being restricted to

CR/MAP2 co-expressing subpopulation (while some CR interneuron-like cells may have only been *TUJ1* expressing) (Figure 4.1E and Chapter 3).

Figure 4.1A

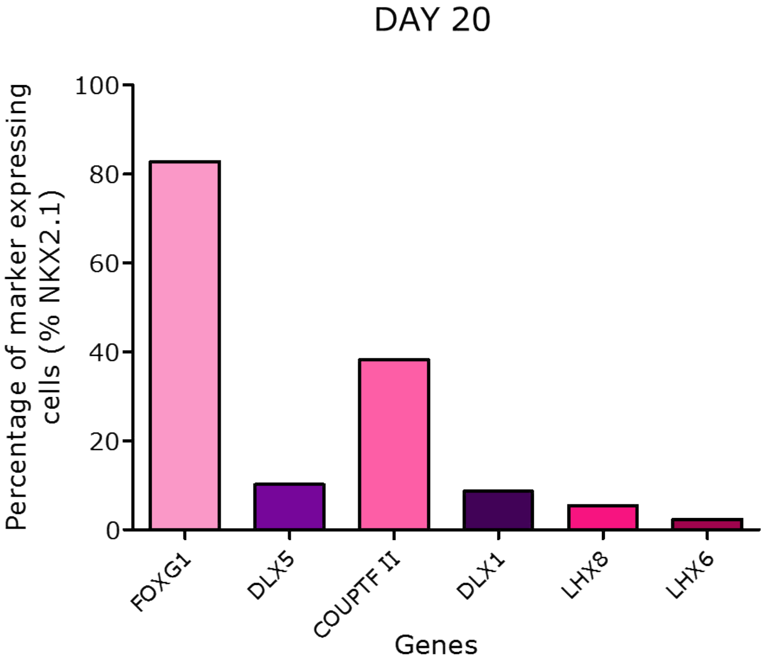


Figure 4.1B

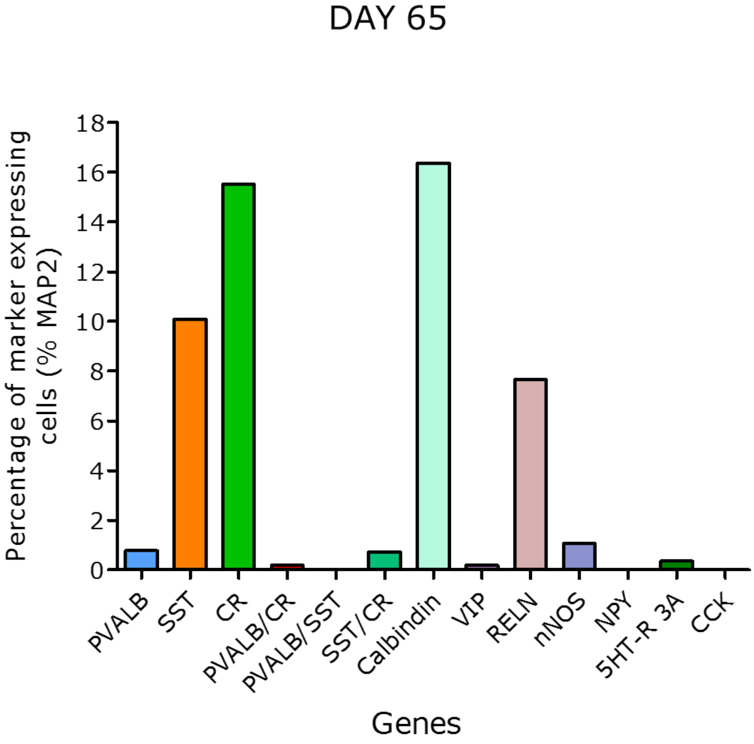


Figure 4.1C

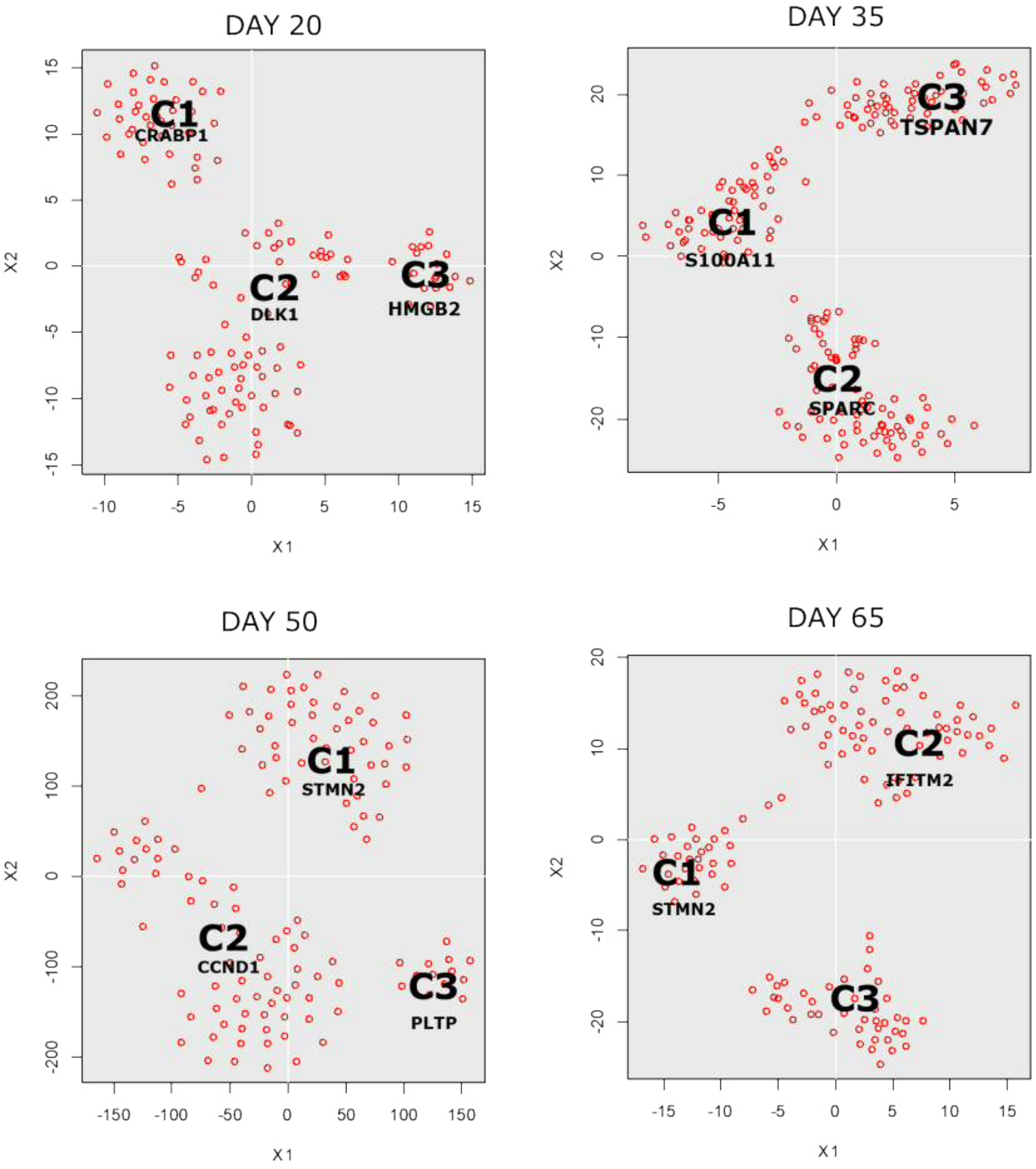

DAY 20
C1
CRABP1
C2
DLK1
C3
HMGB2
X2
15
10
5
0
-5
-10
-15
-10
-5
0
5
10
15
X1
DAY 35
C3
TSPAN7
C1
S100A11
C2
SPARC
X2
20
10
0
-10
-20
-5
0
5
X1
DAY 50
C1
STMN2
C2
CCND1
C3
PLTP
X2
200
100
0
-100
-200
-150
-100
-50
0
50
100
150
X1
DAY 65
C2
IFITM2
C1
STMN2
C3
X2
20
10
0
-10
-20
-15
-10
-5
0
5
10
15
X1

Figure 4.1D

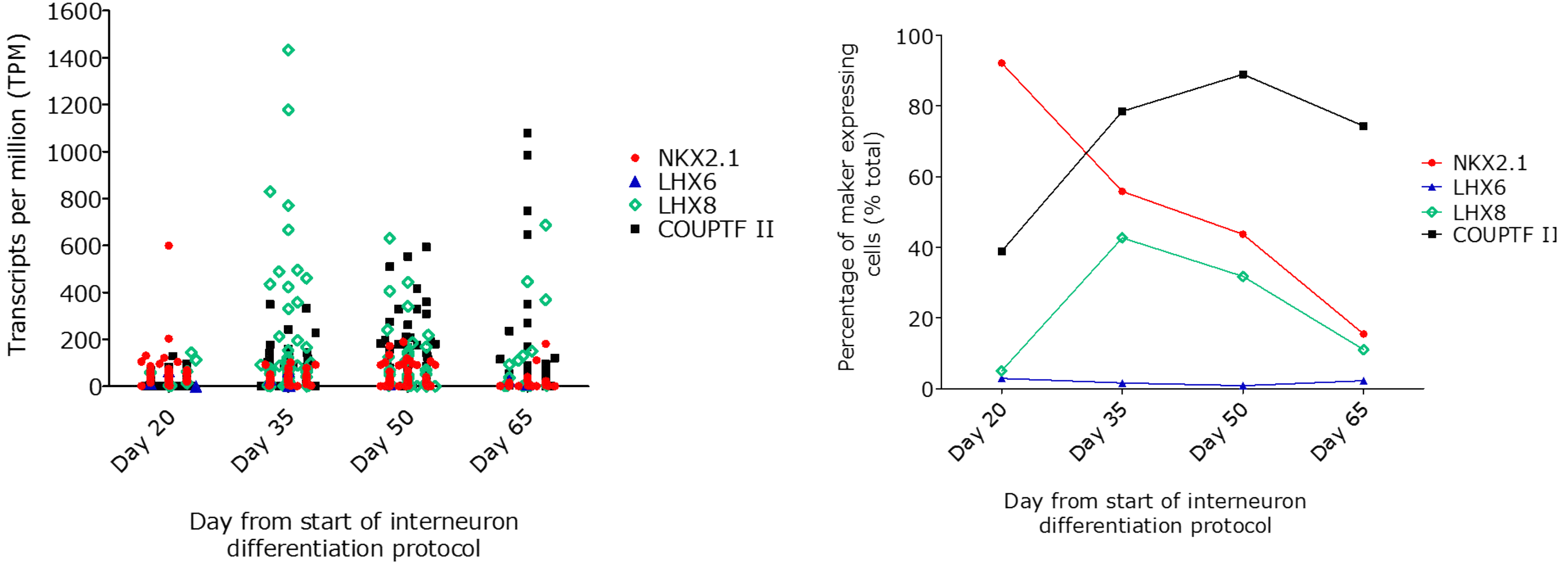

Transcripts per million (TPM)
1600
1400
1200
1000
800
600
400
200
0
Day 20
Day 35
Day 50
Day 65
Day from start of interneuron differentiation protocol
NKX2.1
LHX6
LHX8
COUPTF II
Percentage of maker expressing cells (% total)
100
80
60
40
20
0
Day 20
Day 35
Day 50
Day 65
Day from start of interneuron differentiation protocol
NKX2.1
LHX6
LHX8
COUPTF I]

Figure 4.1E

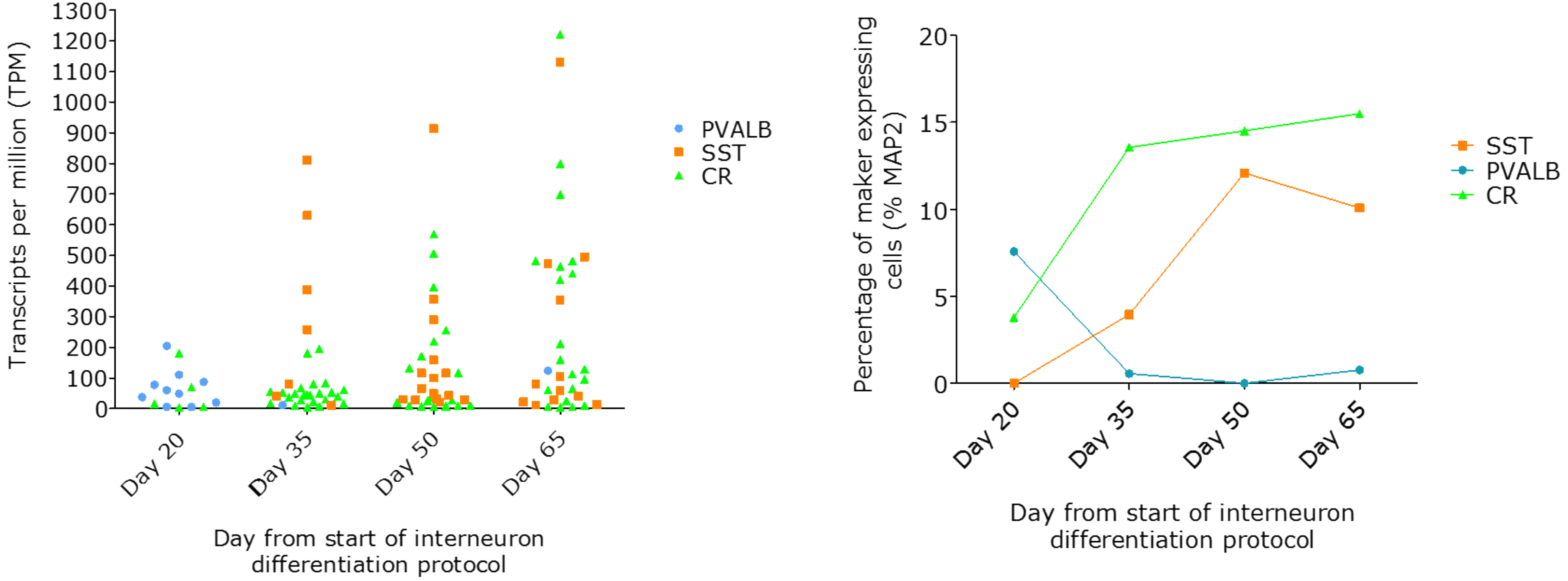

Transcripts per million (TPM)
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
Day 20
Day 35
Day 50
Day 65
Day from start of interneuron differentiation protocol
● PVALB
■ SST
▲ CR
Percentage of maker expressing cells (% MAP2)
20
15
10
5
0
Day 20
Day 35
Day 50
Day 65
Day from start of interneuron differentiation protocol
■ SST
● PVALB
▲ CR

Figure 4.1 Single-cell RNA sequencing analysis of hESC-derived interneurons.

A) Molecular characterisation of hESC-derived interneurons – at single-cell RNA level, day 20. The graph plotted shows the proportion of cells expressing each named marker. The marker positive cells have been counted if the cells are also co-expressing NKX2.1 to restrict analysis to MGE progenitors. Biological replicates (independent differentiations) n=2. Total number of cells per time point can be found in Appendix folder 2 table 4. Total NKX2.1 expressing cells at day 20 = 128 cells.

B) Molecular characterisation of hESC-derived interneurons – at single-cell RNA level, day 65. The graph plotted shows the proportion of cells expressing each named marker. The marker positive cells have been counted if the cells also co-expressing MAP2 to restrict analysis to post-mitotic neurons. Biological replicates (independent differentiations) n=2. Total number of cells per time point can be found in Appendix folder 2 table 4. Total MAP2 expressing cells at day 65 = 129 cells.

C) tSNE plot of the relative similarity in single-cell profiles of hESC-derived interneurons at days 20, 35, 50 and 65. Each symbol on tSNE plot represents a single-cell and the unique gene expression profile of the cell (based on single-cell RNA sequencing data available, trimmed to top 400 genes). Clusters of symbols on tSNE plot represent cells with higher similarity in gene expression profile, with each cluster uniquely labelled with cluster number (e.g. C1) and an example of a highly expressed gene unique to this cluster (where an example exists). For full list of differentially expressed genes between cluster see Appendix folder 2 table 5. Biological replicates (independent differentiations) n=2. Total number of cells per time point can be found in Appendix folder 2 table 4.

D) Trajectory of hESC-derived interneurons expression of MGE/CGE markers– at RNA level, days 20, 35, 50 and 65 in vitro. Two graphs have been presented. One graph plotted shows the sample TPM value for each marker, with each symbol plotted representing the expression value of that marker for a single cell in the sample. The line graph plotted shows the proportion of cells expressing each named marker over time. Biological replicates (independent differentiations) n=2. Total number of cells per time point can be found in Appendix folder 2 table 4.

E) Trajectory of hESC-derived interneurons expression of subtype markers– at RNA level, day 20, 35, 50 and 65 in vitro. Two graphs have been presented. One graph plotted shows the sample TPM value for each marker, with each symbol plotted representing the expression value of that marker for a single cell in the sample. The line graph plotted shows the proportion of cells expressing each named marker over time. The marker positive cells have been counted if the cells also co-expressing MAP2, to restrict analysis to post-mitotic neurons. Biological replicates (independent differentiations) n=2. Total number of cells per time point can be found in Appendix folder 2 table 4. Total MAP2 expressing cells at day 20=132 cells, day 35=177, day 50=124 cells and day 65 = 129 cells.

TPM, transcripts per million. NK2 Homeobox 1, NKX2.1. COUP transcription factor 2, COUPTF II. LHX6, LIM Homeobox 6. FOXG1, Forkhead Box G1. SST, somatostatin. PVALB, parvalbumin. CR, calretinin. LHX8, LIM Homeobox 8. DLX5, Distal-less homeobox 5. DLX1, distal-less homeobox 1. MAP2, microtubule-associated protein 2. Serotonin receptor 3A, 5HT-R 3A. Vasoactive intestinal peptide, VIP. Reelin, RELN. Neuropeptide Y, NPY. Neuronal nitric oxide synthase, nNOS. hESC, human embryonic stem cell. Delta like non-canonical Notch ligand 1, DLK1. Cellular retinoic acid binding protein 1, CRABP1. High mobility group box 2, HMGB2. S100 calcium binding protein A11, S100A11. Secreted protein acidic and cysteine rich, SPARC. Tetraspanin 7, TSPAN7. Stathmin 2, STMN2. Cyclin D1, CCND1. Phospholipid transfer protein, PLTP. Interferon induced transmembrane protein 2, IFITM2

4.3.4 Differentiation trajectory of hPSC-derived interneurons was underpinned by the differential expression of specific transcription factors

Finally, the heterogeneity between samples was explored by performing ANOVA pair-wise analysis (Figure 4.2B-4.2C). On observing all the single-cell profiles presented together on a single tSNE plot, day 20 population would be predicted to have the most unique expression profile relative to other time points (Figure 4.2A). This was indicated by nearly the entire day 20 sample, those corresponding to the early MGE-like progenitors, all clustering together into a single subpopulation (top right cluster, Figure 4.2A). Whereas, the day 35 to day 65 cultures appeared to have many overlapping subpopulations, which could be further divided into three distinct clusters (Figure 4.2A). Furthermore, the ANOVA pair-wise analysis detected the greatest number of significantly differentially expressed genes in the day 20 cross-comparisons (highest being 29 genes for day 20 vs 65 comparisons, Appendix folder 2 table 6 has full list). These significantly differentially expressed genes have been shown in Figure 4.2B and 4.2C, with genes that were more highly expressed at the later time point (e.g. day 35 in a day 20 vs day 35 comparison) assigned a positive fold change value (Figure 4.2B). Whereas, any genes higher expressed at the earlier time point (e.g. day 20 in a day 20 vs day 35 comparison) were assigned a negative fold change value (Figure 4.2B).

To fully evaluate the differentiation trajectory of interneuron-like cells, relative fold changes in differentially expressed genes were determined for all the day 20 cross-comparisons (Figure 4.2B). As seen in Figure 4.2B, certain genes were only differentially expressed relative to day 20 at a single time point, and therefore showed only a transient change in expression over interneuron differentiation. This included *DLX1/5* and fatty acid binding protein 3 (*FABP3*), which were both higher expressed at day 35, but showed no significant change in expression at day 50/65 (Figure 4.2B). This was further reiterated in the day 35 cross-comparisons, which showed that the expression of DLX genes (*DLX1, DLX2, DLX5, DLX6* antisense) peaked at day 35 and then had been downregulated by later time points (Figure 4.2C). In addition, early MGE markers maternally expressed 3 (*MEG3*) and hes family bHLH transcription factor 4/5 (*HES4/5*) were found to be higher expressed at day 20 (Figure 4.2B). Whereas, POU class 3 homeobox 4 (*POU3F4*) was higher expressed at day 50 compared to day 20 and day 35 (Figure 4.2B-4.2C). Also of note was the gene with the highest relative fold-change observed, 220-fold upregulation of nuclear factor I/B (NFI B) at day 65 (day 20 vs 65 comparison). NFI B was also significantly higher expressed at day 65 in both day 20 and day 35 pairwise comparisons (Figure 4.2B-4.2C). Finally, chemokine C-C motif ligand 2 (CCL2) was found to be higher expressed at day 65 relative to day 20 (Figure 4.2B).

The interneuron differentiation of hPSC was also marked by genes consistently up or downregulated between the progenitor (day 20) and post-mitotic stages (day 35 to 65). These genes were detected as significantly differentially expressed genes in all day 20 pairwise cross-comparisons (Figure 4.2B). This included the following genes consistently higher expressed at later time points (upregulated): annexin A1, RNA binding motif single stranded interacting protein 1 (RBMS1) and Na^+ voltage-gated ion channel subunit SCN2A (Figure 4.2B). Overall, the temporal changes in expression pattern reflected the transition of progenitors, with high expression of early transcription factors (*MEG3, HES4/5*) to migratory interneuron-like cells expressing guidance cues (chemokines), ion channels (*SCN2A*) and late MGE markers (*DLX1/5*).

Figure 4.2A

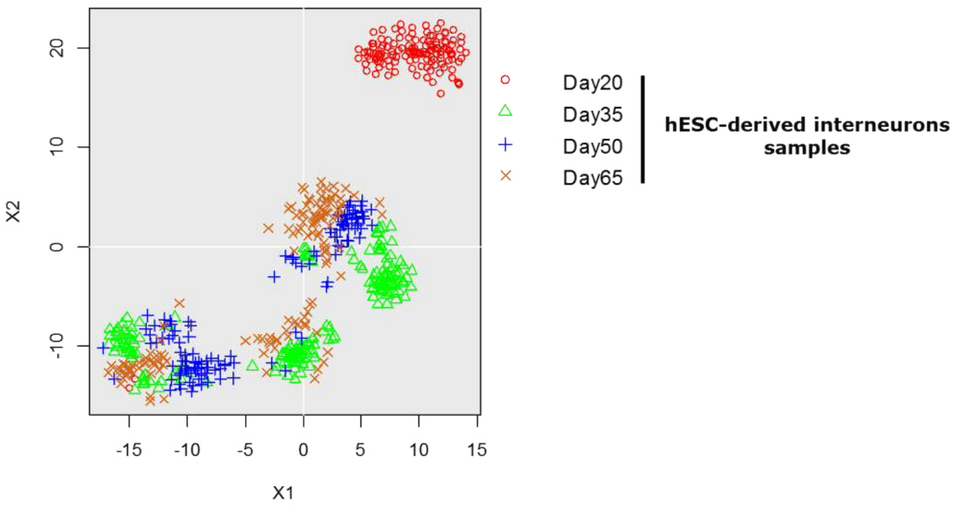


Figure 4.2B

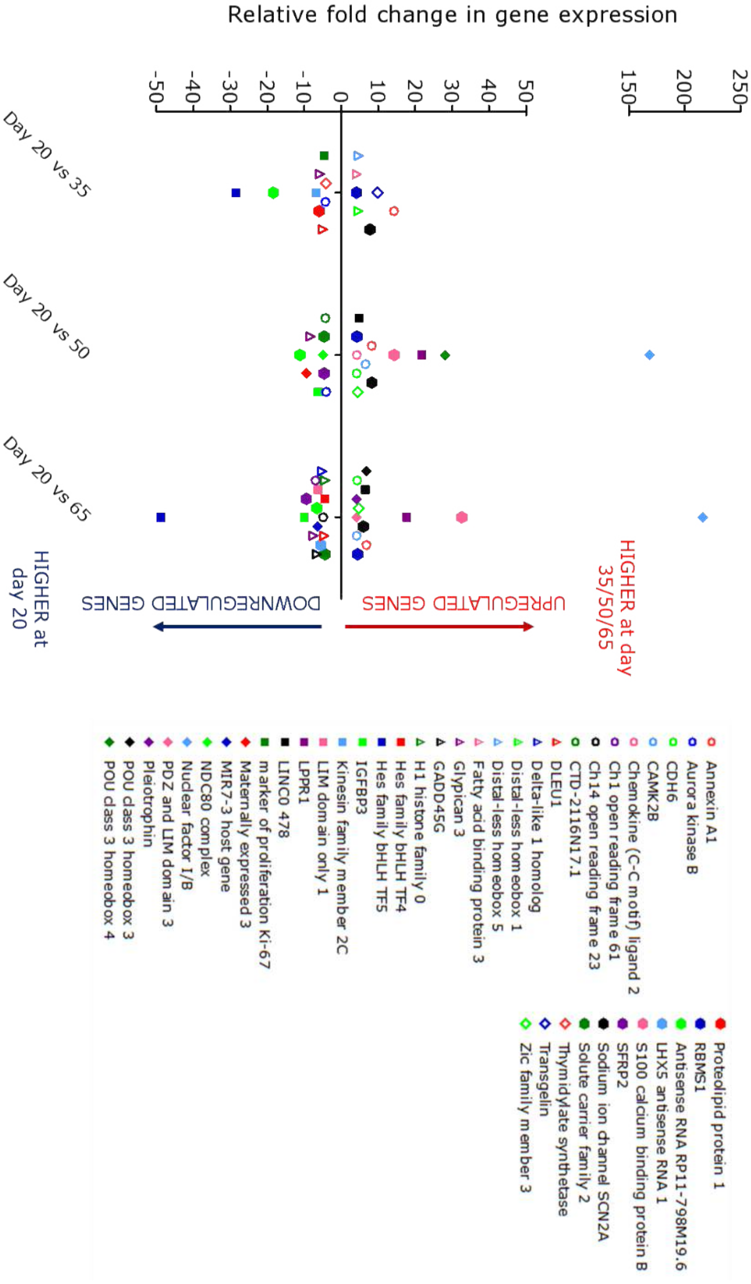


Figure 4.2C

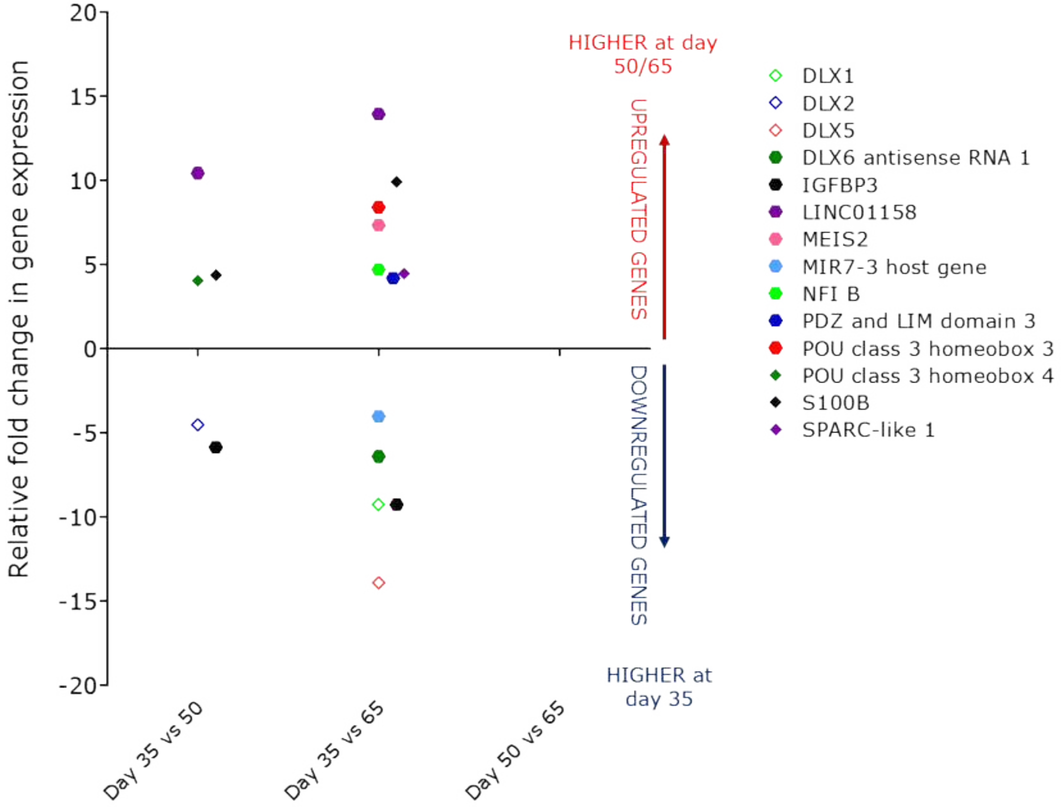

Relative fold change in gene expression
20
15
10
5
0
-5
-10
-15
-20
Day 35 vs 50
Day 35 vs 65
Day 50 vs 65
HIGHER at day 50/65
UPREGULATED GENES
DOWNREGULATED GENES
HIGHER at day 35
◇ DLX1
◇ DLX2
◇ DLX5
● DLX6 antisense RNA 1
● IGFBP3
● LINC01158
● MEIS2
● MIR7-3 host gene
● NFI B
● PDZ and LIM domain 3
● POU class 3 homeobox 3
◆ POU class 3 homeobox 4
◆ S100B
◆ SPARC-like 1

Figure 4.2 Differentially expressed genes marking the developmental trajectory of hESC-derived interneurons.

A) tSNE plot of the relative similarity in single cell profiles of hESC-derived interneurons at days 20, 35, 50 and 65. Each symbol on tSNE plot represents a single cell and the unique gene expression profile of the cell (based on single-cell RNA sequencing data available, trimmed to top 400 genes). Clusters of symbols on tSNE plot represent cells with higher similarity in gene expression profile. Subpopulations within each time point are indicated by spatially isolated clusters. Cells from the same time point along differentiation paradigm share the same symbol (e.g. all day 20 cells are represented by red open circle). Biological replicates (independent differentiations) n=2.

B) Relative fold change in expression of genes significantly differentially expressed between day 20 relative to day 35/50/65. DEGs were determined through ANOVA pairwise analysis, and were included if $p<0.05$ and fold change (either up or downregulated) >4. In this plot, any genes upregulated were relatively higher expressed at the later time point (e.g. day 50 in day 20 vs 50 comparison) and were shown as a positive fold change. Any downregulated genes were higher expressed at the earlier time point (day 20 in all cases), and were shown as a negative fold change. See Appendix folder 2 tables for total cell per sample and DEG list. Biological replicates (independent differentiations) n=2.

C) Relative fold change in expression of genes significantly differentially expressed between day 35 relative to day 50/65, and day 50 relative to day 65. DEGs were determined through ANOVA pairwise analysis, and were included if $p<0.05$ and fold change (either up or downregulated) >4. In this plot, any genes upregulated were relatively higher expressed at the later time point (e.g. day 50 in day 35 vs 50 comparison) and were shown as a positive fold change. Any downregulated genes were higher expressed at the earlier time point (e.g. day 35 in day 35 vs 50 comparison), and were shown as a negative fold change. See Appendix folder 2 tables for total cell per sample and DEG list. Biological replicates (independent differentiations) n=2.

Total number of cells per time point can be found in Appendix folder 2 table 4. CDH6, Cadherin 6 .CAMK2B, Calcium/Calmodulin Dependent Protein Kinase II Beta. Ch1 open reading frame 61, Chromosome open reading frame 61. Ch14 open reading frame 23, Chromosome open reading frame 23. DLEU1, Long Intergenic Non-Protein Coding RNA 21. GADD45G, Growth Arrest And DNA Damage Inducible Gamma. IGFBP3, Insulin Like Growth Factor Binding Protein 3. LPPR1, Phospholipid Phosphatase Related 1. LINC 478, Long intergenic non-protein coding RNA 478. NDC80 complex, NDC80 Kinetochore Complex Component. RBMS1, RNA Binding Motif Single Stranded Interacting Protein 1. SFRP2, Secreted frizzled related protein 2. NK2 Homeobox 1, NKX2.1. COUP transcription factor 2, COUPTF II. LHX6, LIM Homeobox 6. FOXG1, Forkhead Box G1. SST, somatostatin. PVALB, parvalbumin. CR, calretinin. LHX8, LIM Homeobox 8. DLX5, Distal-less homeobox 5. DLX1, distal-less homeobox 1. MAP2, microtubule-associated protein 2. Serotonin receptor 3A, 5HT-R 3A. Vasoactive intestinal peptide, VIP. Reelin, RELN. Neuropeptide Y, NPY. Neuronal nitric oxide synthase, nNOS. DLX1, distal-less homeobox 1. DLX2, distal-less homeobox 2. LINC01158, long intergenic non-protein coding RNA 1158. MEIS2, meis homeobox 2. NFIB, nuclear factor I/B. S100B, S100 calcium binding protein B. hESC, human embryonic stem cell. Differentially expressed genes, DEGs.

4.3.5 Single-cell RNA sequencing data was validated by quality control checks

Before using the single-cell cDNA samples collected from Fludigm C1 IFC for library preparation, several quality control checks were performed to ensure that the single-cell cDNA was of high enough quality for sequencing (Figure 4.3A-4.3D, Appendix folder 2 figures). The reaction chemistry used generated a high enough concentration of single-cell cDNA to allow both single-cell QPCR and RNA sequencing to be performed on the same samples (Figure 4.3A-4.3D).

In agreement to single-cell RNA sequencing data presented above, single-cell QPCR analysis at day 20 detected the high abundance of cells expressing known early MGE and CGE markers, such as *FOXG1, NKX2.1* and *COUPTF II,* and the absence of markers of alternative fates (*CTIP2, FOXP1, PAX6*) (Figure 4.3A and Appendix folder 2 figures). However, single-cell QPCR analysis would also suggest a higher abundance in *NKX2.1* and *COUPTF II* co-expressing cells (90% *NKX2.1/COUPTF II*, Figure 4.3A). However, this finding was not replicated in single-cell RNA sequencing analysis performed on the same cDNA samples, nor in immunohistochemistry of day 20 progenitors shown in Chapter 3, which all found between 25-40% COUPTF II positive population at day 20 (Figure 4.1A and 4.3A).

Another discrepancy reported from single-cell QPCR data was the relatively similar proportion of *MAP2* expressing cells that co-expressed *SST* and *CR* at day 65 (approx. 20% for each subtype marker, Figure 4.3B). This finding was slightly mismatched to the single-cell RNA sequencing, which reported higher *CR* subtype abundance relative to *SST* (Figure 4.1B and 4.3B). Moreover, the longitudinal expression trajectory of *CR* and *SST* was inconsistent with single-cell RNA sequencing findings (Figure 4.1E and 4.3D). The reported abundance of *SST* expressing cells was much higher in the single-cell QPCR analysis (peaking at 45% vs 13% at day 50, Figure 4.1D and 4.3D). This may be due to primer efficiency issues causing overestimation of *SST* levels in single-cell QPCR. Alternatively, single-cell RNA sequencing may be underestimating the abundance of *SST* expressing neurons, due to the lower sensitivity of the technique, which may have prevented cells with low *SST* transcript values being detected (Figure 4.3C). These neurons with low expression of *SST* (delta Ct >14) were of high abundance at day 50, the time point with the greatest discrepancy between the two methodologies (Figure 4.3C). However, what was encouraging was the detection of the rare *PVALB* interneurons, and *SST/CR* and *PVALB/CR* co-expressing cells within the culture using both single-cell QPCR and RNA sequencing methodologies of transcriptome analysis (Figure 4.1B and 4.3A). Furthermore, even though only a limited number of selected genes were used for single-cell QPCR analysis (30 genes tested), heterogeneity in expression of these genes between cells was still detected, thereby revealing some of the molecular heterogeneity that was fully described in single-cell RNA sequencing analysis (Appendix folder 2 figures).

Figure 4.3A

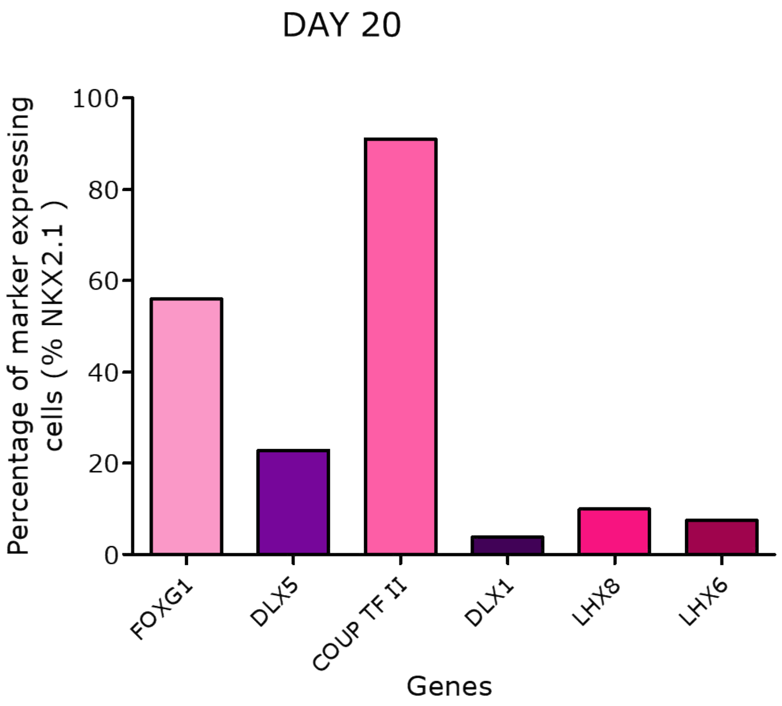


Figure 4.3B

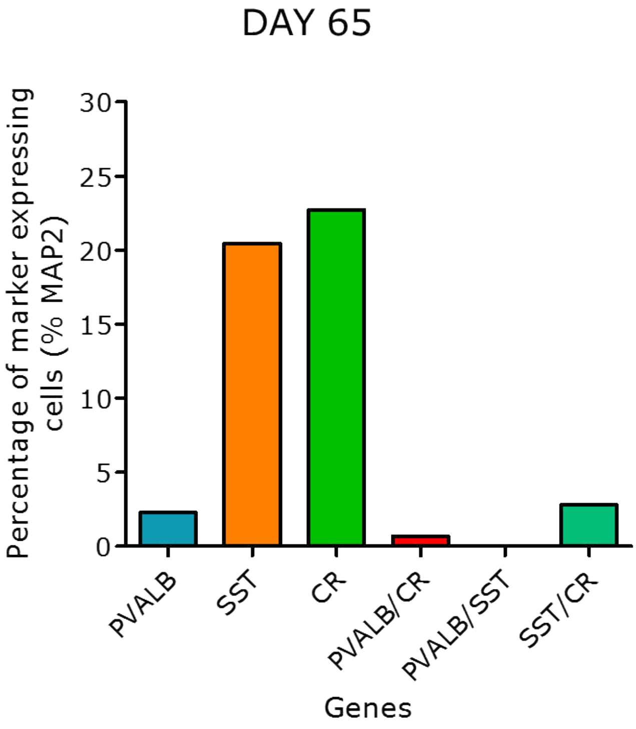


Figure 4.3C

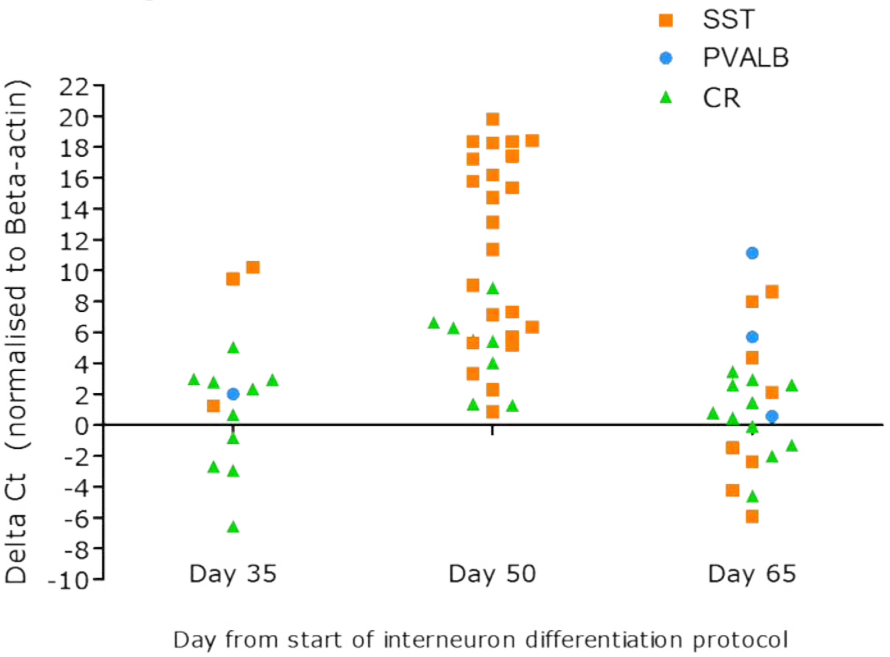


Figure 4.3D

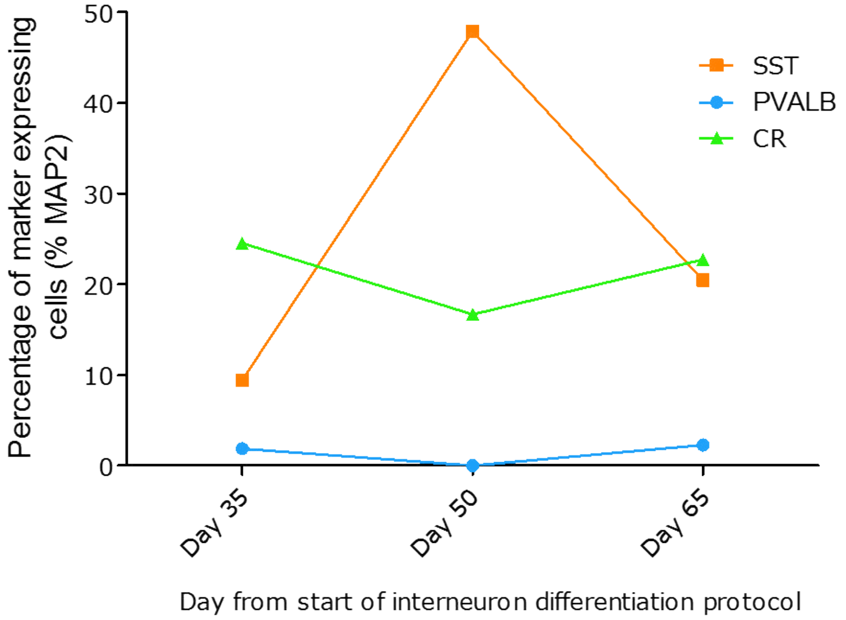


Figure 4.3 Single-cell QPCR analysis: quality control check data.

A) Molecular characterisation of hESC-derived interneurons – at single-cell RNA level, day 20. The graph plotted shows the proportion of cells expressing each named marker. The marker positive cells have been counted if the cells are also co-expressing NKX2.1 to restrict analysis to MGE progenitors. Biological replicates (independent differentiations)= 2. Technical replicates =3. Total NKX2.1 expressing cells at day 20= 93 cells.

B) Molecular characterisation of hESC-derived interneurons – at single-cell RNA level, day 65. The graph plotted shows the proportion of cells expressing each named marker. The marker positive cells have been counted if the cells are also co-expressing MAP2 to restrict analysis to post-mitotic neurons. Biological replicates (independent differentiations)= 2. Technical replicates =3. Total MAP2 expressing cells at day 65 = 46 cells.

C) Trajectory of hESC-derived interneurons subtypes– at RNA level, days 35, 50 and 65 in vitro. The data has been presented in the form of relative gene expression, with delta Ct calculated in reference to beta-actin. Each cell positive for the subtype marker, at each time point, has been individually displayed by a symbol on the graph. The marker positive cells have been counted if the cells also co-expressing MAP2, to restrict analysis to post-mitotic neurons. Biological replicates (independent differentiations)= 2. Technical replicates =3. Total MAP2 expressing cells per time point = 54 cells at day 35, 51 cells at day 50 and 46 cells at day 65.

D) Trajectory of hESC-derived interneurons subtypes– at RNA level, days 35, 50 and 65 in vitro. The graph plotted shows the proportion of cells expressing each named marker. The marker positive cells have been counted if the cells also co-expressing MAP2, to restrict analysis to post-mitotic neurons. Biological replicates (independent differentiations)= 2. Technical replicates =3. Total MAP2 expressing cells per time point = 54 cells at day 35, 51 cells at day 50 and 46 cells at day 65.

NK2 Homeobox 1, NKX2.1. COUP transcription factor 2, COUPTF II. LHX6, LIM Homeobox 6. FOXG1, Forkhead Box G1. SST, somatostatin. PVALB, parvalbumin. CR, calretinin. LHX8, LIM Homeobox 8. DLX5, Distal-less homeobox 5. DLX1, distal-less homeobox 1. MAP2, microtubule-associated protein 2. hESC, human embryonic stem cell.

4.4 Discussion

The longitudinal single-cell RNA sequencing analysis of hPSC-derived interneurons presented here has expanded upon the molecular characterisation presented in Chapter 3. With this data, a novel description of interneuron differentiation of hPSC has been provided.

4.4.1 Examining the subtype diversity within hPSC-derived interneurons

An unbiased and longitudinal evaluation of the hPSC-derived interneuron subtype composition revealed the full extent of subtype diversity achieved using this paradigm. Although the culture was enriched in early MGE-like progenitors, on terminal differentiation a population abundant in *RELN, SST* and *CR* expressing interneuron-like cells was generated. These subtypes are typically derived from the dorsal MGE and CGE (Xu *et al.* 2004; Butt *et al.* 2005; Flames *et al.* 2007; Wonders *et al.* 2008; Miyoshi *et al.* 2010). On examining the differentiation trajectory of the early progenitors, the data revealed a low abundance/absence of late MGE marker (*LHX6, LHX8, DLX5*) expressing cells. Instead the population progressively became more caudal in molecular profile, with the rising expression of CGE enriched transcription factors *COUPTF II* and *NFI B*, with the latter being the most highly upregulated gene from day 50 onwards (Hu *et al.* 2017a; Mayer *et al.* 2018). This expression pattern may explain why the day 65 hESC-derived interneuron population had an absence or low abundance of certain MGE-derived subtypes (*NPY* and *PVALB* respectively) (Vogt *et al.* 2014; Hu *et al.* 201a7; Mayer *et al.* 2018). A similar finding has been reported independently for mESC interneuron differentiation protocol, with SHH signalling alone inefficiently generating cells expressing *Lhx6* and MGE-derived subtype markers *Pvab, Npy,* and *Sst*, which were typically co-expressed with *Lhx6* (Danjo *et al.* 2011; Chen *et al.* 2017). This expression pattern has been linked to the absence of other critical spatial patterning cues, such as FGF8 and BMP4, which when added have selectively enhanced commitment to these lineages (Mukhopadhyay *et al.* 2009; Danjo *et al.* 2011; Chen *et al.* 2013).

In addition, the longitudinal single-cell RNA sequencing analysis provided an understanding of when certain subtype dictated subpopulations emerge, and the dynamics of these subpopulations. In general, subtype marker expressing cells grow in abundance over the course of the hPSC interneuron differentiation paradigm (Close *et al.* 2017; Xiang *et al.* 2017). This trend was true for most of the interneuron-like subtypes detected here, including *CR*, *SST* and *RELN*. However, the longitudinal analysis presented some novel and non-intuitive findings in relation to the emergence of PVALB interneuron-like cells. At day 20 the hESC-derived interneuron culture had a small proportion of nascent post-mitotic interneuron-like cells, which were identified as *CR, PVALB* and *CR/PVALB* co-expressing. However, the abundance of *PVALB* expressing neurons peaked at day 20 and declined over the remainder of the differentiation paradigm. This trend in *PVALB* expression was a novel finding and was in contrast to the expected trend. Nonetheless, the co-expression pattern of PVALB interneuron-like cells, and the independently reported existence of *PVALB/CR* co-expressing interneurons endogenously (hippocampal Chandelier cells), added weight to these cells being genuine PVALB interneuron-like cells (Ishino *et al.* 2017). This led to the conclusion that the neuronal maturation media was insufficient to maintain the PVALB interneuron-like cells. A similar report has also been made independently for the loss of SST interneuron-like cells generated in a similar hPSC differentiation paradigm, although at a much later time point (day 50 onwards) (Close *et al.* 2017). Together these findings would indicate a missing component in the differentiation paradigm that would appear to be essential for maintaining certain interneuron subtypes. A strong candidate would be the absence of native synaptic partners. Endogenously the signals received as part of a neuronal circuit is necessary to promote the survival and maturation of neurons. Moreover, there is evidence of activity-dependent signalling promoting the terminal differentiation of interneuron, as well as fine-tuning the final subtype commitment (Ishino *et al.* 2017; Quattrocolo *et al.* 2017; Wamsley & Fishell 2017). Therefore, one strategy to explore that may help maintain (and potentially enhance) PVALB interneuron-like cells in vitro would be to co-culture hESC-derived interneurons with their native synaptic partners (such as primary tissue or hPSC-derived glutamatergic neurons) (Mukhopadhyay *et al.* 2009; Colasante *et al.* 2015; Meganathan *et al.* 2017).

4.4.2 The diversification of hPSC-derived interneurons

By tracking the differential gene expression over time an in-depth characterisation of genes marking the differentiation trajectory of hESC-derived interneurons was obtained. Many of these findings were in agreement to previous reports from other PSC models and mouse MGE/CGE, including the expression of specific transcription factors (*HES4/5, MEG3, DLK1, NFI B*) by the interneuron-like progenitors (Chen *et al.* 2017; Close *et al.* 2017; Mayer *et al.* 2018). Moreover, the upregulation in expression of migratory cues and ion channel subunits as the population progressed towards terminal differentiation would indicate a close recapitulation of later stages of native interneuron development (Mukhopadhyay *et al.* 2009; Hansen *et al.* 2013). In addition, novel candidate markers of interneuron differentiation trajectory have also been named, including kinases (calmodulin-dependent protein kinase II beta, *CAMK2B*) and cytoskeleton regulatory proteins (PDZ and LIM domain 3). Furthermore, in contrast to previous reports that detected a gradual increase in the heterogeneity of the hPSC-derived interneuron population over time, this study reported a sharp increase in the heterogeneity between day 20 and day 35 (Xiang *et al.* 2017). This sharp transition has been proposed as the emergence of subtype dictated interneuron-like cells in the culture, and the depletion of MGE-like progenitors. The subtypes detected at day 35 were maintain over time, but grew in abundance between day 35 and 65. This would encourage additional sampling between day 20 and day 35 to fully characterise the gene expression changes occurring in this critical period.

In addition to which genes are expressed, single-cell transcriptomics also provides information concerning the between-cell differences in the expression level of each gene. However, what factors caused the number of transcripts generated in each cell to differ, and whether this has any consequences on function/phenotype is uncertain, and is likely gene-dependent. Previous reports in mouse have demonstrated the ability to predict particular physiology parameters of a neuron from the expression level of specific genes (Dehorter *et al.* 2015; Fuzik *et al.* 2015; Bardy *et al.* 2016). For example, the firing pattern of mouse cortical PVALB interneurons (either non-delayed or delayed spiking) was dependent on the protein level of ER81 (Dehorter *et al.* 2015). This highlights the importance of not just which genes are expressed, but also by how much. Therefore, methods that allow direct manipulation over the levels of selected transcripts or proteins, using RNA inference or small molecule-assisted shut-off (SMASh) platforms respectively, may be an advantageous means to generate a more defined interneuron subtype composition (Chung *et al.* 2015; Anton *et al.* 2018).

4.4.3 Candidate lineage determining genes

Finally, in-depth evaluation of genes co-expressed with *PVALB, SST* and *CR* using single-cell RNA sequencing allowed novel candidate genes that may contribute to the unique phenotype and fate commitment of these interneuron subtypes to be named. This has included subtype-specific association of cyclin I with SST interneuron-like cells. Previous rodent studies had reported the importance of cyclins, cell cycle regulatory proteins, in maintaining the appropriate balance of SST and PVALB subtypes during development (Glickstein *et al.* 2007; Gilani *et al.* 2014). Moreover, the role of cyclin I as a cyclin-dependent kinase 5 (CDK5) specific activator, has indicated a function beyond the cell cycle. The activated CDK5 is responsible for regulating cytoskeleton dynamics in neuronal migration and synaptogenesis, as well as promoting neuronal survival (Brinkkoetter *et al.* 2010; Rakić *et al.* 2015; Shah & Lahiri 2017). Moreover, deficits in CDK5 activity in mice has been shown to cause depletion in SST and PVALB interneuron subtypes, and in humans has been associated with neurodevelopmental disorders, including schizophrenia and intellectual disability (Rakić *et al.* 2015; Shah & Lahiri 2017). Overall, this evidence would support the role of cyclin I in SST (and possibly PVALB) lineage commitment.

Another gene of interest was the transcription co-activator uniquely expressed by PVALB interneuron-like cells, SS18 like 2. Little research has been performed on SS18 like 2, with the function inferred from the role of the gene homologue SS18 (de Bruijn *et al.* 2001). Nonetheless, the predicted function of SS18 like 2 has been linked to chromatin remodelling, and therefore SS18 like 2 holds the potential to regulate the expression of a large number of genes, with which genes depending on the accessory proteins present in the cell (de Bruijn *et al.* 2001).

4.4.4 Conclusions

In summary, the series of experiments described in this Chapter have demonstrated the multifaceted analysis possible with single-cell transcriptomics. Future experiments performing enforced expression of named candidate lineage determining genes (cyclin I and SS18-like 2), along with the underrepresented (*LHX6* and *DLX5*) genes, in hPSC-derived interneurons may help to refine the interneuron subtype composition/abundance.

Chapter 5 - Single-cell transcriptome analysis of primary human foetal MGE

5.1 Introduction

PSC-derived models of interneuronal development aim to recapitulate the diversification of interneurons into the spectrum of subtypes present in the adult brain. Such a remarkable feat has been challenging to reproduce in vitro (Cambray *et al.* 2012; Liu *et al.* 2013; Maroof *et al.* 2013; Nicholas *et al.* 2013; Kim *et al.* 2014; Sun *et al.* 2016). This largely stems from the lack of clarity concerning the when, where and how of native interneuron differentiation (Kessaris *et al.* 2014; Hu *et al.* 2017b; Wamsley & Fishell 2017). As described in detail in the General Introduction (Chapter 1), there are several alternative models of interneuron lineage commitment that propose different answers to the following questions (Hu *et al.* 2017b; Wamsley & Fishell 2017). First, when do interneurons diverge into lineage dictated subtypes (at birth or later)? Second, where is the final fate determined, in the region of origin or the region of destination? Third, how are interneurons directed to commit to a specific lineage in favour of another? Depending on the selected model, to answer these questions and identify the Holy Grail in interneuron research (lineage determining transcription factors), molecular characterisation of interneurons in the region of origin and destination at single-cell resolution has been performed (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018; Zhong *et al.* 2018; Loo *et al.* 2019; Mayer *et al.* 2019).

5.1.1 Single-cell transcriptomics

As illustrated in Chapter 4, longitudinal single-cell transcriptomics is the ideal tool to detect the emergence of distinct, lineage committed subpopulations across development (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018). The precise and unbiased nature of single-cell RNA sequencing allows the genes uniquely shared by each subpopulation to be identified, with any transcription factors named as candidate genes for driving lineage commitment (Mayer *et al.* 2018; Mi *et al.* 2018). Although several longitudinal single-cell transcriptomics studies have been conducted on mouse and human foetal derived interneurons, the ability to name candidate lineage determining transcription factors has been inconsistent between studies, and only obtained from studies in the region of origin (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018; Zhong *et al.* 2018; Loo *et al.* 2019; Mayer *et al.* 2019). To date, single-cell profiling of the region of origin (MGE/CGE) has only been performed in mouse (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018).

5.1.2 Molecular characterisation of primary mouse MGE and CGE

Native interneuron development in mouse MGE and CGE can be monitored over a shorter timeframe compared to humans (a matter of weeks relative to months). Thus, making longitudinal analysis encompassing the full development of the MGE/CGE a more achievable task in rodent models (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018). As a result, a large body of research concerning interneuron development has been obtained from rodent models (Marín *et al.* 2000; Kessaris *et al.* 2014).

A comparative assessment of the primary mouse CGE and (dorsal and ventral) MGE profiles during embryonic development revealed region-specific markers, some novel markers and some previously reported from immunostaining and bulk-QPCR (Butt *et al.* 2005; Wonders *et al.* 2008; Miyoshi *et al.* 2010; Miyoshi *et al.* 2015; Mayer *et al.* 2018; Mi *et al.* 2018). Mouse CGE-derived population abundantly expressed forkhead box P2 (*Foxp2),* prospero homeobox 1 *(Prox1), Meis2, 5ht-r 3a,* and *Nfi b* (Mayer *et al.* 2018; Mi *et al.* 2018). Whereas, mouse MGE had an enrichment of *Nkx2.1, Lhx6, Hes5, Mef2c*, Patched (*Ptch1*) and *Lhx8* (Mayer *et al.* 2018; Mi *et al.* 2018). These named markers were either absent or at low abundance in the alternative region (Mayer *et al.* 2018; Mi *et al.* 2018). Although there were distinctions in the expression profile of the MGE and CGE-derived populations, some similarities were observed. Of significance was the overlap in gene expression profile between the dorsal subdomain of the MGE and CGE, especially in terms of shared *Couptf II* and *Sst* expression (Mayer *et al.* 2018; Mi *et al.* 2018).

Furthermore, evidence of the biased commitment of progenitors born in CGE and MGE was detected in the single-cell RNA sequencing analysis. Firstly, in the detection of different variety and abundance of subtype marker expressing cells in the MGE and CGE, which included *Sst* and *Npy* (more abundant in MGE), and *Vip* and *Reln* (more abundant in CGE) (Mayer *et al.*

2018). Secondly, transcription factors selectively associated with committed PVALB and SST interneurons were identified and named as candidate lineage determining transcription factors (*Maf* for SST lineage and *Mef2c* for PVALB lineage) (Mayer *et al.* 2018; Mi *et al.* 2018).

5.1.3 Age-dependent changes in primary mouse MGE and CGE

Although there was a clear origin-dependent distinction in single-cell expression profiles, age-dependent changes were well conserved between the CGE and MGE. Over time, the primary mouse MGE and CGE populations transitioned towards a more heterogeneous, post-mitotic neuronal rich population (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018). At each developmental age the MGE/CGE neuronal population included neuronal committed cells, from highly proliferative progenitors to migratory immature interneurons (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018). The emergence of additional and distinct immature interneuron subpopulations over time led to the greater heterogeneity observed in the older MGE/CGE populations (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018).

Moreover, fate mapping studies characterising the grafts obtained from temporally distinct MGE/CGE tissues (E13.5 and E15.5) have demonstrated the differential lineage commitments of MGE/CGE progenitors born at different ages of development (Butt *et al.* 2005; Inan *et al.* 2012). This has extended the view on age-dependent changes in the MGE/CGE beyond maturation to also reflect the temporal dynamics in the differentiation potential of MGE and CGE progenitors (Chen *et al.* 2017). Therefore, longitudinal single-cell characterisation of the MGE and CGE over development could tease apart the differences between early and late-born progenitors, and lead to the identification of lineage determining transcription factors for additional subtypes. For example, this could be particularly relevant to determining the genes underpinning the divergence of progenitors into SST or PVALB committed cells in the MGE, which in mouse have shown to have an age-dependent regulation (early-born, SST bias, late-born PVALB bias) (Inan *et al.* 2012; Bandler *et al.* 2017).

5.1.4 Species differences and importance of primary human tissue studies

The studies described above in rodents have provided invaluable information concerning native MGE and CGE development, and helped fill the unknown gaps of human neurodevelopment, which (due to the legal restriction) cannot be monitored in the same complete manner (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018). With promising findings from the mouse work, longitudinal single-cell RNA sequencing analysis of human MGE and CGE is eagerly awaited. The importance of carrying out the parallel analysis in human foetal tissue has been demonstrated by between-species differences in interneuron abundance and diversity (Wu & Parent 2000; Bakken *et al.* 2016; Lake *et al.* 2016; Boldog *et al.* 2018). Therefore, rodent models may only provide an incomplete picture of interneuron development in human. This has encouraged, for the first time, the characterisation of human MGE using single-cell RNA sequencing in this study. Previous studies characterising human ganglionic eminence have guided a timeframe of 10WPC to 15WPC for the analysis (Zecevic *et al.* 2011; Reinchisi *et al.* 2012; Hansen *et al.* 2013; Onorati *et al.* 2014; Al-Jaberi *et al.* 2015).

5.2 Aims

To characterise native interneuron development by performing longitudinal single-cell RNA sequencing analysis on human MGE (10WPC to 15WPC).

To identify potential temporal and subtype-dependent gene expression signatures in the MGE population.

To evaluate the authenticity of the hPSC model of interneuron differentiation by performing a cross-comparison between the human MGE and hESC-derived interneuron single-cell RNA sequencing data.

5.3 Results

Single-cell RNA sequencing analysis was conducted on human primary MGE tissue aged from 10WPC to 15WPC. The same single-cell capture methodology (Fluidigm C1) and reaction chemistry (SMART-seq) were used as the single-cell work described previously for hESC-derived interneurons (Chapter 4), thereby making tissue source the only changed factor.

5.3.1 The human MGE samples had consistently low contamination by neighbouring CGE and LGE regions

Before performing single-cell transcriptome analysis on human primary foetal tissue, pilot data was generated to determine whether the MGE was being consistently and accurately dissected from neighbouring CGE and LGE regions. This analysis was performed at both the RNA and protein level. Additionally, the collection of primary tissue samples aged between 9WPC and 15WPC (≥3 foetuses per age) for QPCR analysis enabled a bulk-level evaluation of the age-dependent changes in expression trajectory of known MGE, CGE and LGE markers (in Appendix folder 3 figures).

Bulk-QPCR analysis on human MGE, CGE and LGE from the same foetal tissue source was performed across 10WPC to 15WPC period (Figure 5.1A and Appendix folder 3 figures). As examples, only data from the 15WPC MGE samples have been provided (Figure 5.1A). From this data set the appropriate expression pattern of both transcription factors and post-mitotic markers across regions was demonstrated (Figure 5.1A). For example, MGE-specific transcription factors, *NKX2.1, LHX8* and *LHX6*, were consistently higher expressed (relatively lower average delta Ct value) in MGE than the other regions (Figure 5.1A). This trend was also true for MGE-derived subtype marker *PVALB,* although *SST* expression level was similar between CGE and MGE (Figure 5.1A). In addition, *COUPTF II* expression level in 15WPC MGE matched the level of expression in the CGE (average delta Ct of 4.5, Figure 5.1A). Although the MGE contains *COUPTF II* expressing cells (typically restricted to the dorsal subdomain), the high level of *COUPT II* would suggest a possible CGE contamination of the MGE (Figure 5.1A). The low levels of contamination reported here, and seen in other cases (LHX6 and SOX6 levels, Figure 5.1A) may be assigned to the migratory interneuron precusors, which were migrating through neighbouring region at time of collection.

Independent analysis was conducted at protein level by performing immunohistochemistry, using a single foetal tissue per age at 10WPC and 15WPC (Figure 5.1B-5.1C and Appendix folder 3 figures). As examples, only data from the 15WPC MGE samples have been provided in Figure 5.1B-5.1C. In a similar manner to the bulk-QPCR data, the seeded 15WPC MGE, CGE and LGE samples were checked for the abundance of region-specific markers in the population, along with common pan-ganglionic markers (Figure 5.1B-5.1C). The predicted expression patterns were seen in the immunostainings, with 15WPC MGE population relatively sparse in LGE-markers (<5% FOXP1 and DARPP32) relative to LGE (20% FOXP1, 5% DARPP32, Figure 5.1C). Moreover, the MGE exhibited a higher abundance of NKX2.1, SST and ChAT positive cells than CGE or LGE (Figure 5.1C). The similar abundance of CR positive cells (25%) in MGE and CGE samples may be due to migratory CGE-derived CR interneuron precursors being located in the MGE at age of collection, although CR interneurons are derived from both regions (Figure 5.1C).

By using two complementary approaches, the expression patterns in the MGE, CGE and LGE were consistent with the known expression profile of these regions, which indicated accurate dissection of each region. However, due to the migratory nature of interneuron precursors, low level of contamination from CGE should be expected and taken into account when analysing the single-cell RNA sequencing data.

Figure 5.1A

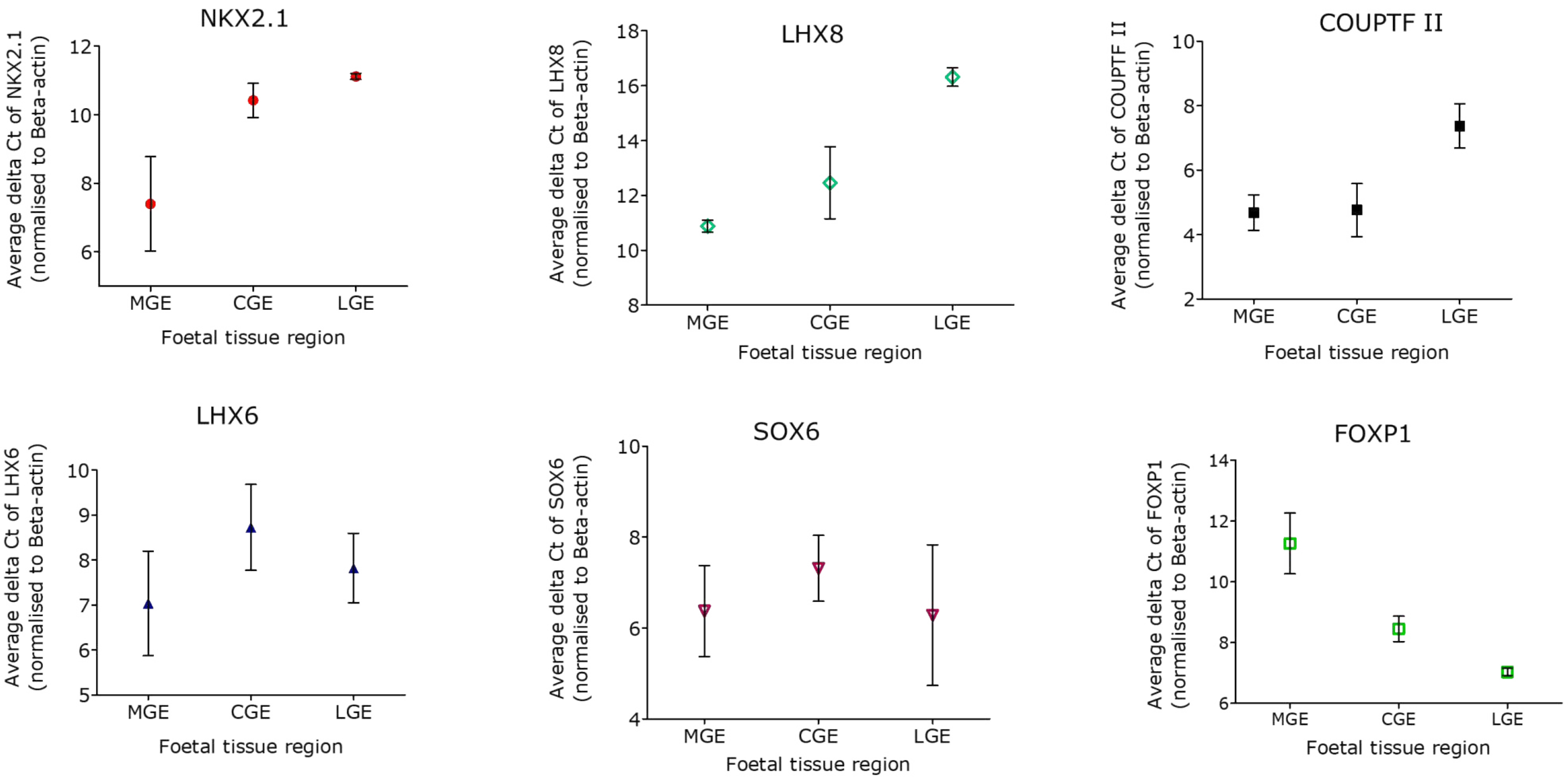

NKX2.1
Average delta Ct of NKX2.1 (normalised to Beta-actin)
6
8
10
12
MGE
CGE
LGE
Foetal tissue region
LHX8
Average delta Ct of LHX8 (normalised to Beta-actin)
8
10
12
14
16
18
MGE
CGE
LGE
Foetal tissue region
COUPTF II
Average delta Ct of COUPTF II (normalised to Beta-actin)
2
4
6
8
10
MGE
CGE
LGE
Foetal tissue region
LHX6
Average delta Ct of LHX6 (normalised to Beta-actin)
5
6
7
8
9
10
MGE
CGE
LGE
Foetal tissue region
SOX6
Average delta Ct of SOX6 (normalised to Beta-actin)
4
6
8
10
MGE
CGE
LGE
Foetal tissue region
FOXP1
Average delta Ct of FOXP1 (normalised to Beta-actin)
6
8
10
12
14
MGE
CGE
LGE
Foetal tissue region

Figure 5.1A

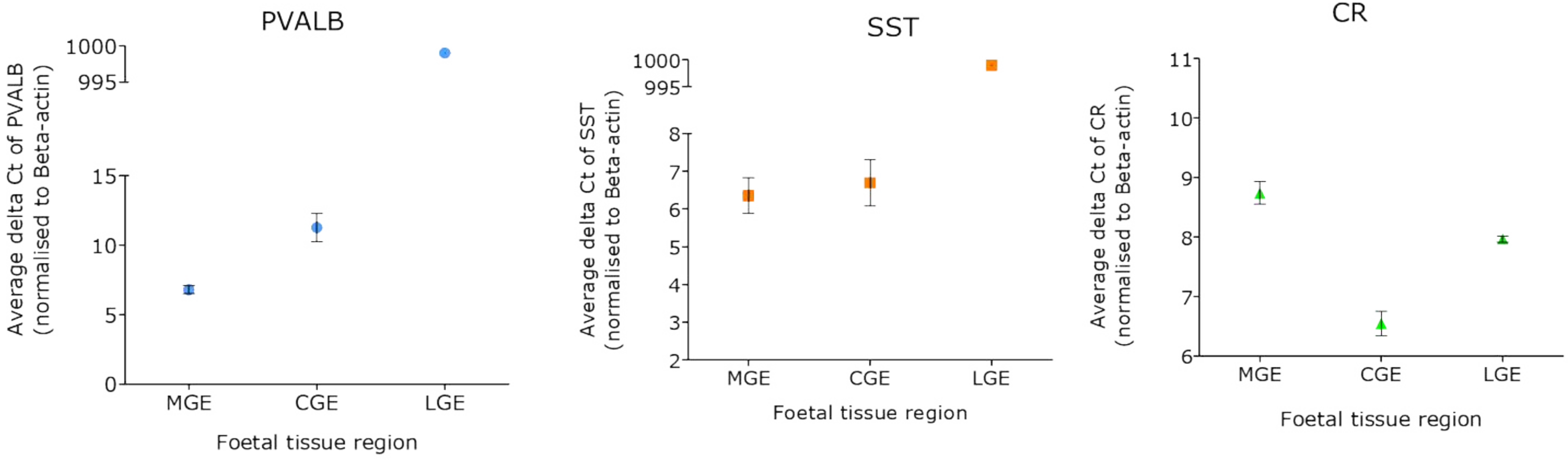

PVALB
Average delta Ct of PVALB (normalised to Beta-actin)
1000
995
15
10
5
0
MGE
CGE
LGE
Foetal tissue region
SST
Average delta Ct of SST (normalised to Beta-actin)
1000
995
8
7
6
5
4
3
2
MGE
CGE
LGE
Foetal tissue region
CR
Average delta Ct of CR (normalised to Beta-actin)
11
10
9
8
7
6
MGE
CGE
LGE
Foetal tissue region

Figure 5.1B

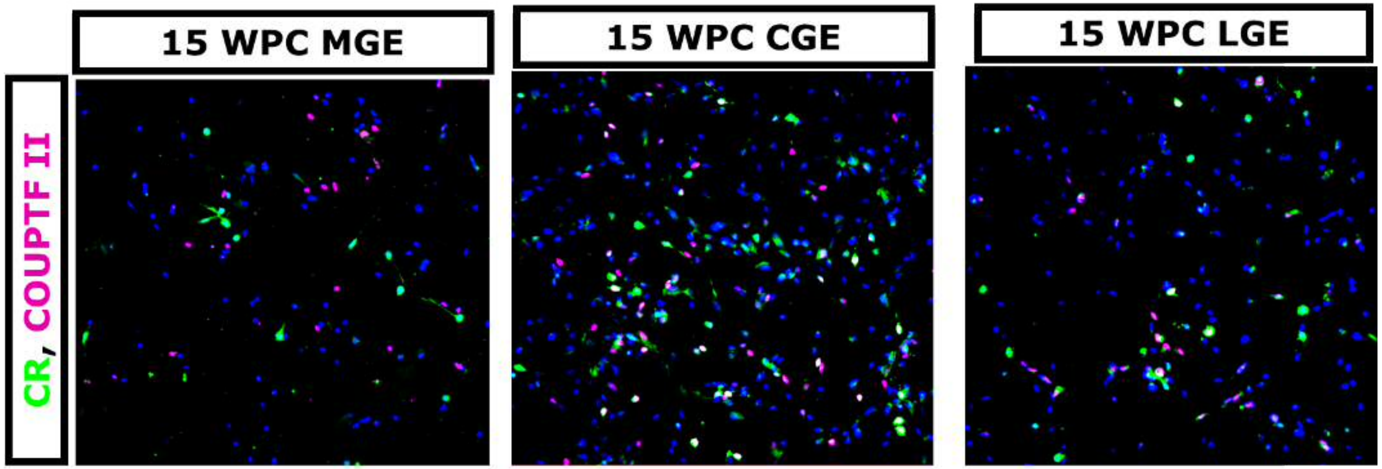
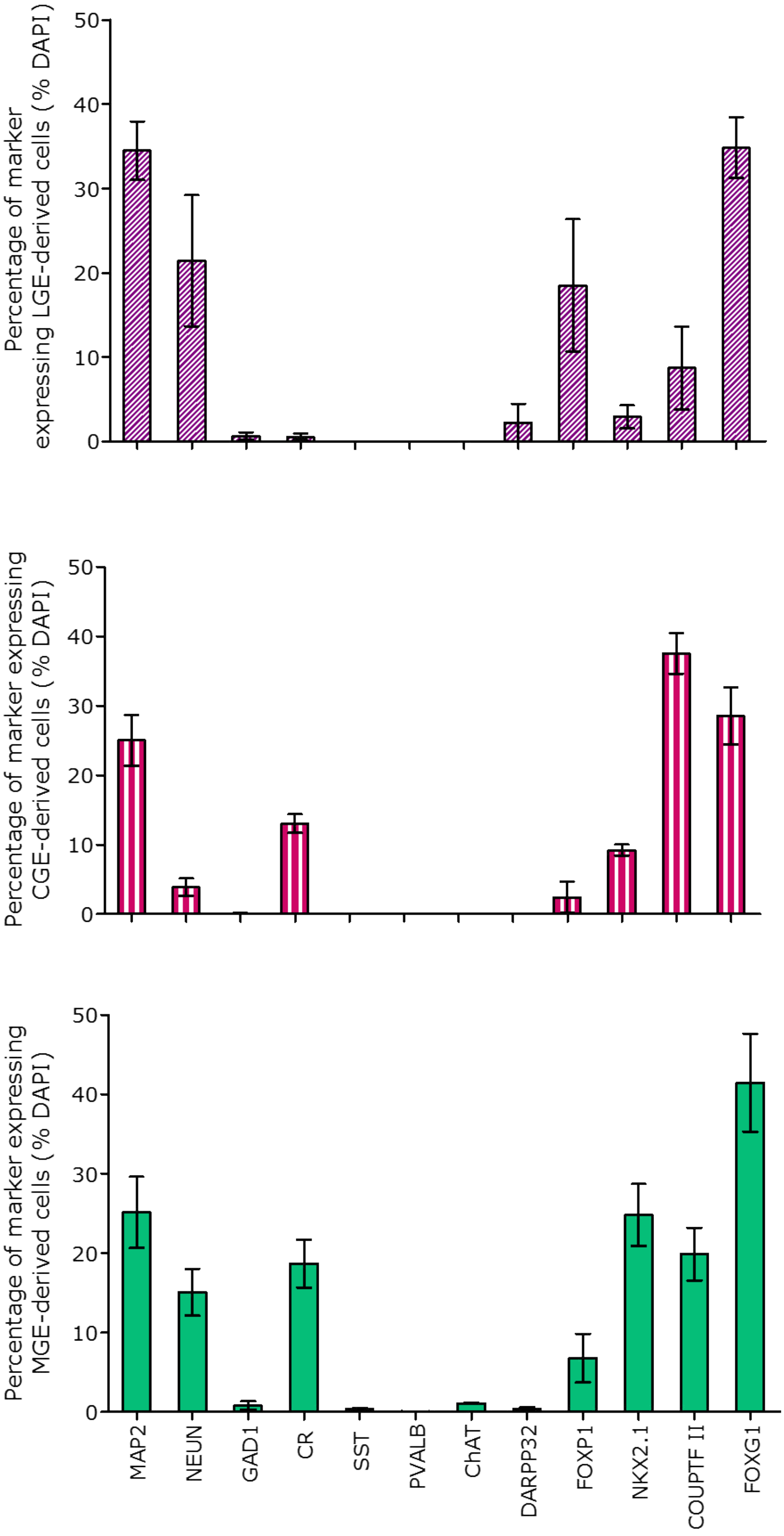

15 WPC MGE
15 WPC CGE
15 WPC LGE
CR, COUPTF II

Figure 5.1C



Percentage of marker expressing LGE-derived cells (% DAPI)
Percentage of marker expressing CGE-derived cells (% DAPI)
Percentage of marker expressing MGE-derived cells (% DAPI)
0
10
20
30
40
50
MAP2
NEUN
GAD1
CR
SST
PVALB
ChAT
DARPP32
FOXP1
NKX2.1
COUPTF II
FOXG1

Figure 5.1 Primary human foetal tissue characterisation of MGE, CGE and LGE: pilot data.

A) Molecular characterisation of primary human MGE, CGE and LGE – at RNA level, at 15WPC. The data has been presented in the form of expression level provided as an average delta Ct, along with +/- SEM. The delta Ct was calculated in reference to beta-actin as the standard housekeeping gene. Minimum biological replicates (independent foetal tissue) n=3. Technical replicates n=3.

B) Molecular characterisation of primary human MGE, CGE and LGE cultures– at protein level, at 15WPC. The isolated tissues were seeded and maintained for 2-3 days in vitro before fixing. The data has been presented in the form of a panel of deconvoluted images for the immunostaining against CR (green) and COUPTF II (purple). The CGE, MGE and LGE were obtained from the same source. Biological replicates (independent foetal tissue) n=1. Minimum technical replicates n=6. The total number of cells counted per technical replicate can be found in Appendix folder 7.

C) Molecular characterisation of primary human MGE, CGE and LGE cultures– at protein level, at 15WPC. The isolated tissues were seeded and maintained for 2-3 days in vitro before fixing. The graph shows the mean percentage of marker expressing cells out of the total cell population (total cell count measured using DAPI count), along with +/- SEM. The CGE, MGE and LGE were obtained from the same source. Biological replicates (independent foetal tissue) n=1. Minimum technical replicates n=6. The total number of cells counted per technical replicate can be found in Appendix folder 7.

CGE, caudal ganglionic eminence. MGE, medial ganglionic eminence. LGE, lateral ganglionic eminence. WPC, weeks post-conception. NeuN, neuronal nuclei. SST, somatostatin. PVALB, parvalbumin. CR, calretinin. NK2 Homeobox 1, NKX2.1. COUP transcription factor 2, COUPTF II. ChAT, choline acetyltransferase. GAD1, glutamate decarboxylase 1. MAP2, microtubule-associated protein 2. DARPP32, dopamine- and cAMP-regulated neuronal phosphoprotein. FOXP1, Forkhead Box P1. FOXG1, Forkhead Box G1. LHX6, LIM Homeobox 6. LHX8, LIM Homeobox 8. DLX5, Distal-less homeobox 5. COUP-TF-Interacting Protein 2, CTIP2. SRY-Box 6, SOX6. Gamma-aminobutyric acid-ergic, GABAergic. SEM, standard error of the mean.

5.3.2 A discrete set of interneuron subtypes emerged within the human MGE

Initial inspection of the single-cell RNA sequencing data began with each of the MGE samples being checked for the expression of known MGE and CGE markers (Figure 5.2A-5.2B). The proportion of marker expressing cells was then averaged for the entire data set (MGE samples of all ages) to provide a principal characterisation of the human MGE (Figure 5.2A-5.2B). The human MGE (10WPC to 15WPC) had a large population of *NKX2.1* expressing cells, which equated to an average of 60% across samples (Figure 5.2E). Out of these *NKX2.1* expressing cells there was high co-expression within known late MGE markers, including *DLX1/5* and *LHX6* (Figure 5.2A). This would indicate that the MGE had a larger proportion of transitioning progenitors and interneuron precursors (*LHX6, DLX5* expressing), then early proliferating progenitors (*NKX2.1* expressing) during this developmental period. Furthermore, the detection of *NKX2.1/LHX6* and *NKX2.1/COUPTF II* co-expressing subpopulations likely corresponded to ventrally and dorsally located MGE progenitors respectively, with a good representation of both in the samples (Figure 5.2A). On characterising the post-mitotic neuronal population obtained from all MGE samples, *NPY* was the most abundant subtype marker (60%), followed by *SST* (30%) and *CR* (10%) (Figure 5.2B). In addition to *NPY* neurons, co-expression of *NPY* with *SST* and *CR* were also detected in the population, although these subpopulations were rarer than the *NPY* expressing interneurons without a secondary subtype marker (Figure 5.2B). *SST/CR* co-expressing neurons were also detected in the population (5%), and were likely derived from the dorsal MGE *NKX2.1/COUPTF II* co-expressing subpopulation described above (Figure 5.2B). Whereas, the proportion of CGE-derived interneuron subtype markers, *VIP, CCK, RELN* and *nNOS*, and late MGE-derived interneuron subtype marker *PVALB* were at low abundance (<5%) or absent across samples (Figure 5.2B). Moreover, no *MAP2* positive cells were found to express the CGE restricted marker *5HT-R 3A* (Figure 5.2B). The favoured expression of *NPY* and *SST* would indicate a high proportion of MGE-derived interneurons in each sample, with little contamination from the CGE (which was in agreement to pilot data presented above, Figure 5.1A-C).

To further explore the molecular profile of interneuron subtypes in the MGE sample, the three most abundant subtypes detected, *SST, NPY* and *CR*, were compared to identify genes uniquely expressed by each one (Appendix folder 3). These genes included transcription factors that may have directed the lineage, for example S-phase kinase associated protein 1 (*SKP1*) was uniquely expressed by *NPY* interneurons (Appendix folder 3). Many of the genes uniquely co-expressed with *SST* or *CR* interneurons were also identified in hESC-derived interneuron analysis (Chapter 4), and included intracellular signalling proteins (GNAS complex locus, small EDRK-rich factor 2 , casein kinase 1 epsilon), chaperons (heat shock protein 90 alpha family class B member 1) and cytoskeletal motor proteins (dynein light chain LC8-type 1) (Appendix folder 3). This would provide an initial favourable comparison between the in vitro and in vivo derived interneurons.

5.3.3 The human MGE was composed of a diverse population

The molecular heterogeneity within the human MGE for each developmental age sampled has been represented in a series of tSNE plots (Figure 5.2C). Within each time point there appeared to be two or three distinct clusters (labelled C1, C2 or C3, Figure 5.2C). Each cluster represented a molecularly distinct subpopulation in the MGE, with the cluster size and degree of segregation of each subpopulation variable (Figure 5.2C). This would indicate that some single-cell expression profiles were rarer, and showed greater differences to others. The full list of differentially expressed genes between clusters for each age can be found in Appendix folder 3, with a single example of highly differentially expressed genes shown in Figure 5.2C. To explore heterogeneity based on between-subject variability, tSNE plot was generated for the two independently sourced 12WPC MGE samples A and B (Figure 5.2C). The tSNE plot showed a general clustering pattern based on foetal tissue source (A or B), indicating a molecular heterogeneity largely driven by between-subject variability (Figure 5.2C). This would encourage a greater number of independent samples to be used in future experiments to limit the effects of between-subject variability clouding the outcome of the data analysis.

Finally, when all the MGE single-cell samples from 10WPC to 15WPC were analysed together to generate a single tSNE plot, the samples were found to be broadly overlapping in gene expression profiles (Figure 5.2D). As seen in Figure 5.2D, single-cell samples from each time point distributed themselves into a least two visible clusters, with samples from each time point present in both subpopulations, although with a slight age-dependent bias. This may represent a division according to the maturity of the MGE cells, because samples from earlier developmental ages (10WPC and 12WPC derived MGE) were more abundant in one cluster (bottom right cluster) and 14WPC-15WPC derived MGE single-cells were more abundant in the other (top left cluster) (Figure 5.2D). This would suggest that there were differences in gene expression profiles based on the age of the sample.

5.3.4 Subtype dictated subpopulations become more abundant over time, adding to the age-dependent heterogeneity

At each time point there was a mix of single-cell profiles belonging to neuronal and non-neuronal (glial precursors) lineages. For example, single-cell subpopulations were found to be segregated according to the differential expression of meteorin, a glial cell differentiation regulator (Appendix folder 3 table 11 for full list of genes).

Single-cell RNA sequencing data detected variability between cells, at each time point and across samples, in terms of the expression level of MGE-specific markers and the proportion of cells expressing these markers in the population (Figure 5.2E-F). In particular, the proportion of *COUPTF II* expressing cells decreased in older samples, which may reflect a slight ventral bias in the dissection of the older MGE samples (especially 14WPC) (Figure 5.2E-F). *LHX8* average TPM values and abundance of *LHX8* expressing cells showed a progressive decline over developmental age (Figure 5.2E). The depletion of MGE-derived subpopulation expressing *LHX8* from 12WPC onwards may represent a temporal shift in the lineage commitment of progenitors over this period, as *LHX8* is strongly associated with ChAT interneuron lineage (which were not abundantly detected in these samples). *LHX6* and *NKX2.1* showed a mirrored change in abundance over the developmental period, which would support previous reports demonstrating *NKX2.1*-driven upregulation of *LHX6* expression in MGE progenitors (Figure 5.2E).

Further variability was detected in the expression of interneuron subtype markers, *NPY*, *CR* and *SST* (Figure 5.2F). *CR* expressing cells (out of *MAP2*) progressively increased in abundance between 13WPC to 15WPC (from approx. 10% to 35%, Figure 5.2F). The increase in *CR* positive population was also accompanied by *CR* average transcript level increasing (Figure 5.2F). A similar trajectory was reported for *NPY*, which was the most abundant subtype present at each time point (peaking at 85% at 15WPC), and progressively increased in abundance over time (Figure 5.2F). Furthermore, the transcript levels of *NPY* were, in general, the highest at each time point relative to *CR* and *SST* (Figure 5.2F). Whereas, between 10WPC to 15WPC samples, *SST* abundance and average transcript levels typically remained consistent at approximately 30-40% abundance and 100 TPM expression level, with only a large deviation from this range at 13WPC (Figure 5.2F). This may be due to between-subject variability, as only a single independent foetal tissue sample was used for this time point.

In conclusion, although each sample had a full spectrum of cells along interneuron development, from highly proliferative early MGE progenitors (*SOX2, NKX2.1, FOXG1* positive), to immature interneurons (*NPY, SST, CR* expressing), older MGE samples had a greater abundance of the latter (Figure 5.2E-5.2F).

5.3.5 The developmental trajectory of the MGE could be distilled down to a shortlist of differentially expressed genes

To identify the genes significantly differentially expressed across the developmental period, a series of pairwise ANOVAs were carried out between samples (Figure 5.2G). Only pairwise comparisons made relative to 10WPC have been presented here, although the full list for all pairwise comparisons can be found in Appendix folder 3 (Figure 5.2G). Any genes shown as a positive fold change (upregulated over time) were relatively higher expressed at later developmental ages (e.g. 15WPC in 10WPC vs 15WPC comparison). The reciprocal is true for genes assigned a negative fold change value (downregulated over time), which were relatively higher expressed at 10WPC (Figure 5.2G). In general, only a small number of genes (<6) were found to be significantly differentially expressed between 10WPC MGE population and older samples, with the exception being for the 10WPC and 15WPC comparison (Figure 5.2G). This would suggest that only a small temporally-induced change had occurred in MGE profile over 10WPC to 14WPC period. Whereas, a relatively greater age-dependent transition had occurred by 15WPC in the MGE, with 14 genes detected as significantly differentially expressed between 10WPC and 15WPC sample (Figure 5.2G). A small proportion of these genes presented in Figure 5.2G had also been detected as differentially expressed in other pairwise comparisons, including X inactive specific transcript (*XIST,* highest upregulated gene at 13WPC, 14WPC and 15WPC) and cyclin D1 (*CCND1,* the highest downregulated gene at 13WPC, 14WPC and 15WPC) (Figure 5.2G). Many of the uniquely differentially expressed genes between 10WPC and 15WPC MGE samples encoded for transcription factors, which were downregulated at 15WPC (relative to 10WPC) (Figure 5.2G). Theses transcription factors included, CAMP responsive element binding protein 5 (*CREB5*) and nuclear receptor subfamily 2 group E member 1 (*NR2E1*), and may be linked to the temporal dynamics of gene expression in MGE ventricular zone responsible for generating progenitors committed to different lineages (Figure 5.2G). Therefore, these identified transcription factors add to the list of candidate lineage determining transcription factors, although which lineage was undetermined.

Additional genes identified as differentially expressed relative to 10WPC sample were associated with specific signalling pathways, including insulin-like growth factor binding protein 2 (*IGFBP2*), neuroendocrine signalling protein secretogranin V, and activin and BMP4 signalling inhibitor, follistatin like 1 (Figure 5.2G). These signalling pathways may act indirectly to alter gene expression, and could potentially have a critical role in directing interneuron fate.

Moreover, many of the downregulated genes (relative to 10WPC) were related to RNA and protein synthesis (Figure 5.2G). These genes included the enzyme responsible for the conversion of deoxyribonucleotides from ribonucleotides, ribonucleotide reductase regulatory subunit M2 (*RRM2*), membrane trafficking protein Yip1 interacting factor homolog B (*YIF1B)* and adaptor related protein complex 1 subunit sigma 2 (*AP1S2*) responsible for protein sorting in trans-Golgi/endosomes (Figure 5.2G). The relatively higher expression of these genes at 10WPC would indicate the transition from a population abundant in proliferating progenitors at 10WPC, to a population with a greater proportion of immature interneurons (which have a reduced rate of both transcription and translation) by 15WPC.

Figure 5.2A

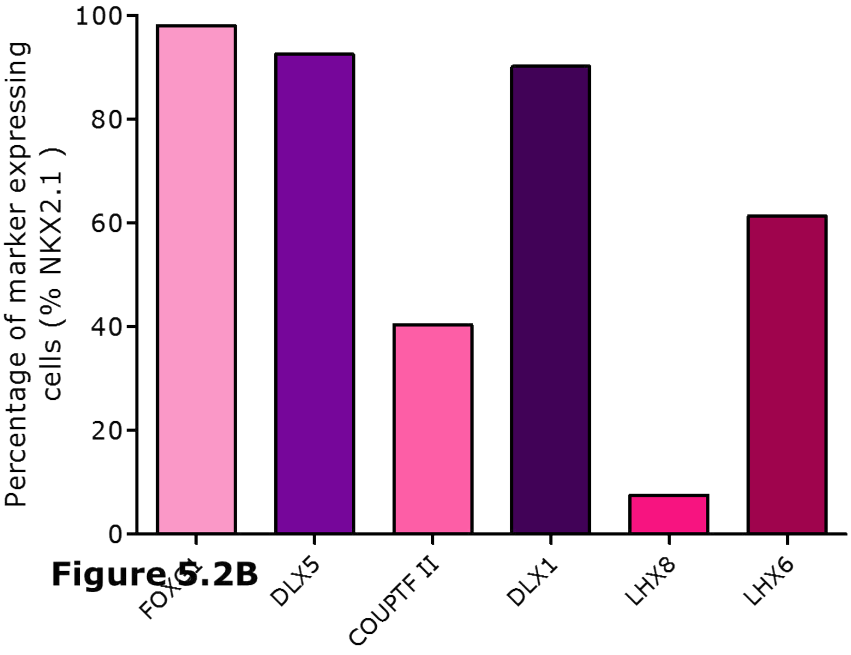


Figure 5.2B

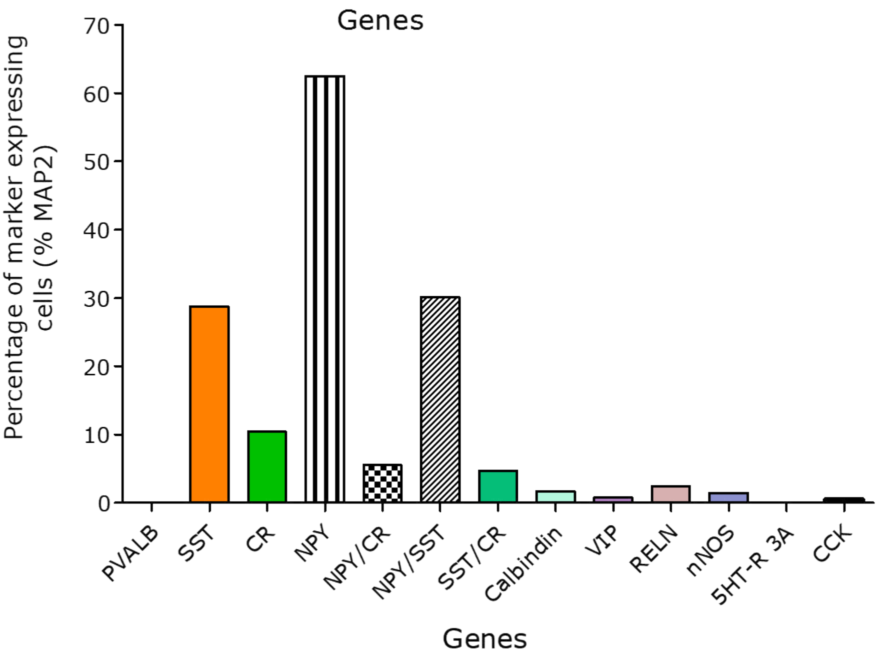
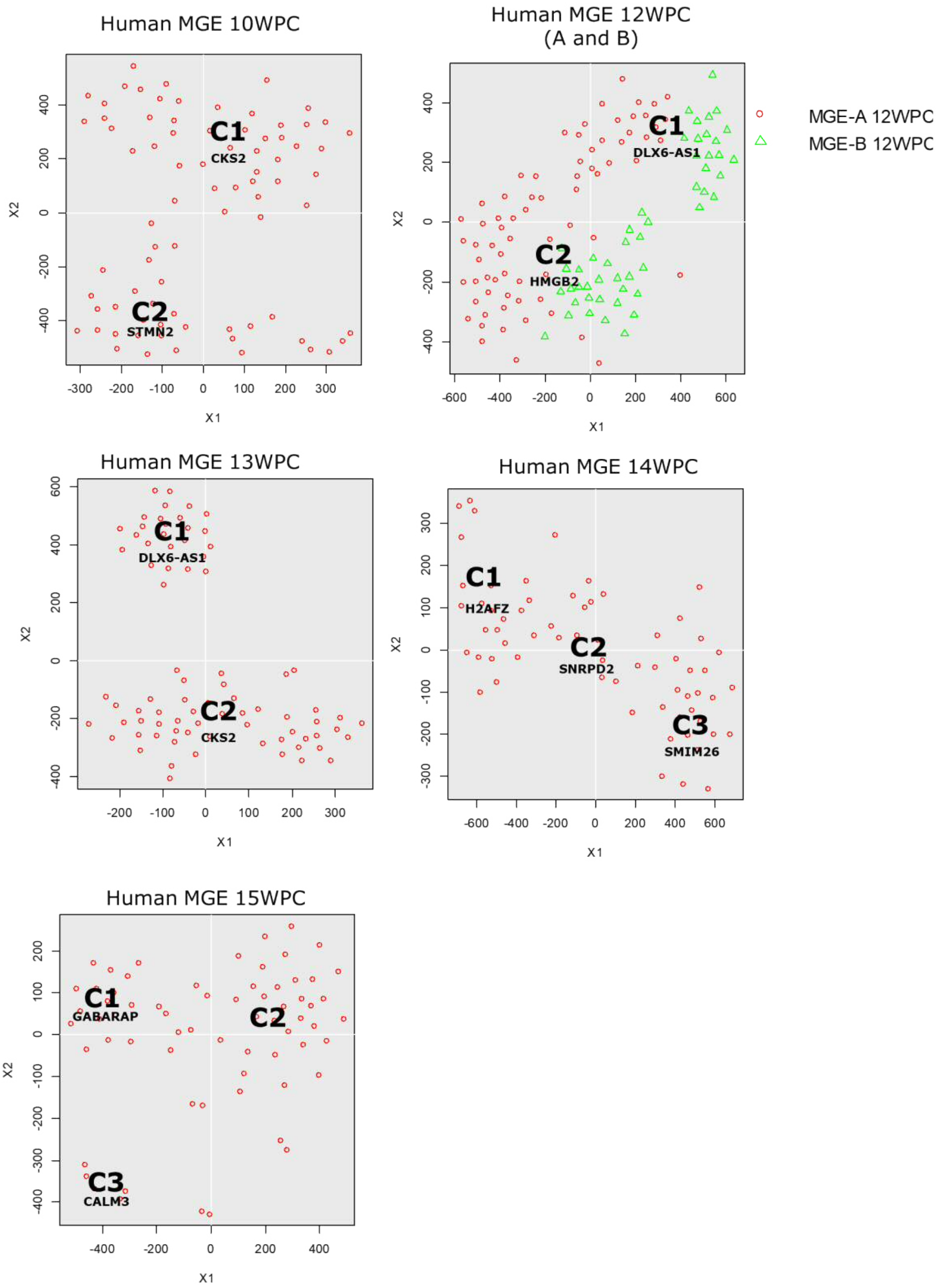


Figure 5.2C



Human MGE 10WPC
C1
CKS2
C2
STMN2
X1
X2
Human MGE 12WPC
(A and B)
C1
DLX6-AS1
C2
HMGB2
MGE-A 12WPC
MGE-B 12WPC
Human MGE 13WPC
C1
DLX6-AS1
C2
CKS2
Human MGE 14WPC
C1
H2AFZ
C2
SNRPD2
C3
SMIM26
Human MGE 15WPC
C1
GABARAP
C2
C3
CALM3

Figure 5.2D

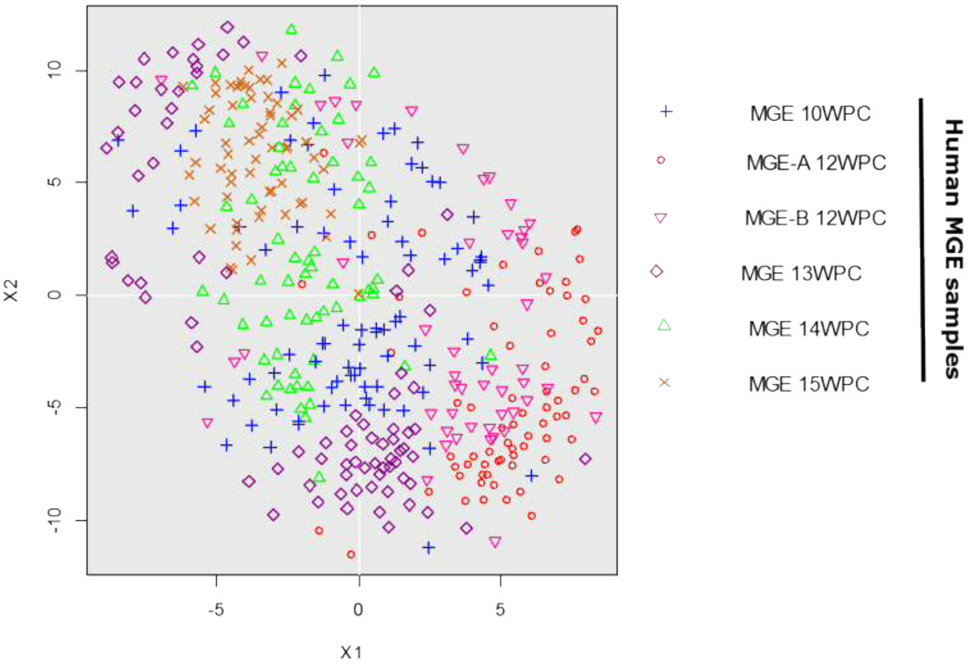

10
5
0
-5
-10
X2
-5
0
5
X1
MGE 10WPC
MGE-A 12WPC
MGE-B 12WPC
MGE 13WPC
MGE 14WPC
MGE 15WPC
Human MGE samples

Figure 5.2E

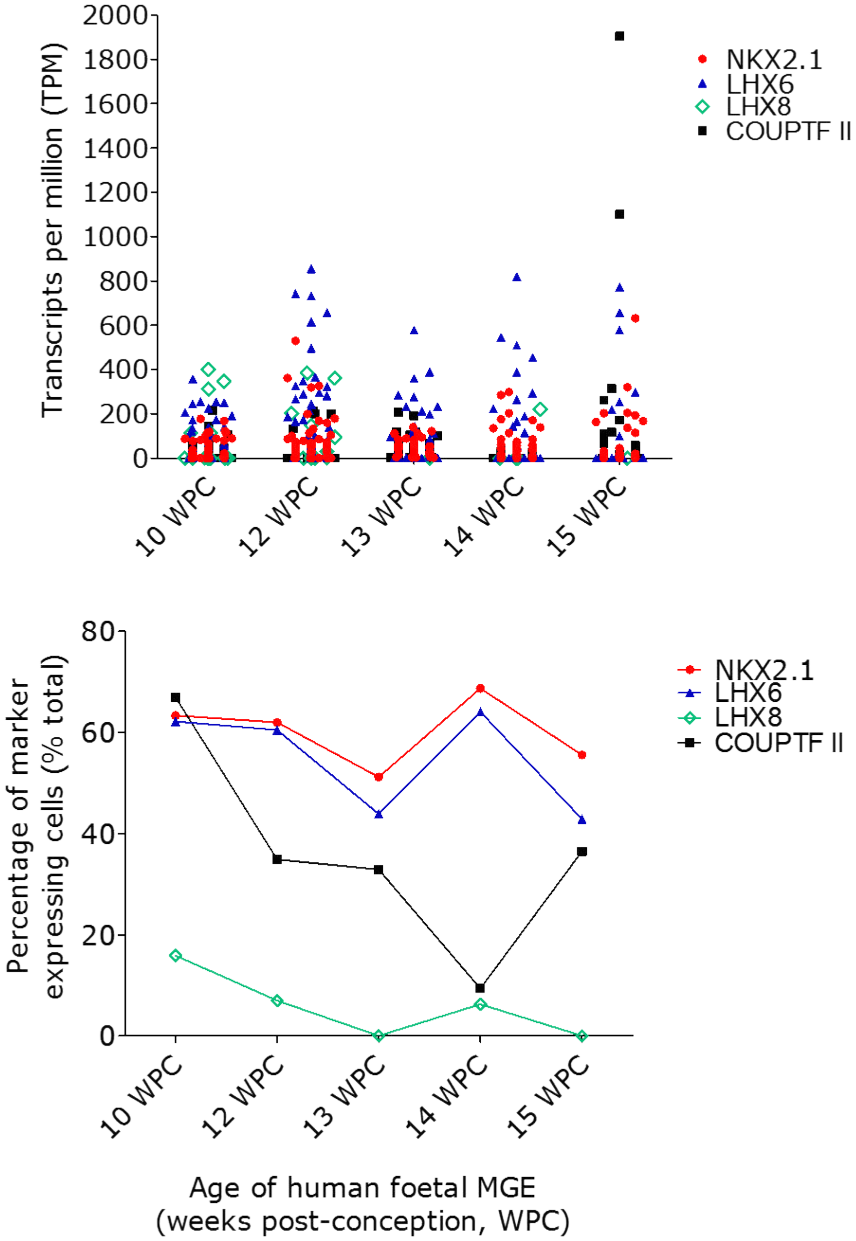

Transcripts per million (TPM)
2000
1800
1600
1400
1200
1000
800
600
400
200
0
10 WPC
12 WPC
13 WPC
14 WPC
15 WPC
NKX2.1
LHX6
LHX8
COUPTF II
Percentage of marker expressing cells (% total)
80
60
40
20
0
10 WPC
12 WPC
13 WPC
14 WPC
15 WPC
NKX2.1
LHX6
LHX8
COUPTF II
Age of human foetal MGE (weeks post-conception, WPC)

Figure 5.2F

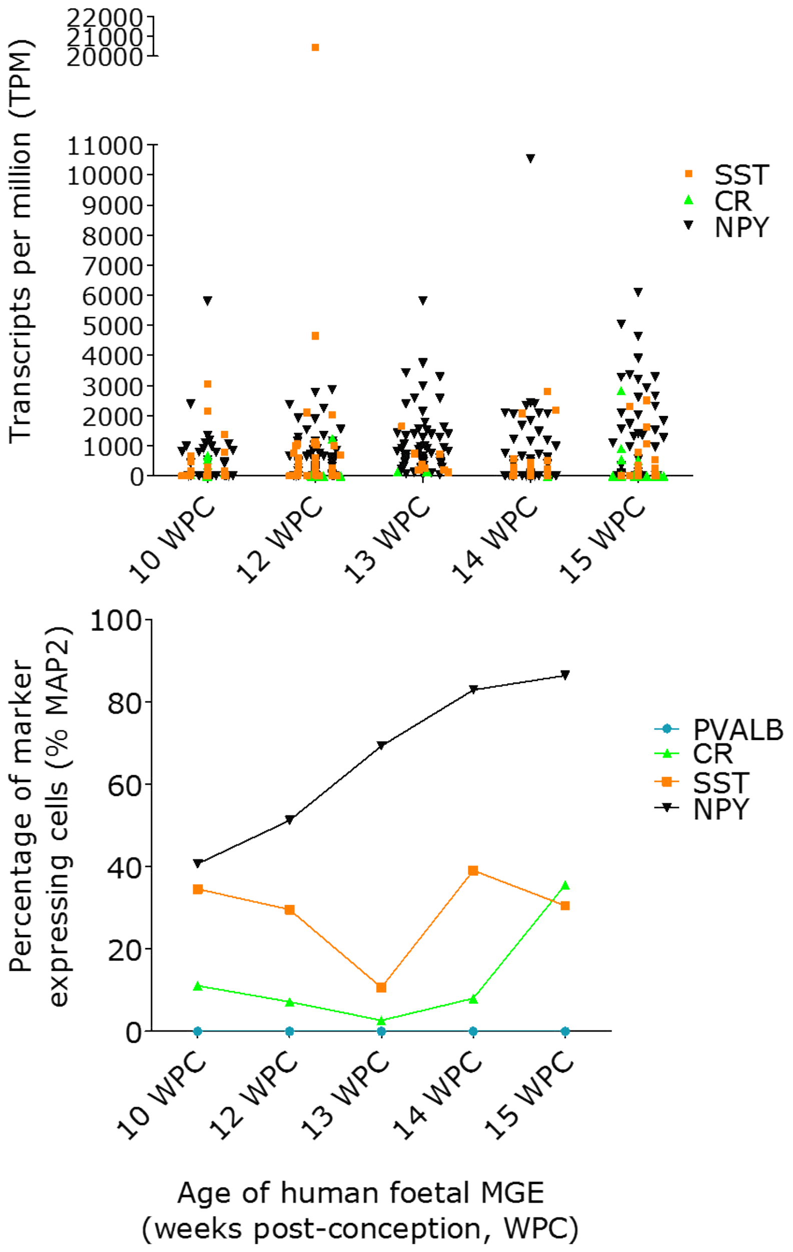

Transcripts per million (TPM)
22000
21000
20000
11000
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
SST
CR
NPY
10 WPC
12 WPC
13 WPC
14 WPC
15 WPC
Percentage of marker expressing cells (% MAP2)
100
80
60
40
20
0
PVALB
CR
SST
NPY
10 WPC
12 WPC
13 WPC
14 WPC
15 WPC
Age of human foetal MGE (weeks post-conception, WPC)

Figure 5.2G

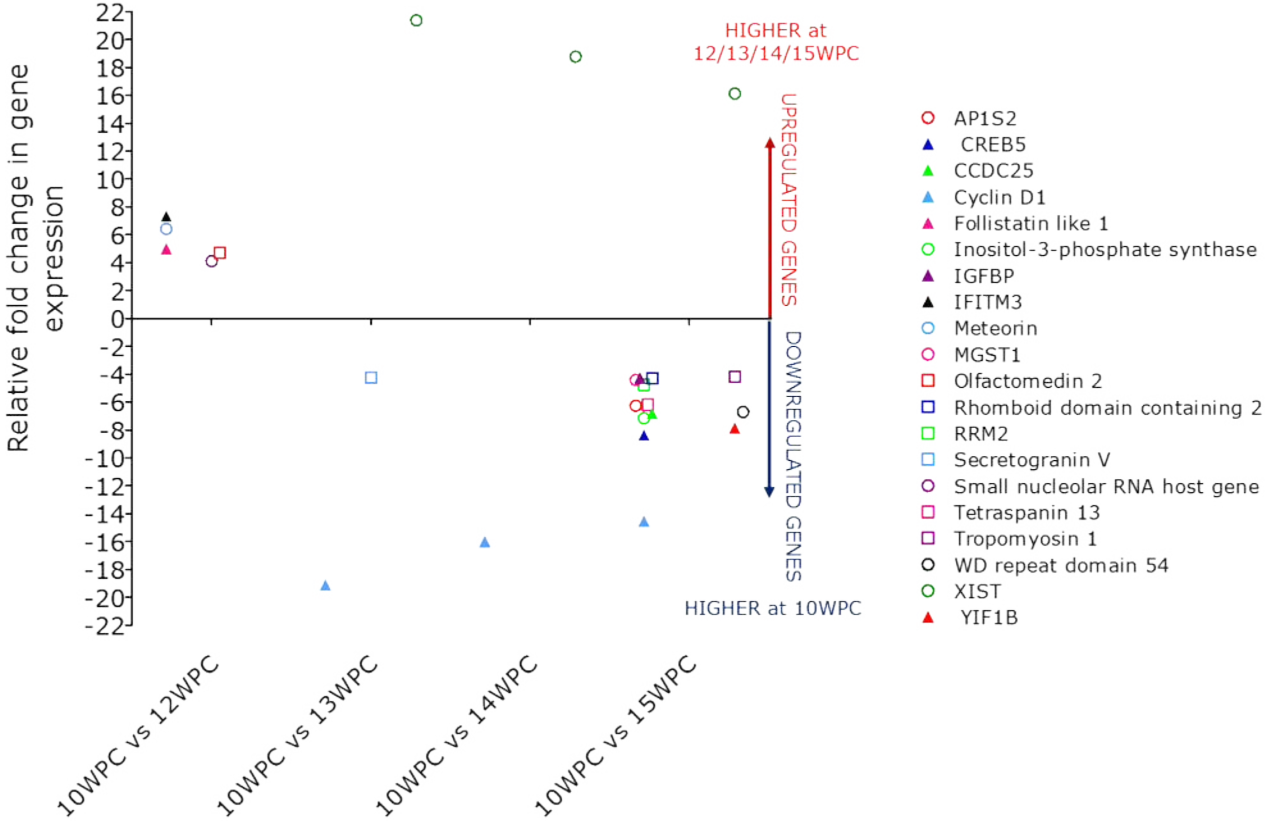

Relative fold change in gene expression
22
20
18
16
14
12
10
8
6
4
2
0
-2
-4
-6
-8
-10
-12
-14
-16
-18
-20
-22
10WPC vs 12WPC
10WPC vs 13WPC
10WPC vs 14WPC
10WPC vs 15WPC
HIGHER at 12/13/14/15WPC
UPREGULATED GENES
DOWNREGULATED GENES
HIGHER at 10WPC
AP1S2
CREB5
CCDC25
Cyclin D1
Follistatin like 1
Inositol-3-phosphate synthase
IGFBP
IFITM3
Meteorin
MGST1
Olfactomedin 2
Rhomboid domain containing 2
RRM2
Secretogranin V
Small nucleolar RNA host gene
Tetraspanin 13
Tropomyosin 1
WD repeat domain 54
XIST
YIF1B

Figure 5.2 Single-cell RNA sequencing analysis of primary human foetal MGE.

A-B) Molecular characterisation of primary human MGE– at single-cell RNA level, total for all MGE samples (10WPC -15WPC). The graph plotted shows the proportion of cells expressing each named marker. The marker positive cells have been counted if the cells also co-expressing 5.2A) NKX2.1 or 5.2B) MAP2, to restrict analysis to MGE progenitors and post-mitotic neurons respectively. Minimum biological replicates (independent foetal tissue per age) n=1. 5.2A) Total number of cells expressing NKX2.1 =253. 5.2B) Total number of cells expressing MAP2 = 403 cells.

C-D) tSNE plot of the relative similarity in single cell profiles of primary human MGE at 10WPC, 12WPC, 13WPC, 14WPC, 15WPC. Each symbol on tSNE plot represents a single cell and the unique gene expression profile of the cell (based on single-cell RNA sequencing data available, trimmed to top 400 genes). 5.2C) Each cluster uniquely labelled with cluster number (e.g. C1) and an example of a highly expressed gene unique to this cluster (where an example exists). For full list of differentially expressed genes between cluster see Appendix folder 3 table 11. 5.2D) All samples presented on a single tSNE plot. Cells from the same MGE sample share the same symbol (e.g. all 12WPC MGE from sample A cells are represented by red open circle). Minimum biological replicates (independent foetal tissue per age) n=1. Total number of cells per age can be found in Appendix folder 3 table 10.

E) Trajectory of primary human MGE/CGE markers – at RNA level, at 10WPC, 12WPC, 13WPC, 14WPC, 15WPC. The data has been presented as a quantification of the expression level of selected markers (TPM) and the proportion of cells expressing each named marker. On the graph displaying the TPM values, each symbol plotted represents the expression value of that marker for a single cell. Minimum biological replicates (independent foetal tissue per age) n=1. Total number of cells per age can be found in Appendix folder 3 table 10.

F) Trajectory of primary human MGE derived interneurons subtypes – at RNA level, at 10WPC, 12WPC, 13WPC, 14WPC, 15WPC. The data has been presented as a quantification of the expression level of selected markers (TPM) and the proportion of cells expressing each subtype marker. On the graph displaying the TPM values, each symbol plotted represents the expression value of that marker for a single cell. The marker positive cells have been counted if the cells also co-expressed MAP2, to restrict analysis to post-mitotic neurons. Minimum biological replicates (independent foetal tissue per age) n=1. Total number of MAP2 expressing cells per age: 81 cells 10WPC, 125 cells 12WPC, 75 cells 13WPC, 64 cells 14WPC and 59 cells 15WPC.

G) Relative fold change in expression of genes significantly differentially expressed between 10WPC and 12/13/14/15 WPC. Differentially expressed genes were determined through ANOVA pairwise analysis, and were included if $p<0.05$ and fold change (either up or downregulated) >4. In this plot, any genes upregulated had a relatively higher expression at the later age (e.g. 15WPC in 10WPC vs 15WPC comparison) and were shown as a positive fold change. Any downregulated genes were higher expressed at 10WPC (relative to the later age in the pair) and were shown as a negative fold change. See Appendix folder 3 tables for list. Minimum biological replicates (independent foetal tissue per age) n=1. Total number of cells per age can be found in Appendix folder 3 table 10.

Single independent source of tissue for each age of primary human MGE, except for 12WPC which had two independent tissues (labelled A and B). TPM, transcripts per million. WPC, weeks post-conception. NK2 Homeobox 1, NKX2.1. COUP transcription factor 2, COUPTF II. LHX6, LIM Homeobox 6. FOXG1, Forkhead Box G1. SST, somatostatin. PVALB, parvalbumin. CR, calretinin. LHX8, LIM Homeobox 8. DLX5, Distal-less homeobox 5. DLX1, distal-less homeobox 1. MAP2, microtubule-associated protein 2. Serotonin receptor 3A, 5HT-R 3A. Vasoactive intestinal peptide, VIP. Reelin, RELN. Neuropeptide Y, NPY. Neuronal nitric oxide synthase, nNOS. MGE, medial ganglionic eminence. Adaptor related protein complex 1 subunit sigma 2, AP1S2. CAMP responsive element binding protein 5, CREB5. Coiled-coil domain containing 25, CCDC25. Insulin like growth factor binding protein 2, IGFBP. Interferon induced transmembrane protein 3, IFITM3. Microsomal glutathione S-transferase 1, MGST1. Ribonucleotide reductase regulatory subunit M2, RRM2. X inactive specific transcript, XIST. Yip1 interacting factor homolog B, membrane trafficking protein, YIF1B. GABA type A receptor-associated protein, GABARAP. Calmodulin 3, CALM3. H2A.Z variant histone 1, H2AFZ. Small integral membrane protein 26, SMIM26. Small nuclear ribonucleoprotein D2 polypeptide, SNRPD2. CDC28 protein kinase regulatory subunit 2, CKS2. Distal-less homeobox 6 antisense 1, DLX6-AS1. High mobility group box 2, HMGB2. Stathmin 2, STMN2.

5.3.6 The MGE derived single-cell cDNA passed quality control checks

Prior to using the single-cell cDNA in library preparation for RNA sequencing, multiple quality control measures were performed to ensure the material for sequencing was of high standards (Figure 5.3A-5.3D and in the Appendix folder 3 figures). A large body of this work involved performing single-cell QPCR on the human MGE derived samples. This data was primarily used as a means to determine the quality of the single-cell cDNA, to prevent poor quality samples being used in single-cell libraries for RNA sequencing. Moreover, the data has provided a means to detect any inconsistencies between the two single-cell transcriptomic approaches, as the same samples were used in both methodologies. Overall, the findings were consistent between single-cell QPCR and RNA sequencing data, with a similar co-expression pattern trend described for each marker (Figure 5.2A and 5.3A). Furthermore, a hint to the heterogeneity fully explored in single-cell RNA sequencing analysis was also observed in the single-cell QPCR data, in terms of the mix of immature neurons and progenitors, and the between-cell differences in expression levels (delta Ct values) of specific markers (Figure 5.3A-5.3D and Appendix folder 3 figures).

However, there were also several differences to note. Firstly, although the trend may be consistent between the two methodologies, the actual reported proportion of marker expressing cells did not match between the two single-cell transcriptomic data sets. For example, single-cell RNA sequencing data reported almost 100% co-expression of *NKX2.1, FOXG1* and *DLX1/5* (Figure 5.2A). Whereas, single-cell QPCR reported only 60% of *NKX2.1* expressing cells co-expressed *FOXG1* and *DLX5* (Figure 5.3A). Moreover, there was a higher proportion of *NKX2.1/COUPTF II* co-expressing cells detected by single-cell RNA sequencing analysis (40% vs 17%, Figure 5.2A and 5.3A). These discrepancies may be due to primer issues in the Biomark QPCR reaction causing lower efficiency in the QPCR amplification of the transcripts, thus impairing the detection.

A final inconsistency of significance concerns the proportion of *PVALB* expressing cells. The single-cell QPCR analysis detected *PVALB* expressing cells (including *PVALB/CR* and *PVALB/SST* co-expressing cells) in the human MGE samples from 12WPC onwards, although at low abundance (<10% out of MAP2 positive cells) and low expression levels (>12 delta Ct) (Figure 5.3C-5.3D). However, no *PVALB* expressing cells were detected at RNA level using single-cell RNA sequencing (Figure 5.2F). The discrepancy may have been caused by the lower sensitivity of single-cell RNA sequencing paradigm, leading to *PVALB* transcripts being left undetected. This leaves the question over whether *PVALB* expressing cells were detected in the human MGE over this period under doubt.

Nonetheless, the single-cell QPCR analysis confirmed the high quality of the single-cell cDNA samples generated, and the overall abundance of MGE-derived cells in the sample. This conclusion enabled the samples to be carried forward for single-cell RNA sequencing analysis.

Figure 5.3A

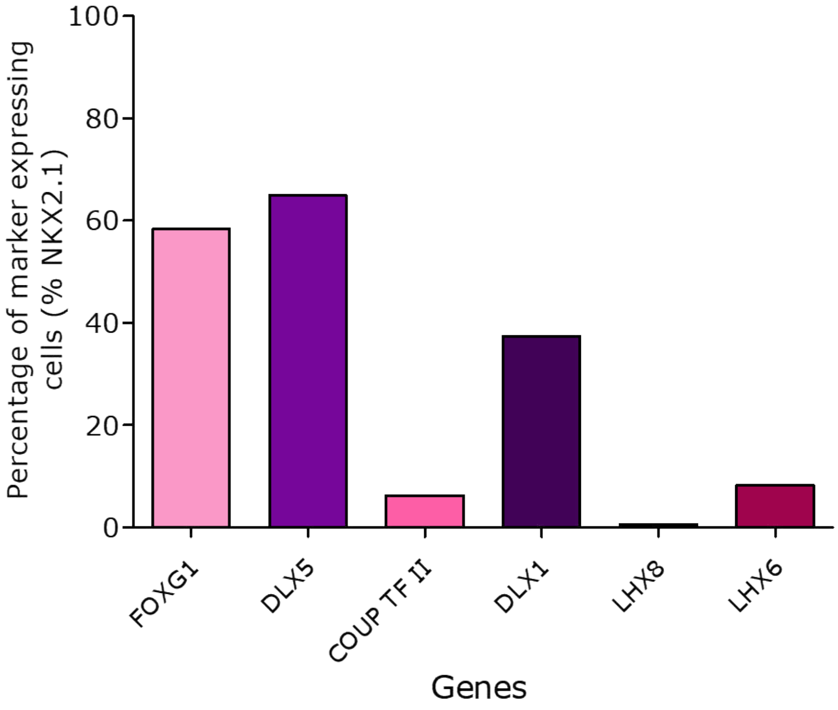


Figure 5.3B

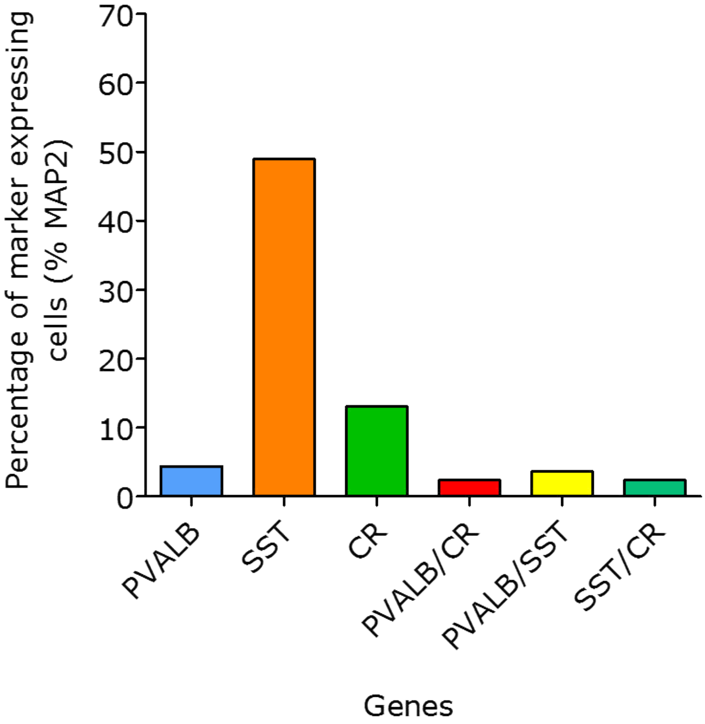


Figure 5.3C

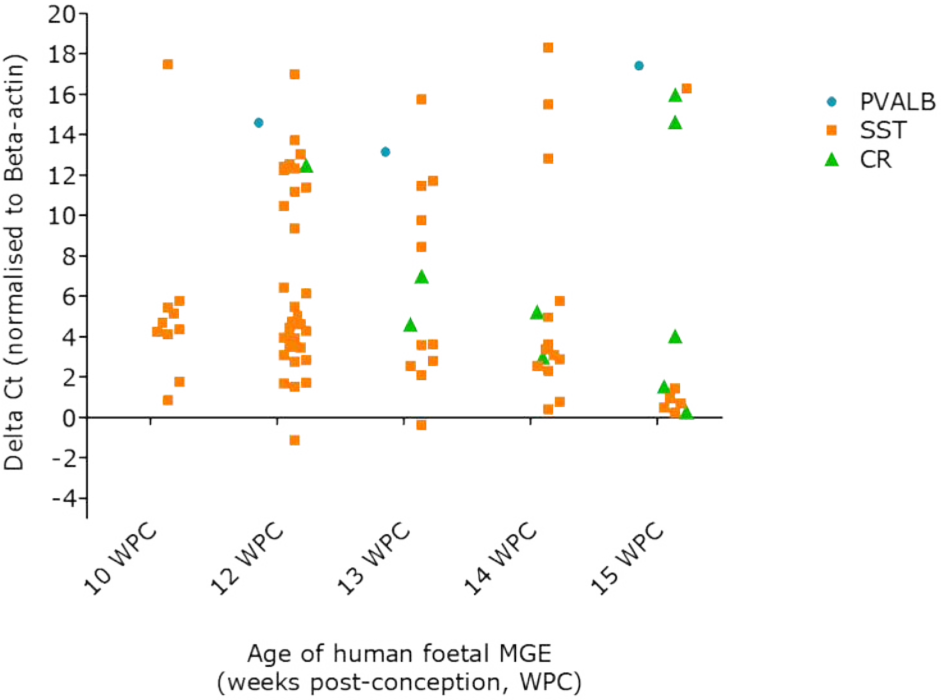


Figure 5.3D

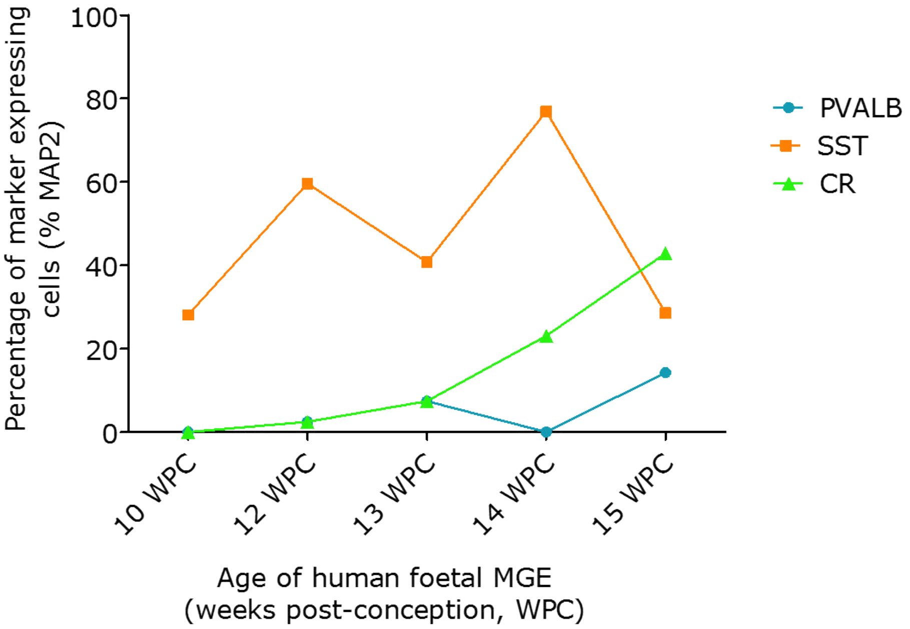


Figure 5.3 Single-cell QPCR analysis: quality control check data.

A-B) Molecular characterisation of primary human MGE– at single-cell RNA level, average across all MGE samples (10WPC -15WPC). The graph plotted shows the proportion of cells expressing each named marker. The marker positive cells have been counted if the cells also co-expressing 5.3A) NKX2.1 or 5.3B) MAP2, to restrict analysis to MGE progenitors and post-mitotic neurons respectively. Total number of cells expressing NKX2.1 =116 cells. Total number of cells expressing MAP2 =163 cells. Minimum biological replicates (independent foetal tissue per age) n=1.

C) Trajectory of primary human MGE derived interneurons subtypes – at RNA level, at 10WPC, 12WPC, 13WPC, 14WPC, 15WPC. The data has been presented as a quantification of the expression level of selected marker (delta Ct). On the graph, each symbol plotted represents the expression value of that marker for a single cell. The marker positive cells have been counted if the cells also co-expressed MAP2, to restrict analysis to post-mitotic neurons. Total number of cells expressing MAP2 at each age = 48 cells 10WPC, 48 cells 12WPC, 35 cells 13WPC,13 14WPC and 19 cells 15WPC. Minimum biological replicates (independent foetal tissue per age) n=1.

D) Trajectory of primary human MGE derived interneurons subtypes – at RNA level, at 10WPC, 12WPC, 13WPC, 14WPC, 15WPC. The data has been presented as the proportion of cells expressing each subtype marker. The marker positive cells have been counted if the cells also co-expressed MAP2, to restrict analysis to post-mitotic neurons. Total number of cells expressing MAP2 at each age = 48 cells 10WPC, 48 cells 12WPC, 35 cells 13WPC,13 14WPC and 19 cells 15WPC. Minimum biological replicates (independent foetal tissue per age) n=1.

Single independent source of tissue for each age of primary human MGE, except for 12WPC which had two independent tissues (samples A and B) and have been averaged. NK2 Homeobox 1, NKX2.1. COUP transcription factor 2, COUPTF II. LHX6, LIM Homeobox 6. FOXG1, Forkhead Box G1. SST, somatostatin. PVALB, parvalbumin. CR, calretinin. LHX8, LIM Homeobox 8. DLX5, Distal-less homeobox 5. DLX1, distal-less homeobox 1. MAP2, microtubule-associated protein 2. MGE, medial ganglionic eminence. WPC, weeks post-conception.

5.3.7 Cross-comparison to hESC-derived interneurons revealed a large source-dependent difference in expression profiles

The single-cell RNA sequencing data described in Chapter 4 for hESC-derived interneurons (day 20-day 65) and this Chapter were combined to allow a comparative assessment of the native human MGE molecular profile to one obtained in vitro (Figure 5.4A-5.4C). Analysis began by a comparison of all single-cell samples, in a single tSNE plot, to visualise the degree of similarity in single-cell expression profiles from the human MGE and hESC-derived interneuron cultures over time (Figure 5.4A). As seen in Figure 5.4A, all single-cells derived from the same source (human foetal MGE or hESC) shared a greater similarity in gene expression profiles, which led to the segregation of all MGE derived samples (bottom right) from all hESC-derived interneurons derived samples (top left). However, the hESC-derived interneuron cluster was more disperse, forming three divisions within (Figure 5.4A). The wider clustering pattern on the tSNE plot would imply a greater molecular heterogeneity in the gene expression profile of hESC-derived interneurons between day 20 to day 65, relative to the differences in the molecular profile of MGE-derived populations between 10WPC to 15WPC (Figure 5.4A). This is not surprising as the in vitro cultures showed a much sharper transition over the sample period, from one abundant in MGE-like proliferating progenitors at day 20, to an abundant post-mitotic interneuron-like population by day 35 onwards (see Chapters 3 and 4). Whereas, the single-cell expression profiles of the human MGE would suggest a greater consistency in the progenitor vs neuron population ratio, with a gradual rise in the neuronal population over time (Figure 5.2F).

Ultimately the clustering pattern on the tSNE plot was underpinned by a large number of differentially expressed genes (>150) between MGE and hESC-derived interneuron samples (Figure 5.2B-5.2C). These genes were identified through a series of pairwise ANOVAs between human MGE samples and hESC-derived interneuron samples, which named between 150-200 significantly differentially expressed genes (Appendix folder 4 tables). This number of differentially expressed genes was much greater than any other pairwise comparisons made between samples from the same source (<40 genes). Hence, the differences between the two sources outweighed the expression profile changes observed in relation to age of MGE/time along the differentiation protocol.

5.3.8 Post-mitotic interneuron-like cells showed the greatest resemblance to 15WPC MGE

To investigate whether any particular subpopulation of hESC-derived interneurons, generated across the differentiation paradigm, showed greater resemblance to a particular age of MGE, the two most molecularly distinct MGE ages (10WPC and 15WPC) were compared in turn to all hESC-derived samples (Figure 5.4B-5.4E and Appendix folder 4 tables). The comparison involved performing pairwise ANOVAs. As predicted from the source-dependent clustering pattern on the tSNE plot of hESC-derived interneurons and 10WPC MGE, a large number of genes were identified as differentially expressed between 10WPC and in vitro samples (>150 genes) (Figure 5.4B-5.4C). Due to the large number of genes, only the differentially expressed genes within the top 5 highest fold changes (in both directions) relative to 10WPC MGE have been shown as examples in Figure 5.4C (full list in Appendix folder 4 tables). All genes assigned a positive fold change value were significantly higher expressed in hESC-derived interneuron sample (in vitro) relative to 10WPC MGE (Figure 5.4C). Many of the genes with relatively higher expression in vitro were repeatedly detected across pairwise comparisons, and included calcyon neuron specific vesicular protein (*CALY*), developmental pluripotency associated 4 (*DPPA4*) and interferon induced transmembrane protein 2 (*IFITM2*) (Figure 5.4C). This would indicate a similar mis-match in the single-cell expression profiles of 10WPC MGE to each hESC-derived interneuron sample (independent of time along differentiation). The genes that showed a negative fold change in expression (relatively higher expressed in 10WPC MGE), in general encoded for transcription factors, such as *NFI B*, maternally expressed gene 3 (*MEG3,* maternally imprinted gene), *DLX5*, *HES5*, and POU Class 3 Homeobox 4 (*POU3F4*) (Figure 5.4C). The failure to induce as high expression of these transcription factors in vitro may link to the inability to achieve the wide-diversity of MGE-derived interneuron subtypes in the culture (especially *NPY* expressing neurons).

On the other hand, 15WPC MGE single-cell gene expression profiles showed a higher resemblance to hESC-derived interneurons, in particular, a subpopulation abundant in day 35 and day 65 cultures (Figure 5.4B and 5.4D). Although there were still a large number of differentially expressed genes between 15WPC MGE and hESC-derived interneuron samples (approx. 150 genes), the smallest number was detected between day 65 samples and 15WPC MGE (Appendix folder 4, table 16). This was reflected in the closer clustering of these two samples on the tSNE plot (Figure 5.4D). Certain genes were consistently found to be relatively higher or lower expressed in all hESC-derived interneuron samples relative to 15WPC MGE, and these were distinct from those described for 10WPC MGE comparisons (Figure 5.4C and 5.4E). The genes relatively higher expressed in vitro included, calcium/calmodulin-dependent protein kinase II beta (*CAMK2B*), neurotrophic factor serpin family F member 1 (*SERPINF1*), FGF2 signalling inhibitor sprouty RTK signalling antagonist 2 (*SPRY2*), calcium signalling proteins S100 beta (*S100B*) and S100 A13 (*S100A13*) (Figure 5.4C).

As previously reported for 10WPC MGE pairwise comparisons (Figure 5.4C), a large number of the genes identified as relatively higher expressed in 15WPC MGE were transcription factors (*MEG3, DLX5, NFI B, POU3F4*) (Figure 5.4E). As these transcription factors were repeatedly named in both 10WPC and 15WPC comparisons, this finding would emphasise the importance of these transcription factors in native human MGE development (Figure 5.4C and 5.4E). Similarly, a gene consistently detected across all MGE vs hESC-derived interneuron comparisons, and found to be under-represented in vitro, encoded for the paternally imprinted gene, neuronatin, which encoded for a neuronally expressed proteolipid (Figure 5.4C and 5.4E). The relatively lower expression of neuronatin (up to 1000-fold) in vitro may reflect the inability to generate this particular native MGE subpopulation (restricted to lateral MGE border) in the hESC-derived interneuron culture (Figure 5.4C and 5.4E).

In summary, encouraging similarities were found between 15WPC human MGE and post-mitotic interneurons-like cells generated in the in vitro model, at the molecular level (Figure 5.4D). Although, neither 10WPC or 15WPC MGE sample shared many similarities in expression profile with the day 20 MGE-like progenitors enriched culture, likely due to the limited capability of the in vitro differentiation paradigm to induce the expression of certain transcription factors (*MEG3, DLX5, POU3F4, NFI B*) as potently as in the MGE.

Figure 5.4A

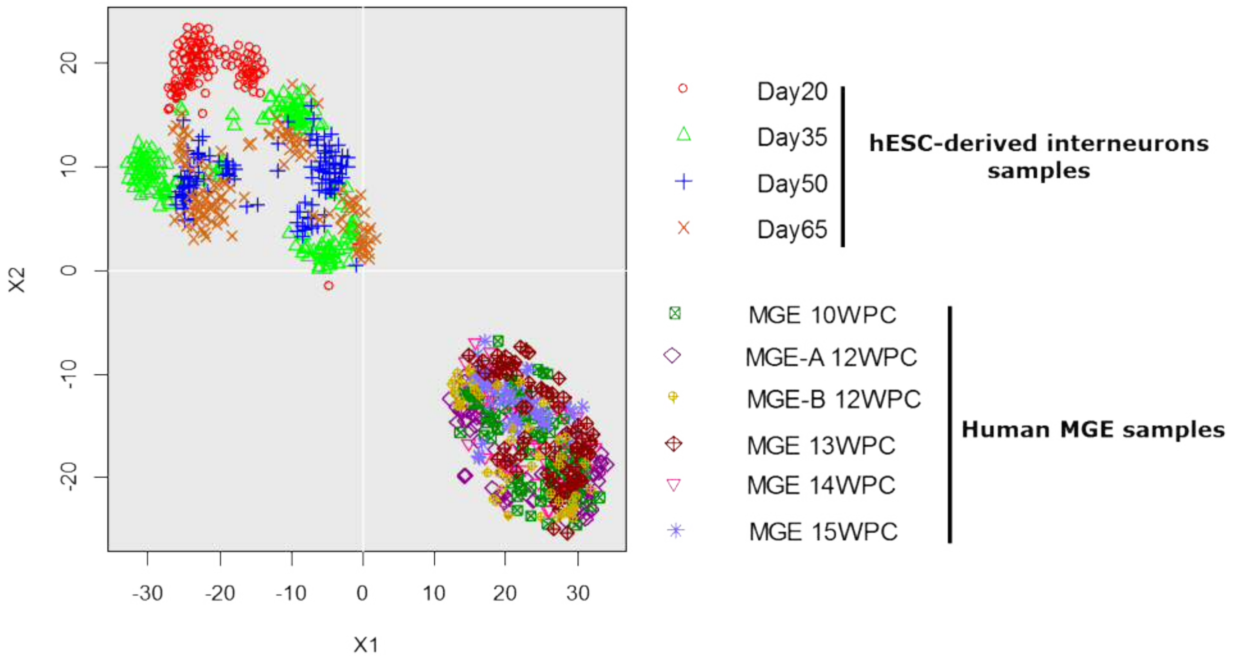

X2
20
10
0
-10
-20
-30
-20
-10
0
10
20
30
X1
○ Day20
△ Day35
+ Day50
× Day65
hESC-derived interneurons samples
⊠ MGE 10WPC
◇ MGE-A 12WPC
⊕ MGE-B 12WPC
◈ MGE 13WPC
▽ MGE 14WPC
✳ MGE 15WPC
Human MGE samples

Figure 5.4B

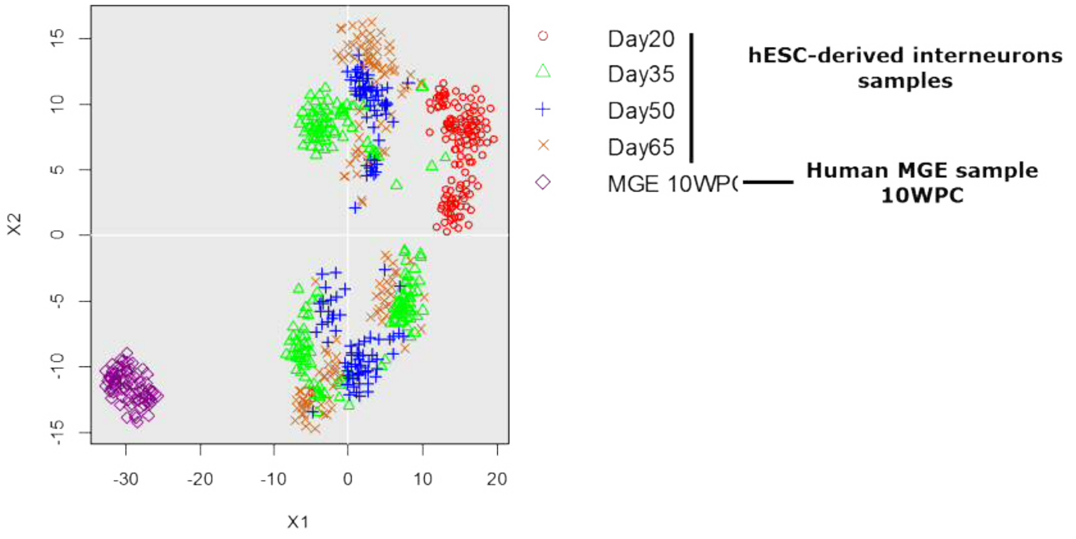


Figure 5.4C

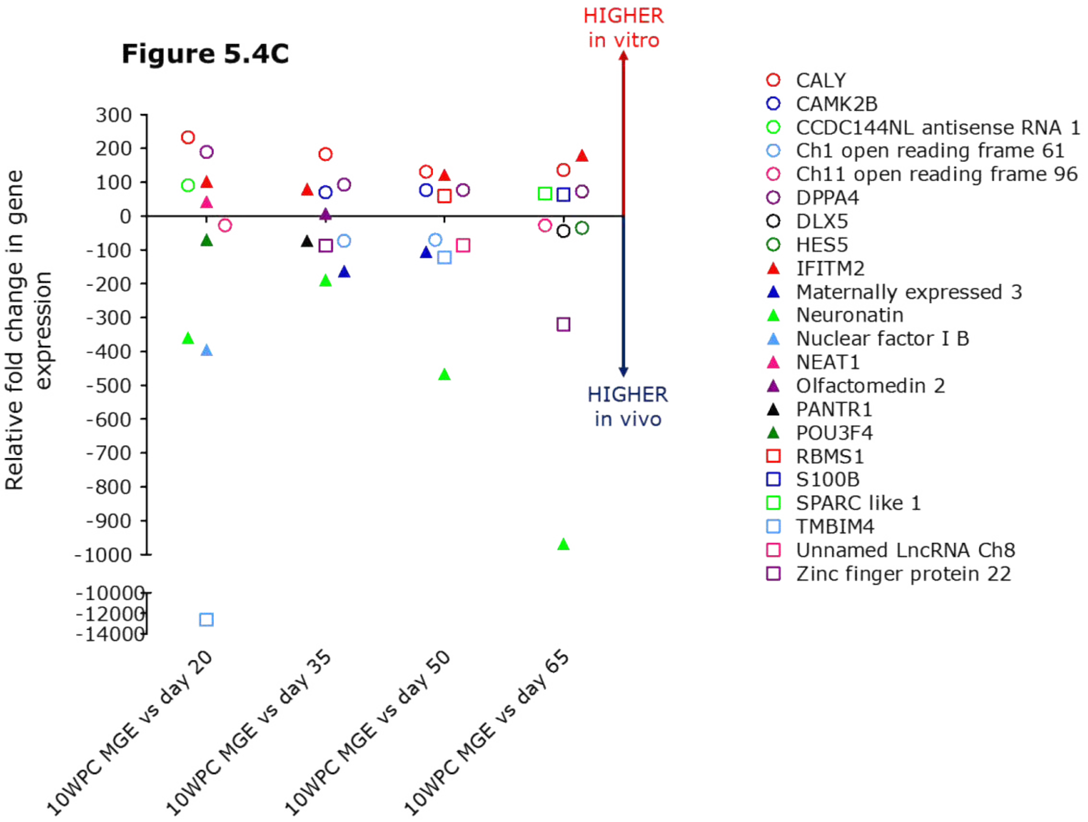


Figure 5.4D

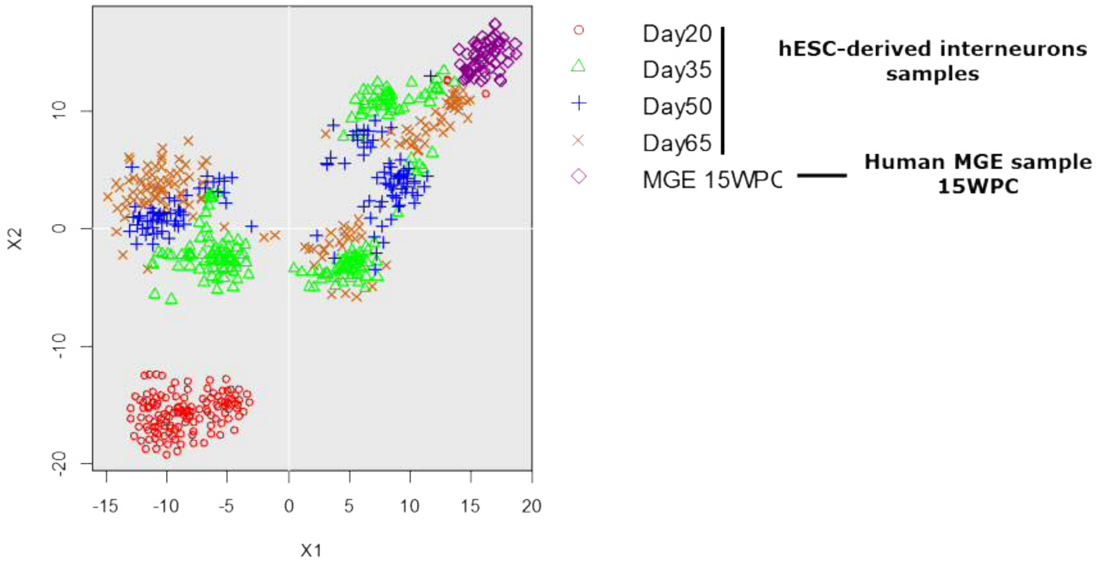


Figure 5.4E

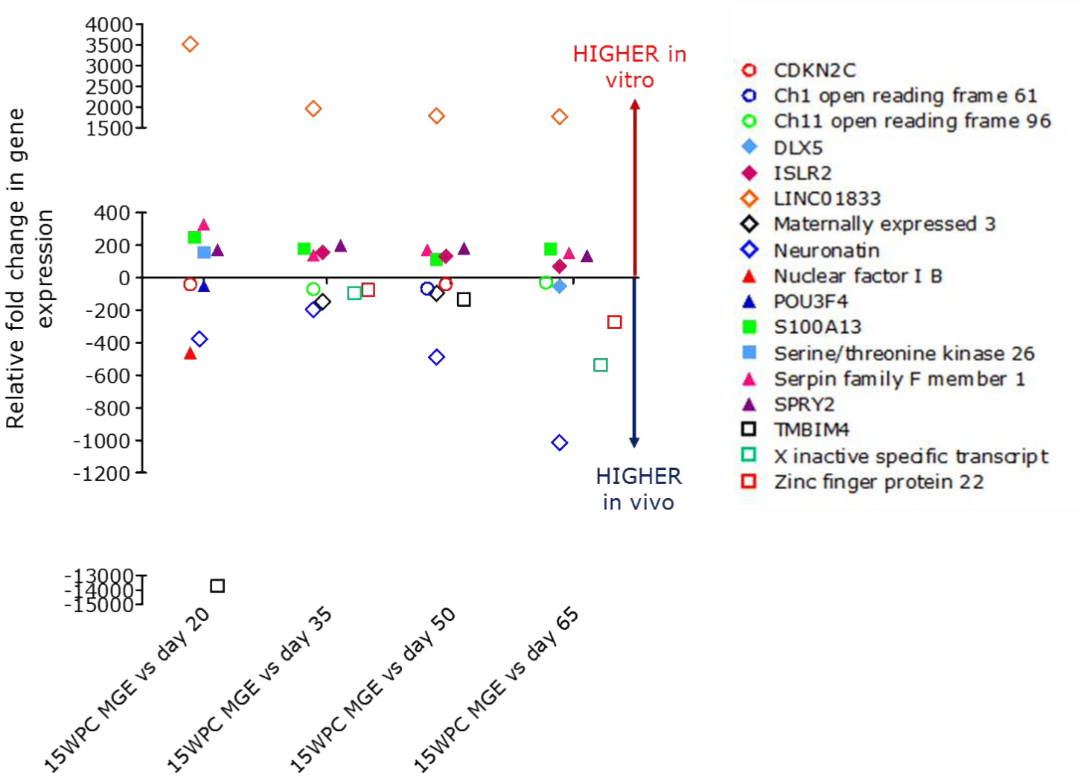


Figure 5.4 Cross-comparison to hESC-derived interneuron single-cell RNA sequencing data.

A) tSNE plot of the relative similarity in single cell profiles of hESC-derived interneurons and human foetal MGE – all time points and samples combined. Each symbol on tSNE plot represents a single cell and the unique gene expression profile of the cell. Cells from the same time point along differentiation paradigm/ foetal tissue sample share the same symbol (e.g. all day 20 cells are represented by red open circle). Clusters of symbols on tSNE plot represent cells with higher similarity in gene expression profile. Minimum biological replicates for MGE, n=1. Biological replicates for in vitro differentiations n=2. Total number of cells per time point/age can be found in Appendix folder 2 table 4 and 3 table 10.

B) tSNE plot of the relative similarity in single cell profiles of hESC-derived interneurons and human foetal MGE 10WPC. Cells from the same time point along differentiation paradigm or from 10WPC foetal tissue sample share the same symbol (e.g. all day 20 cells are represented by red open circle). Clusters of symbols on tSNE plot represent cells with higher similarity in gene expression profile. Biological replicates for 10WPC MGE, n=1. Biological replicates for in vitro differentiations n=2. Total number of cells per time point/age can be found in Appendix folder 2 table 4 and 3 table 10.

C) Relative fold change in expression of genes significantly differentially expressed between primary foetal MGE at 10WPC and hESC-derived interneurons. DEGs were determined through ANOVA pairwise analysis, and were included if $p<0.05$ and fold change (in either direction) >4. In the graphs only the top 5 DEGs in both directions (higher in vitro or in vivo) have been shown as examples (due to the large number of DEGs). The full list can be found in the Appendix folder 4. Genes with a relatively higher expression in hESC-derived interneuron sample (compared to 10WPC MGE) were shown as a positive fold change. Any genes with a relatively higher expression in 10WPC MGE (compared to hESC-derived interneuron sample) were shown as a negative fold change. Biological replicates for 10WPC MGE, n=1. Biological replicates for in vitro differentiations n=2.

D) tSNE plot of the relative similarity in single cell profiles of hESC derived interneurons and human foetal MGE 15WPC. Cells from the same time point along differentiation paradigm or from 15WPC foetal tissue sample share the same symbol (e.g. all day 20 cells are represented by red open circle). Clusters of symbols on tSNE plot represent cells with higher similarity in gene expression profile. Biological replicates for 15WPC MGE, n=1. Biological replicates for in vitro differentiations n=2. Total number of cells per time point/age can be found in Appendix folder 2 table 4 and 3 table 10.

E) Relative fold change in expression of genes significantly differentially expressed between primary foetal MGE at 15WPC and hESC-derived interneurons. DEGs were determined through ANOVA pairwise analysis, and were included if $p<0.05$ and fold change (in either direction) >4. In the graphs only the top 5 DEGs in both directions (higher in vitro or in vivo) have been shown as examples (due to the large number of DEGs). The full list can be found in the Appendix folder 4. Genes with a relatively higher expression in hESC-derived interneuron sample (compared to 15WPC MGE) were shown as a positive fold change. Any genes with a relatively higher expression in 15WPC MGE (compared to hESC-derived interneuron sample) were shown as a negative fold change. Biological replicates for 15WPC MGE, n=1. Biological replicates for in vitro differentiations n=2.

DEG, differentially expressed gene. MGE, medial ganglionic eminence. hESC, human embryonic stem cell. Calcyon neuron specific vesicular protein, CALY. CAMK2B, calcium/calmodulin dependent protein kinase II beta. CDKN2C, Cyclin dependent kinase inhibitor 2C. Chromosome (Ch) 1 open reading frame 61. Chromosome (Ch)

11 open reading frame 96. Developmental pluripotency associated 4, DPPA4. DLX5, Distal-less homeobox 5. HES5, Hes family bHLH transcription factor 5. Interferon induced transmembrane protein 2, IFITM2. ISLR2, Immunoglobulin superfamily containing leucine rich repeat 2. LINC01833, Long intergenic non-protein coding RNA 1833. Maternally expressed 3, MEG3. Neuronatin, NNAT. Nuclear factor I B, NFI B. Nuclear paraspeckle assembly transcript 1, NEAT1. Olfactomedin 2, OLFM2. PANTR1, POU3F3 adjacent non-coding transcript 1. POU3F4, POU class 3 homeobox 4. RBMS1, RNA binding motif single stranded interacting protein 1. S100A13, S100 calcium binding protein A13. S100B, S100 calcium binding protein B. Serine/threonine kinase 26, AKT. SPRY2, Sprouty RTK signaling antagonist 2. TMBIM4, transmembrane BAX inhibitor motif containing 4. Unnamed LncRNA, Chromosome (Ch) 8. X inactive specific transcript, XIST. WPC, weeks post-conception.

5.4 Discussion

To date, longitudinal single-cell RNA sequencing analysis of human MGE has not been reported in the literature. As a result, the data presented here has provided the first in-depth, unbiased analysis of human MGE molecular profile.

5.4.1 Human MGE developmental trajectory

Characterising the human MGE at single-cell resolution illustrated the heterogeneity within the population and across time. Over the late first trimester to early second trimester period (10WPC to 15WPC) the MGE samples shared a common mix of proliferating progenitors and immature neurons (mainly *NPY* expressing interneurons). These populations can be spatially divided into restricted zones within the developing MGE, namely the ventricular/subventricular zones and mantle zones respectively (Zechel *et al.* 2014). The proliferating progenitors were identified by the expression of early interneuron markers (*NKX2.1, FOXG1*), transcription related genes (*RRM2, CREB5*), and cell cycle regulatory genes (*CCND1*). The downregulation of these genes, and upregulation of post-mitotic neuronal markers (*MAP2*, *CR, SST, NPY*) over the developmental period reflected the shift towards a more neuronal rich population in the MGE. These findings at RNA level were supported by immunostaining performed on human MGE over the same period (10WPC to 14WPC), which reported a significant shrinkage of the MGE ventricular zone (reduced abundance of newborn progenitors) (Hansen *et al.* 2013). Therefore, together these findings would indicate the maturation of the MGE was a major factor contributing to the heterogeneity over the developmental period.

Furthermore, over time different subpopulations emerged within the MGE, which differed in the expression of signalling factors, including insulin-like growth factor signalling, neuroendocrine signalling protein secretogranin V, and activin and BMP4 signalling inhibitor follistatin like 1 (Cole *et al.* 2016). These signalling pathways may represent steps underpinning interneuron lineage commitment, in particular, differential levels of BMP4 signalling has previously been demonstrated to influence interneuron commitment to SST or PVALB fate (Mukhopadhyay *et al.* 2009). Therefore, the higher levels of BMP4 inhibitor, follistatin like 1, in MGE samples may be linked to the greater proportion of *SST* expressing interneurons (at the expense of *PVALB*) detected in the MGE compared to hESC-derived interneuronal culture (Mukhopadhyay *et al.* 2009). Therefore, exploring the effects of these named signalling pathways during the in vitro differentiation paradigm may enrich the hESC-derived interneuronal cultures with subtypes that we absent or at low abundance (*NPY* and *SST*, respectively).

However, due to the nature of longitudinal primary tissue analysis, and the small number of independent samples collected in this study (typically only one foetus per time point), some of the variations between different aged samples may be explained by between-subject variations. One example was the differential expression of *XIST*, a gene selectively expressed in females to regulate X-linked gene dosage (Sahakyan *et al.* 2018). As the foetal tissues were used in a gender unbiased manner, the upregulation of *XIST* over development was likely due to the change from male to female tissue samples rather than temporal changes in expression observed during MGE development (Zechel *et al.* 2014). Thus, gender-selective studies may be favourable in the future, especially as gene expression changes during brain development have been illustrated to be sex-dependent (Kang *et al.* 2011).

5.4.2 Interneuron lineage commitment in the MGE

Details concerning interneuron diversification into specific subtypes remain to be conclusively determined, with several competing theories in the literature (Hu *et al.* 2017b; Wamsley & Fishell 2017). However, the data described here would support the conclusion that lineage dictated subpopulations emerge within the MGE. This is in support to the most recent single-cell RNA sequencing analysis in mouse MGE and CGE that named *Maf and Mef2c* as candidate lineage determining transcription factors for SST and PVALB lineages respectively (Mayer *et al.* 2018; Mi *et al.* 2018). Although these findings were not reproduced in the data presented here, a novel gene associated with NPY lineage commitment was named: *SKP1*. The known activity of SKP1 in regulating the protein levels of RNA polymerases, transcription factors and cycle cell regulatory proteins through ubiquitination, means SKP1 effects many downstream pathways by initiating a cascade effect by regulating protein levels in the cell (Galan & Peter 1999). As a result, this makes *SKP1* a potential candidate for lineage commitment to NPY fate.

Although these findings favour an intrinsic preprogramming of interneuron commitment, initiated in the MGE, there are several caveats to this statement to consider. Firstly, although the data described here has detected lineage dictated subpopulations, this cannot rule out the possible further refinement of the interneuron profile in the region of destination (through extrinsic cues) (Kessaris *et al.* 2014; Wamsley & Fishell 2017). Secondly, only certain subtype committed subpopulations were detected (*NPY, SST, CR*) in the MGE. The absence of other MGE-derived interneuron lineages (e.g. PVALB) may be as a result of the developmental period selected or due to the emergence of this lineage only within the region of destination (Bandler *et al.* 2017; Wamsley & Fishell 2017). Therefore, depending on the interneuron subtype, a different interneuron lineage commitment model may better fit the true developmental steps (Hu *et al.* 2017b; Wamsley & Fishell 2017). In summary, this study provided favourable evidence for early interneuron lineage commitment, directed through intrinsic programming of cells in the MGE, but the importance of extrinsic cues in lineage commitment (especially for alternative lineages, like PVALB), and in determining the final phenotype of the interneuron, cannot be ruled out.

5.4.3 hPSC-derived interneuronal models – the good and the bad

Of final consideration was the cross-comparison of the single-cell profiles of hESC-derived interneurons and human MGE samples. This assessment of the in vitro model relative to the authentic counterpart in the foetal brain provided insight into how PSC-derived neurons mimic some, but not all, of the gene expression patterns observed in vivo. Many of the genes under-represented in the hESC-derived interneuron cultures were transcription factors, including *DLX5, MEG3* and *NFI B,* and regulators of the cell-cycle (cyclin-dependent kinase inhibitor 2C, *CDKN2C*). The low/absent expression of these genes in vitro may reflect the limited ability of in vitro models to recapitulate the generation of the distinct progenitor pools in the MGE, and the subsequent absence of these subtypes in the terminally differentiated culture. If the molecular profile of the hPSC-derived progenitors could be directed to express these under-represented genes, this could provide an efficient means to widen the variety of interneuron subtypes that can be generated in vitro. Furthermore, neuronatin was an additional gene consistently detected as under-expressed in the hESC-derived interneuron cultures. Neuronatin expression has previously been detected in the mouse MGE at RNA level, mainly restricted to lateral boundary of the MGE mantle zone, and therefore, is expressed by migratory interneurons (Zechel *et al.* 2014). The limited expression of neuronatin in the hESC-derived interneuron culture may be due to the absence of critical morphogens in vitro needed to mimic the spatial patterning in the lateral MGE. In summary, these discrepancies between PSC neuronal models and their primary tissue derivatives have been repeatedly reported across PSC- neuronal models, and point to common faults of PSC neuronal models that need to be addressed to achieve more authentic neurons in vitro (Camp *et al.* 2015; Chen *et al.* 2017; Close *et al.* 2017).

Nonetheless, many of the genes upregulated in vitro highlight the validity of PSC-derived interneurons as a neurodevelopmental model. These genes included neuron specific vesicular protein (*CALY*) and calcium signalling proteins (*S100B, S100A13, CAMK2B*), with calcium signalling known as an important modulator of synaptogenesis and maintenance of synaptic connections (Fink *et al.* 2003; Wayman *et al.* 2008). In particular, *CAMK2B* is a calmodulin-dependent kinase highly expressed during early development and regulates actin-filament polymerisation to promote synapse formation in a calcium-dependent manner (Fink *et al.* 2003; Wayman *et al.* 2008). The higher expression of these synaptic-related markers in vitro likely reflects the generation of neuronal networks in the culture which are absent in the MGE. This would indicate that the hPSC-derived interneuron model encompasses a wider range of interneuron development, beyond the stages in the MGE. This emphasises the importance of comparing PSC-interneuronal models to authentic counterparts in the region of destination (e.g. cortical tissue) to understand the later stages of interneuron development (dendritic branching and synaptogenesis), as impairments in these processes have been implicated in numerous neuropathologies (Brennand *et al.* 2011; Lima Caldeira *et al.* 2019; Shao *et al.* 2019). Although comparisons of hPSC-derived interneurons to cortical foetal interneurons were consistent with findings presented here, and also concluded a good similarity to early second trimester (14-15WPC) foetal tissue (Close *et al.* 2017).

5.4.4 Conclusion

In summary, human MGE single-cell transcriptome analysis has provided novel insight into MGE expression profile which has been utilised to critically evaluate the hPSC model of interneuron differentiation. Encouragingly, the hPSC model, in many aspects, reproduced the findings reported in the human MGE. However, the differences in molecular profile detected emphasised the need for future refinements to the differentiation paradigm, which can be guided from the novel cross-comparison analysis presented here. Of significance were pathways regulating gene expression (directly or indirectly) that were under-represented in vitro relative to the native MGE, and therefore if intrinsically or extrinsically elevated during the interneuron differentiation paradigm, may aid the generation of more authentic and defined interneurons in vitro.

Chapter 6 - Physiology and connectivity of hESC derived interneurons

6.1 Introduction

As illustrated in the previous results Chapters, the expression profiles of hESC-derived interneurons at single-cell resolution was heterogeneous and displayed temporal dynamics (see Chapters 4-5). The transcriptome is the foundation underpinning the phenotype of the interneuron. Thus, interneuron diversity goes beyond molecular profile and is also reflected in the morphology, connectivity and physiology of the neuron (Kepecs & Fishell 2014).

6.1.1 The connectivity of neurons

As introduced in Chapter 1, monosynaptic rabies viral tracers have become the prime tool used to identify and manipulate neuronal circuits of interest (Wickersham *et al.* 2007; Osakada *et al.* 2011; Osakada & Callaway 2013). The system is based on three main components. First, EnvA pseudotyped ΔG rabies tracer, which is a non-self replicating rabies virus that can only enter neuronal population expressing the complementary cell surface receptor TVA. This neuronal population is termed starter neurons. Second, the viral tracer induces the expression of reporters and all the viral genes, except the viral glycoprotein, in the infected host neuron (Wickersham *et al.* 2007; Osakada *et al.* 2011; Osakada & Callaway 2013). The ability to choose the reporters has made viral tracers a highly versatile tracing platform, which can both label traced circuits using a fluorescent protein (e.g. mCherry) and manipulate the traced circuit by using neuronal modifying proteins such as the light-gated cation channels (Nagel *et al.* 2003; Osakada *et al.* 2011). Channelrhodopsin 2 (ChR2) is common example, used to stimulate neuronal activity, due to the optogenetic manipulation being of high temporal precision, reliability and sensitivity (Nagel *et al.* 2003; Boyden *et al.* 2005; Osakada *et al.* 2011).

Third, the starter neurons are genetically engineered to express the rabies glycoprotein, allowing replication of the viral tracer through glycoprotein trans-complementation (Osakada & Callaway 2013). Once synaptically transferred to a non-starter neuron, the loss of the exogenous source of the viral glycoprotein prevents viral amplification necessary for efficient synaptic transmission, thereby terminating the tracing signal to this traced neuron, which is directly one synapse from the starter neuron (see Figure 1.5). Hence, the EnvA pseudotyped ΔG rabies tracer provides information concerning the direct (monosynaptic) inputs onto a selected neuronal population (the starter neurons) (Brennand *et al.* 2011; Osakada *et al.* 2011; Wall *et al.* 2013; Tian *et al.* 2016; Tornero *et al.* 2017).

In conclusion, this new generation of tracers has allowed the combined determination of neuronal connectivity and manipulation of neuronal physiology of selected circuits (Tian *et al.* 2016; Tornero *et al.* 2017).

6.1.2 Neuronal physiology

Although the principle requirement of all neurons is the ability to code information in the form of spike patterns, neuronal physiology shows high neuronal class specificity. Interneurons are a classic example, with subtypes categorised according to differences in spike patterns and intrinsic physiology, such as fast spiking PVALB interneurons, delayed regular spiking bipolar CR/VIP interneurons, and burst spiking SST interneurons (Butt *et al.* 2005; Flames & Marín 2005, Kepecs & Fishell 2014).

Depending on the research questions, there are many different approaches to measure the physiology of neurons. Indirect measures of firing activity can be achieved using calcium imaging (Sun *et al.* 2016; Xiang *et al.* 2017; Mayer *et al.* 2019). Although calcium imaging has the advantage of providing a high-throughput neuronal response to a stimulus at single-cell resolution, the indirectness of the measure could potentially lead to inaccurate conclusions, as fluctuations in intracellular Ca^{2+} ion levels can have numerous causes (Bagur & Hajnóczky 2017). Population level responses can also be gained from multi-electrode array (MEA) recordings, which detect extracellular changes in potential (termed extracellular field potentials), caused by the firing of neurons in that region of the array (Spira & Hai 2013). Both calcium imaging and MEA have been particularly useful in the understanding of network activity, and how neuronal dysfunction can impact the ability of neuronal populations to achieve rhythmic and synchronised patterns of firing (Hedrich *et al.* 2014; Sun *et al.* 2019). However, neither calcium imaging or MEA recordings allow determination of the intrinsic physiology or the evoked spike parameters of a single neuron.

To functionally characterise neurons at a single-cell resolution an alternative approach must be taken, namely whole-cell patch clamp (Varga *et al.* 2015). Whole-cell patch clamping requires direct contact to be made to the neuron using a micro-pipette, by breaking into the cell the intracellular solution and pipette solution become one. This allows the detection of intracellular changes in ion concentrations, and subsequently membrane potential, that occurs on the generation of an action potential (Varga *et al.* 2015; Fuzik *et al.* 2015).

6.1.3 Physiology of primary tissue derived interneurons

Longitudinal whole-cell patch clamp recordings at different stages along neuronal differentiation have shown how the intrinsic and evoked parameters of the neuron change as the integrated circuits are moulded in an activity-dependent manner during development (Moore *et al.* 2009; Calvigioni *et al.* 2016; Pan *et al.* 2016; Zhong *et al.* 2018). Longitudinal functional characterisation of interneurons derived from mouse primary tissue has included SST interneurons in the anterior cingulate cortex (P3 to P45) and CCK interneurons in the hippocampus (E12.5 to P16) (Calvigioni *et al.* 2016; Pan *et al.* 2016). The age-dependent changes in the intrinsic physiology of neurons included decreased resting membrane potential (more hyperpolarised) and input resistance, and increased capacitance with age (Calvigioni *et al.* 2016; Pan *et al.* 2016). Moreover, action potential properties were also found to change from the first postnatal week, with a significant reduction in action potential duration and consequentially, rise in maximal firing frequency (Calvigioni *et al.* 2016; Pan *et al.* 2016). Ultimately the functional differentiation trajectory of interneurons led to the diverse set of physiology profiles seen within the adult brain (Varga *et al.* 2015).

6.1.4 Physiology of hPSC-derived neurons

Whole-cell patch clamp recordings have also been performed on PSC-derived interneurons (Cambray *et al.* 2012; Maroof *et al.* 2013; Sun *et al.* 2016; Close *et al.* 2017; Xiang *et al.* 2017; Noakes *et al.* 2019). In vitro differentiated PSC-derived interneurons reportedly fail to show the intrinsic physiological parameters observed in their authentic counterparts in the adult brain, and lack the full spectrum of firing patterns (often no fast-spiking or bursts spiking neurons) (Sun *et al.* 2016; Close *et al.* 2017; Noakes *et al.* 2019). However, there are inconsistencies across paradigms concerning the level of functional maturity achieved in vitro (Cambray *et al.* 2012; Maroof *et al.* 2013; Sun *et al.* 2016; Close *et al.* 2017; Xiang *et al.* 2017; Noakes *et al.* 2019). Nevertheless, acquiring neuron-like cells with native molecular and physiological maturation properties remains one of the major challenges and limitations of in vitro neuron differentiation paradigms to date. This has encouraged several different approaches to obtain more functionally mature neuronal populations (Nicholas *et al.* 2013; Colasante *et al.* 2015; Kemp *et al.* 2016; Sun *et al.* 2016; Telezhkin *et al.* 2016).

6.1.5 Improving the functional maturity of hPSC-derived neurons

One strategy has been the development of in vitro neuronal differentiation paradigms that apply a cocktail of cell cycle inhibitors (DAPT and PD0332991), neurogenic/neurotrophic factors (ascorbic acid, BDNF and LM22A4) and other factors known to directly influence neuronal physiology (GABA, calcium chloride, CHIR99021 and forskolin) to PSC-derived neuronal precursors (Kemp *et al.* 2016; Telezhkin *et al.* 2016). These protocols have arguably provided the simplest strategy to enhance both the speed and synchronicity of maturation of the PSC-derived neuronal precursors to post-mitotic neurons in vitro (Kemp *et al.* 2016; Telezhkin *et al.* 2016). The use of SCM1 factors described by Telezhkin *et al.* has been utilised in the experiments described in this Chapter in an attempt to improve the functional maturation of the hESC-derived interneuronal culture within the 65 day differentiation paradigm (Telezhkin *et al.* 2016; Telezhkin *et al.* 2018).

6.2 Aims

To characterise the differentiation trajectory of hESC-derived interneurons at the functional level by performing longitudinal whole-cell patch clamp recording.

To compare the longitudinal functional and molecular characterisation of hESC-derived interneurons (introduced in Chapter 4), to investigate whether any changes in physiology were mirrored by changes in expression profile.

To provide preliminary evidence for the successful application (and non-toxicity) of the EnvA pseudotyped ΔG ChR2-mCherry rabies viral tracer on hESC-derived interneurons, using immunostaining and optogenetics.

6.3 Results

Refinements to the hPSC interneuron differentiation protocol were introduced in this Chapter to enhance functional maturation of the interneuronal culture in vitro. As shown in Figure 6.1A, SCM1 factors were used at the beginning of the neuronal maturation stage (approx. day 21 to day 27 in vitro). As a result, post-mitotic neuronal-like branching was observed earlier, typically by the end of the SCM1 treatment, and was maintained with little to no contamination from proliferating progenitors over the remaining 40 days in vitro (Figure 6.1A).

6.3.1 The hESC-derived interneurons had a heterogeneous, but overall immature physiology

Whole-cell patch clamp recordings were conducted at days 40, 50 and 60 on hESC-derived interneurons to track changes in neuronal physiology over time (Figure 6.1B-6.1F). Moreover, the recordings from single-cells enabled the variation in the population to be characterised, as shown by the distribution of points on the graph (each point represents the results from a single patch neuron) (Figure 6.1B and 6.1D). The population average resting membrane potential was weakly hyperpolarised at all time points, but did show a slight (and significant, ***$P<0.001$) hyperpolarisation over the course of the in vitro differentiation paradigm to a population average resting membrane potential of -20mV (*$p<0.05$, ***$p< 0.001$, Figure 6.1B). In general, the differentiation trajectory of the hESC-derived interneurons was marked by the progressive increase in the membrane time constant (tau) over time, with a significant increase by day 60 to 90ms (***$p< 0.001$, Figure 6.1B). The capacitance for the population was overall very low at each time point (average between 20-30pF), with patched neurons rarely found to have a capacitance higher than 50pF (Figure 6.1B). Nonetheless, the population average capacitance significantly increased over time (*$p< 0.05$, Figure 6.1B).

The evoked responses of the same patched neurons were then investigated through a series of different current clamp paradigms (see Appendix folder 6 figures). These paradigms involved the injection of current in the patched neuron and simultaneous measure of changes in the membrane potential. From the patched neurons that successfully spiked on current stimulation, the following properties of the evoked action potential were determined: spike threshold, spike amplitude and spike half-width (Figure 6.1C -6.1D). An illustrative example defining these action potential parameters has been provided in Figure 6.1C. Broadly speaking, the average evoked action potential parameters for responsive hESC-derived interneurons did not significantly change over the day 40 to day 60 period (Figure 6.1D). Although there was a large variation in spike parameters within the population at each time point, the average remained relatively constant over time (Figure 6.1D). The population average spike amplitude showed a small, but significant reduction between day 40 and day 50, but otherwise also remained relatively unchanged over the day 40 to day 60 in vitro differentiation period (77.63mV to 48.11mV *p<0.05, Figure 6.1D). The cause of this reduced spike amplitude at day 50 was unclear, but possibly the small sample number (n=12) at day 40 may have led to a biased representation of the day 40 population response (Figure 6.1D).

Although the spike parameters may not have changed over time, how did the proportion of passive to responsive neurons change? The evoked responses to a single injection of current were assigned to one of four categories: one, an attempted, but failed action potential; two, a single action potential; three, an attempt at generating more than one action potential (a train); four, an action potential train (Figure 6.1E). The proportion of patched neurons failing to show an action potential response remained at approx. 30% of the population at each of the time points (Figure 6.1F). Similarly, only a small proportion of patched neurons achieved a train of evoked action potentials, with only a small increase in this proportion over time (peak at day 50 at 9%, Figure 6.1E). Whereas, there appeared to be a reciprocal increase in patched neurons that attempted a train of action potentials (6% to 17% day 40 to day 60) at the expense of those generating a single action potential (61% to 48% day 40 to day 60) over time (Figure 6.1F).

In summary, these findings would suggest that the interneuron differentiation of hESC recapitulated some of the functional changes expected in terminally differentiating neurons. However, both the intrinsic and evoked properties of hESC-derived interneurons would indicate that the functional maturity reached was limited, and did not confer to that known of their counterparts in the adult brain. Moreover, the high between-cell variation in innate and evoked physiological parameters corresponded to the heterogeneity within the culture. This may in part reflect the mixed functional maturity of the hESC-derived interneurons at each time point, with both passive and more functionally mature hESC-derived neurons maintained over time (Figure 6.1F). In addition, this may signify differences between discrete interneuron-like subtypes in the culture.

Figure 6.1A

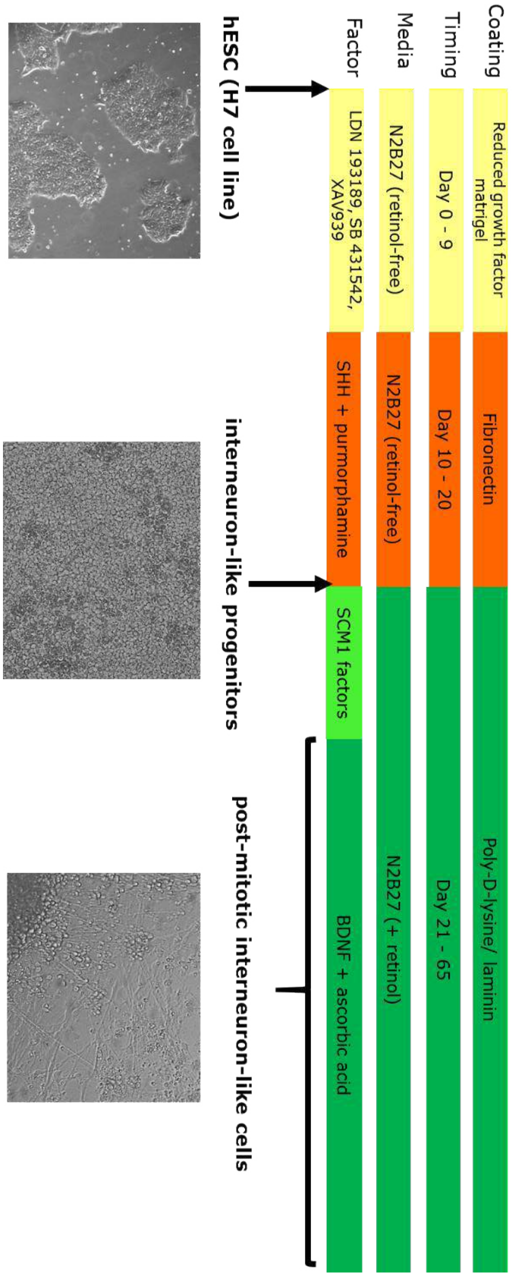

Coating
Reduced growth factor matrigel
Fibronectin
Poly-D-lysine/ laminin
Timing
Day 0 - 9
Day 10 - 20
Day 21 - 65
Media
N2B27 (retinol-free)
N2B27 (retinol-free)
N2B27 (+ retinol)
Factor
LDN 193189, SB 431542, XAV939
SHH + purmorphamine
SCM1 factors
BDNF + ascorbic acid
hESC (H7 cell line)
interneuron-like progenitors
post-mitotic interneuron-like cells

Figure 6.1B

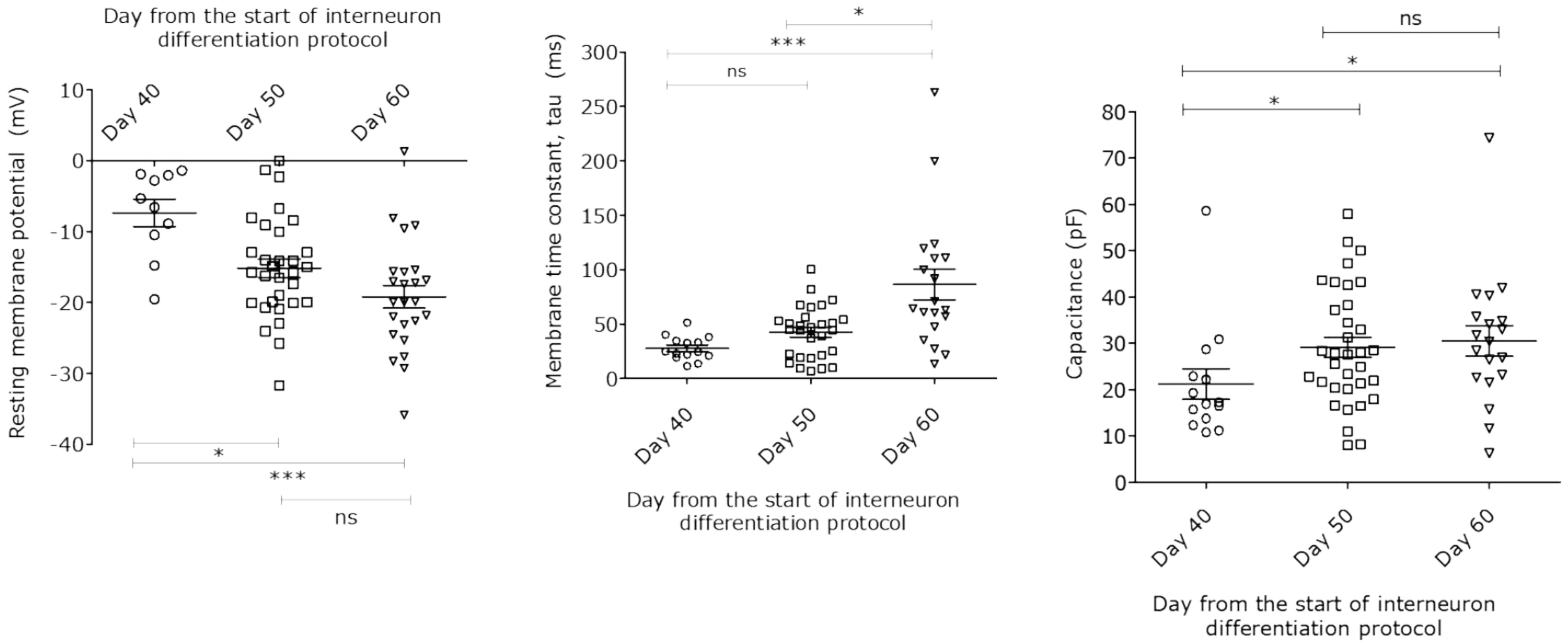

Day from the start of interneuron differentiation protocol
Resting membrane potential (mV)
10
0
-10
-20
-30
-40
Day 40
Day 50
Day 60
*
***
ns
Membrane time constant, tau (ms)
300
250
200
150
100
50
0
*
***
ns
Day 40
Day 50
Day 60
Day from the start of interneuron differentiation protocol
Capacitance (pF)
80
70
60
50
40
30
20
10
0
ns
*
*
Day 40
Day 50
Day 60
Day from the start of interneuron differentiation protocol

Figure 6.1C

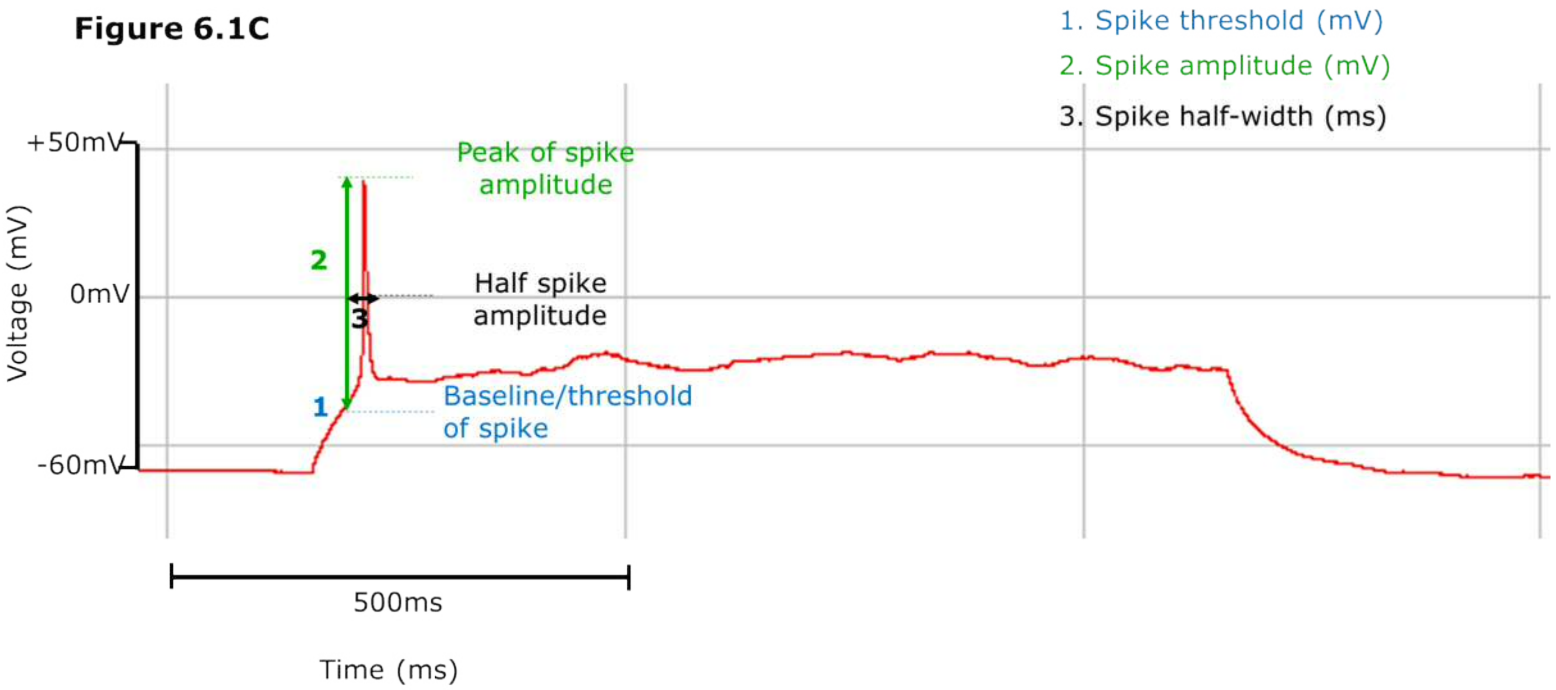

1. Spike threshold (mV)
2. Spike amplitude (mV)
3. Spike half-width (ms)
+50mV
0mV
-60mV
Voltage (mV)
Peak of spike amplitude
2
Half spike amplitude
3
1
Baseline/threshold of spike
500ms
Time (ms)

Figure 6.1D

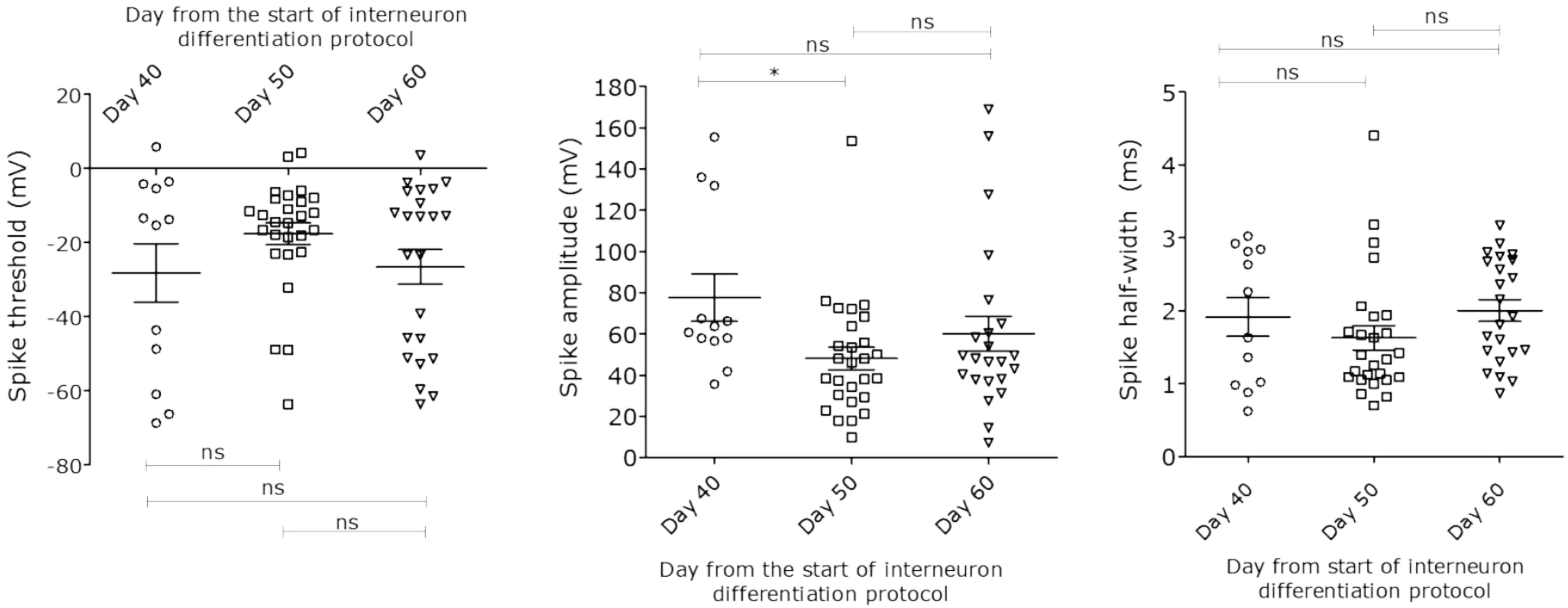
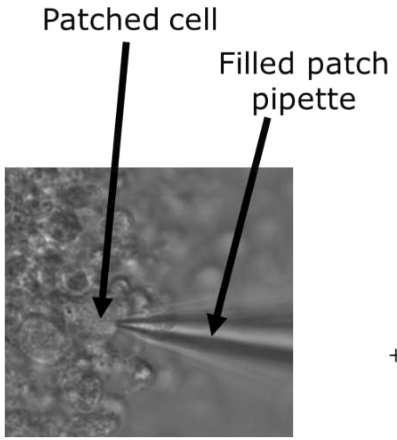
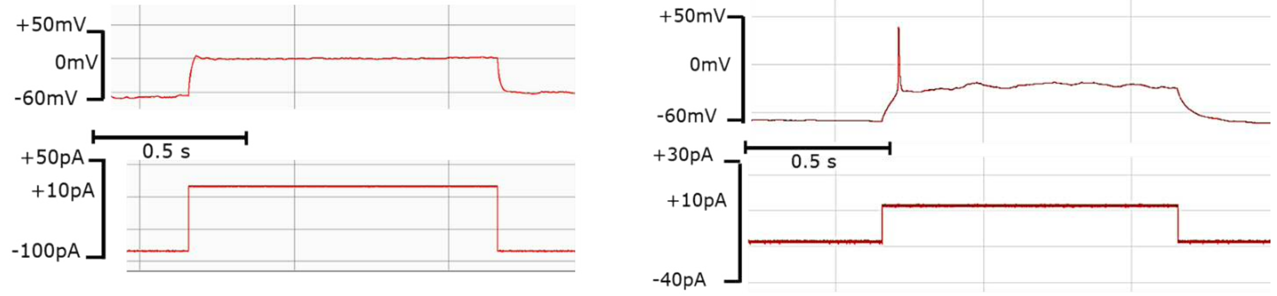
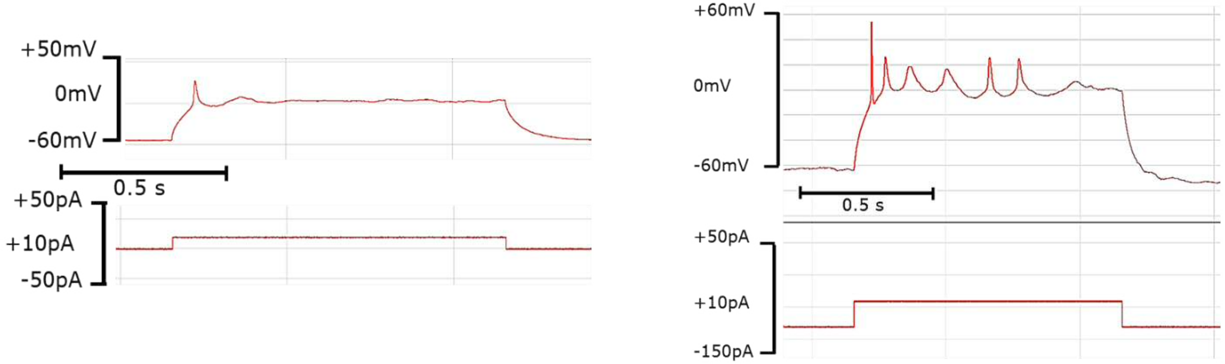

Day from the start of interneuron differentiation protocol
Spike threshold (mV)
20
0
-20
-40
-60
-80
Day 40
Day 50
Day 60
ns
ns
ns
Spike amplitude (mV)
180
160
140
120
100
80
60
40
20
0
*
ns
ns
Day 40
Day 50
Day 60
Day from the start of interneuron differentiation protocol
Spike half-width (ms)
5
4
3
2
1
0
ns
ns
ns
Day 40
Day 50
Day 60
Day from start of interneuron differentiation protocol

Figure 6.1E


Patched cell
Filled patch pipette


1. Attempted action potential

2. Single action potential


3. Attempted train of action potentials

4. Train of action potentials

Figure 6.1F

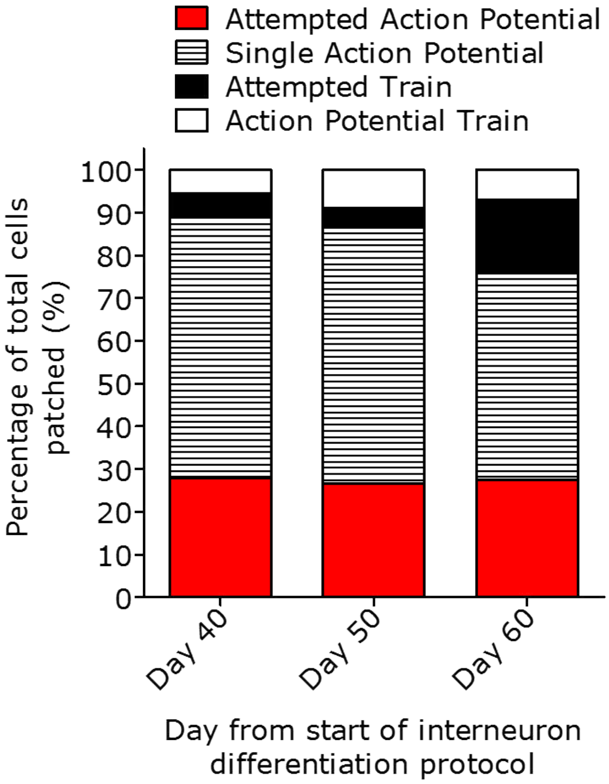


Figure 6.1 Physiological parameters of hESC-derived interneurons.

A) Schematic of the in vitro hPSC interneuron differentiation protocol utilised in this study, with real-time brightfield images of in vitro culture at different time points. The paradigms has been simply divided into three different stages. Stage 1, neural induction (yellow). Stage 2, interneuron fate commitment (orange). Stage 3, terminal differentiation with additional SCM1 factors treatment phase (green). The SCM1 factors consisted of: neurotropic factors, BDNF, ascorbic acid and LM22A4, cell cycle inhibitors, DAPT and PD0332991 and factors known to directly shape neuronal physiology, GABA, calcium chloride, CHIR99021 and forskolin. For more information please refer to Chapter 2.

B) Intrinsic physiology of patched neurons, including the immediate resting potential of the neuron (taken within first 30sec of breaking into neuron), membrane time constant (tau) and capacitance (capacitance (pF) = tau (ms) / input resistance (GΩ)). The graph has been plotted with each symbol on the plot representing the value of a single patched neuron, the mean +/- SEM have been shown for each time point. Biological replicates (independent differentiations) n=3. Minimum technical replicates per time point= 10

C) Representative example of an evoked action potential labelled to define the following action potential parameters: spike threshold (mV), spike amplitude (mV) and spike half-width (ms)

D) Current evoked action potential properties of patched neurons, including spike threshold (mV), spike amplitude (mV) and spike half-width (ms). The graph has been plotted with each symbol on the plot representing the value of a single patched neuron, the mean +/- SEM have been shown for each time point. Holding potential, -60mV. Biological replicates (independent differentiations) n=3. Minimum technical replicates per time point= 10

E) Representative examples of the four different categories of responses to current injection. Category 1. Attempted, but failure to generate an action potential (passive response). Category 2. Single action potential generated. Category 3. Attempt to generate a second action potential (train of action potential), but failed. Category 4. Two or more action potentials generated in response to a single current injection. Holding potential, -60mV.

F) The proportion of patched cells at each time point within each of the four categories of evoked responses (described in 6.1E). The graph shows the percentage of patched cells with the corresponding response (attempted spike, single spike, attempted train, train of action potentials) out of total cells patched at that time point. minimum n= 12 at each time point. Holding potential, -60mV. Biological replicates (independent differentiations) n=3. Minimum technical replicates per time point= 10

Statistics performed to determine significant differences in physiological responses of patched cells over time using nonparametric, Kruskal-Wallis test and Dunn's Multiple Comparison Test (*p<0.05, **p<0.01, ***p<0.001). SHH, sonic hedgehog. BDNF, brain-derived neurotrophic factor. hESC, human embryonic stem cell. SEM, standard error of the mean.

6.3.2 The differentiation trajectory of hESC-derived interneurons was reflected in the expression of markers associated with neuronal physiology

To determine whether the differences in physiology were accompanied by changes in expression of neuronal physiology-related genes (such as ion channel, ion transporters, neurotransmitter receptors), the longitudinal single-cell RNA sequencing data for hESC-derived interneurons (introduced in Chapter 4) was utilised to track changes in selected genes (Figure 6.2A-6.2B).

Analysis of the longitudinal single-cell RNA sequencing data for hESC-derived interneurons began by restricting the analysis to the post-mitotic neuronal population at each time point. The *MAP2* expressing subpopulation was checked for cells expressing genes encoding for ion transporters (ATPase Na+/K+ transporting subunits alpha 1 *ATP1A1* and beta 3 *ATP1B3*), vesicular GABA transporter (*vGAT*), signalling receptors (inositol 1,4,5-trisphosphate receptor type 1, *ITPR1*), and cytoskeletal proteins and cytoskeletal regulatory proteins (tropomyosin 4, *TPM4* and cyclase associated actin cytoskeleton regulatory protein 2, *CAP2* respectively) (Figure 6.2A). The proportion of the neuronal population expressing *ATP1B3* and *TPM4* remained high (approx. >80%) from day 20 to day 65, showing little temporal dynamics (Figure 6.2A). Although the average expression level for *ATP1B3* transcripts were much higher than *TPM4* at each time point (by at least 2-fold, day 50 270 TPM vs 500 TPM) (Figure 6.2A). The proportion of *CAP2* and *ATP1A1* expressing neurons over time showed a mirrored expression trend, with both markers showing a progressive decrease over time and a slight increase at day 65 in abundance (Figure 6.2A). The high co-expression pattern of *CAP2* and *ATP1A1* in the population (>75%) may indicate a shared gene regulatory network, driving a common phenotypic trait. The proportion of *ITPR1* and *vGAT* expressing neurons remained low across the differentiation paradigm, with the proportion of *vGAT* expressing cells peaking at day 35 and *ITPR1* at day 50 (at 12% and 25%, respectively Figure 6.2A). The population average expression level of *ITPR1* and *vGAT* were also both at low levels (<50 TPM per cell average) (Figure 6.2A).

Furthermore, the proportion of neurons expressing genes encoding for several different ion channel subunits including selective Na^+ ion (*SCN1A, SCN2A*), K^+ ion (*KCNQ2, KCNK1*), Ca^{2+} ion (*CACNA1C*) and Cl^- ion (*CLCN3*) voltage-gated channels, was determined (Figure 6.2B). As shown in Figure 6.2B, the temporal expression pattern of the ion channel subunits varied, both in terms of the direction and size of change over time. For example, the abundance of neurons expressing *SCN1A, SCN2A* and *CACNA1C*, peaked at a later time point, although only the proportion of *SCN1A* expressing neurons progressively increased over day 20 to day 65 (approx. 2% to 40%, day 20 vs day 65 Figure 6.2B). Whereas, *KCNQ2* and *CLCN3* were expressed by the greatest proportion of neurons at day 20, and progressively decreased over time (Figure 6.2B). At day 20 only *PVALB* and *CR* expressing interneuron subtypes were detected, with *PVALB* expressing neurons also declining in abundance over time (see Chapter 4). This may suggest a subtype biased expression pattern of *KCNQ2* and *CLCN3*. Finally, K^+ two pore domain channel *KCNK1* expressing neurons were of low abundance at all time points of the differentiation paradigm (remained at approx. 5%) (Figure 6.2B). Although, the proportion of marker expressing cells may have changed over time, the average number of transcripts per cell remained relatively independent of time along the differentiation (Figure 6.2B). Although some markers (*SCN1A, SCN2A* and *CACNA1C*) did show a small rise in population average transcript levels in later populations (typically from day 35 onwards) (Figure 6.2B).

The changes described in single-cell profiles may, in part, reflect the differentiation trajectory of the post-mitotic interneuron-like cells, mirroring the temporal changes reported in the whole-cell patch clamp recordings. In addition, the differences in ion channel expression patterns between temporally distinct cultures could signify changes in subtype diversity of interneuron-like population over time.

6.3.3 hESC-derived interneurons showed subtype-dependent expression pattern of markers associated with neuronal physiology

To investigate whether the synaptic proteins and ion channel expression patterns differ between interneuron subtypes, the *MAP2* expressing population was instead divided according to *SST, CR* or *PVALB* co-expression (rather by time point) (Figure 6.2C). The proportion of *SST, CR* and *PVALB* expressing neurons co-expressing different voltage-gated Na^+ ion channel subunits belonging to Nav1 family was determined, and demonstrated a small subtype bias in the expression of *SCN1A* (SST bias), *SCN8A* (SST bias) and *SCN2A* (CR bias) (Figure 6.2C). This coincides with the temporal expression data for *SCN1A* and *SCN2A* in Figure 6.2B, both of which showed greater expression from day 35 onwards, with this time point corresponding to the emergence and greater abundance of *SST* and *CR* expressing interneurons (with the decline in *PVALB*) (see Chapter 4).

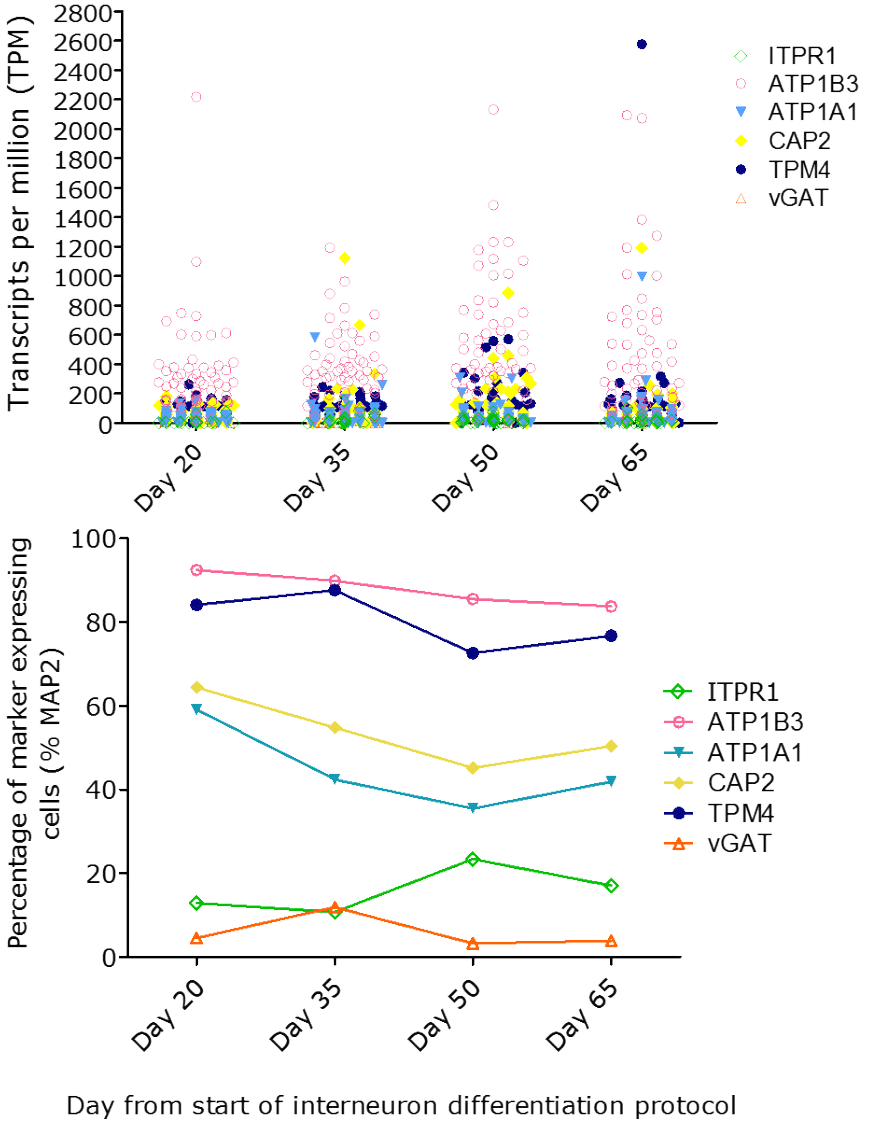
Overall, the potassium voltage-gated ion channels for Kv3 family (*KCNC1, KCNC2)* were at a low (<20%) abundance or absent in *SST, CR* and *PVALB* expressing neuronal subpopulations (Figure 6.2C). As Kv3.1 (*KCNC1)* and Kv3.2 (*KCNC2)* are pan-neuronal markers, the low/absent expression would imply the neurons were functionally immature. Whereas, there was greater variability in the expression pattern of the Kv7 family of ion channels *(KCNQ2, KCNQ3, KCNQ4, KCNQ5*) (Figure 6.2C). Kv7.4 (*KCNQ4)* expressing neurons were not co-expressed by any of the *SST, CR* and *PVALB* expressing neuronal subpopulations (Figure 6.2C). Similarly, Kv7.5 *(KCNQ5)* expressing neurons were either absent or at very low abundance (<5%) (Figure 6.2C). Whereas, Kv7.2 (*KCNQ2)* and Kv7.3 (*KCNQ3)* expressing cells were detected in each subpopulation, with *KCNQ2* being the most abundant K^+ ion channel subunit for all subtypes (Figure 6.2C). Moreover, *KCNQ2* showed greatest co-expression with *PVALB* (approx. 75% of *PVALB* positive post-mitotic neurons Figure 6.2C). This reflected the temporal expression data for *KCNQ2* described above in Figure 6.2B, and *PVALB* described in Chapter 4, both of which had peak abundance in the *MAP2* positive population at day 20.

In a similar manner reported for the Kv7 family of ion channels, the proportion of neurons expressing voltage-gated L-type Cav1 family of ion channel subunits (*CACNA1C, CACNA1D, CACNA1F, CACNA1S*) showed high variability between subunits. For example, in hESC-derived post-mitotic neurons *CACNA1S* and *CACNA1F* were either absent or at low proportion within *SST, PVALB, CR* subpopulations (both at <5%, Figure 6.2C). Whereas, *CACNA1C* and *CACNA1D* showed similar co-expression with each subtype (Figure 6.2C).

Finally, the proportion of subtype marker positive cells expressing glutamate ionotropic receptor subunits (NMDA or AMPA receptors) was determined. NMDA receptor subunit *GRIN2B* and AMPA receptor subunit *GRIA3* were the most highly expressed by all subtypes (45-50% and 25-35% respectively, Figure 6.2C). Although, this expression pattern was subtype dependent, with a slight *PVALB* biased co-expression pattern with *GRIA3* (Figure 6.2C). Moreover, *GRIN2B* displayed a clear co-expression bias with *SST* and *CR,* but no co-expression reported with *PVALB* (Figure 6.2C).

Overall, hESC-derived *PVALB* expressing neurons co-expressed the smallest number and diversity of the ion channels, ion transporters and neurotransmitter receptor subunits checked (Figure 6.2C). Furthermore, often the proportion of *PVALB* /marker co-expressing neurons was lower relative to *CR* and *SST* expressing interneurons (Figure 6.2C). This would suggest the hESC-derived *PVALB* expressing interneurons were less functionally mature than *CR/SST* interneurons, likely due to the emergence of *PVALB* expressing interneurons at an earlier time point in vitro and loss by later time points (see Chapter 4).

Figure 6.2A



Transcripts per million (TPM)
2800
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
600
400
200
0
Day 20
Day 35
Day 50
Day 65
ITPR1
ATP1B3
ATP1A1
CAP2
TPM4
vGAT
Percentage of marker expressing cells (% MAP2)
100
80
60
40
20
0
Day 20
Day 35
Day 50
Day 65
ITPR1
ATP1B3
ATP1A1
CAP2
TPM4
vGAT
Day from start of interneuron differentiation protocol

Figure 6.2B

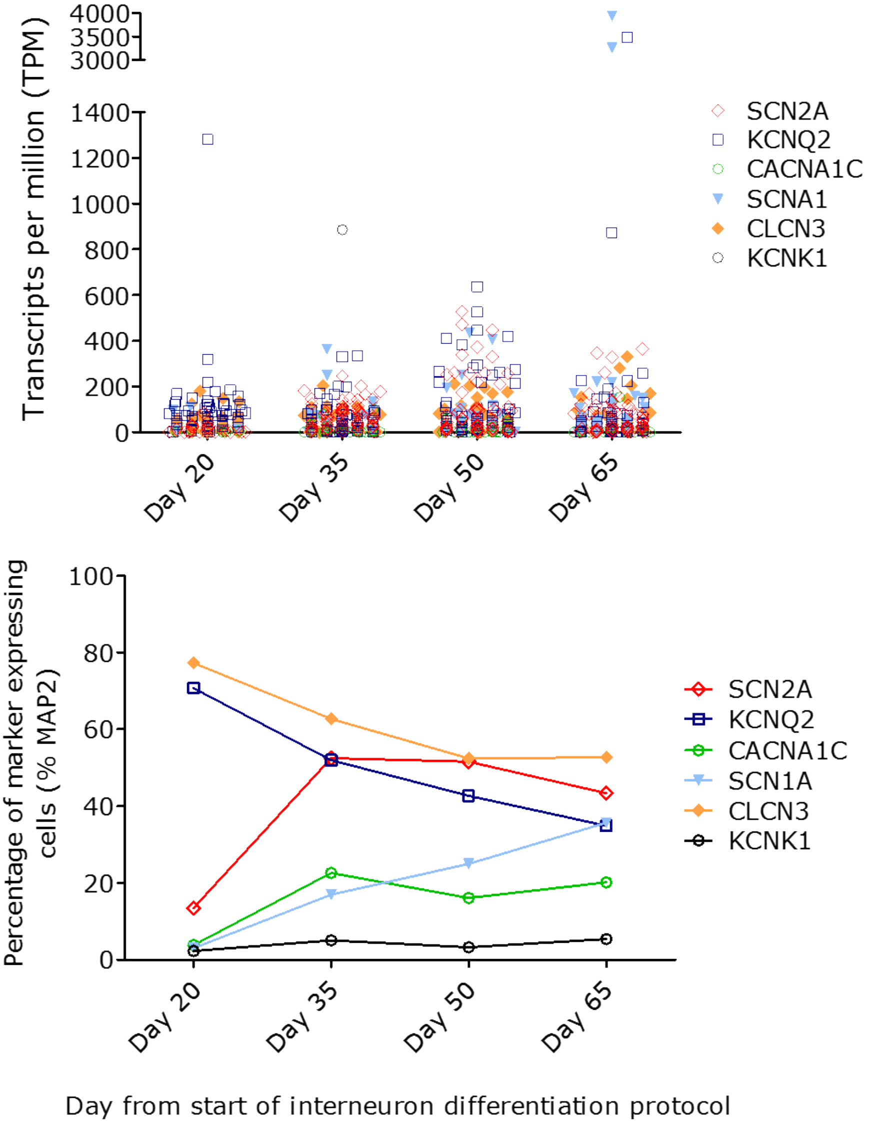

Transcripts per million (TPM)
4000
3500
3000
1400
1200
1000
800
600
400
200
0
Day 20
Day 35
Day 50
Day 65
SCN2A
KCNQ2
CACNA1C
SCNA1
CLCN3
KCNK1
Percentage of marker expressing cells (% MAP2)
100
80
60
40
20
0
Day 20
Day 35
Day 50
Day 65
SCN2A
KCNQ2
CACNA1C
SCN1A
CLCN3
KCNK1
Day from start of interneuron differentiation protocol

Figure 6.2C

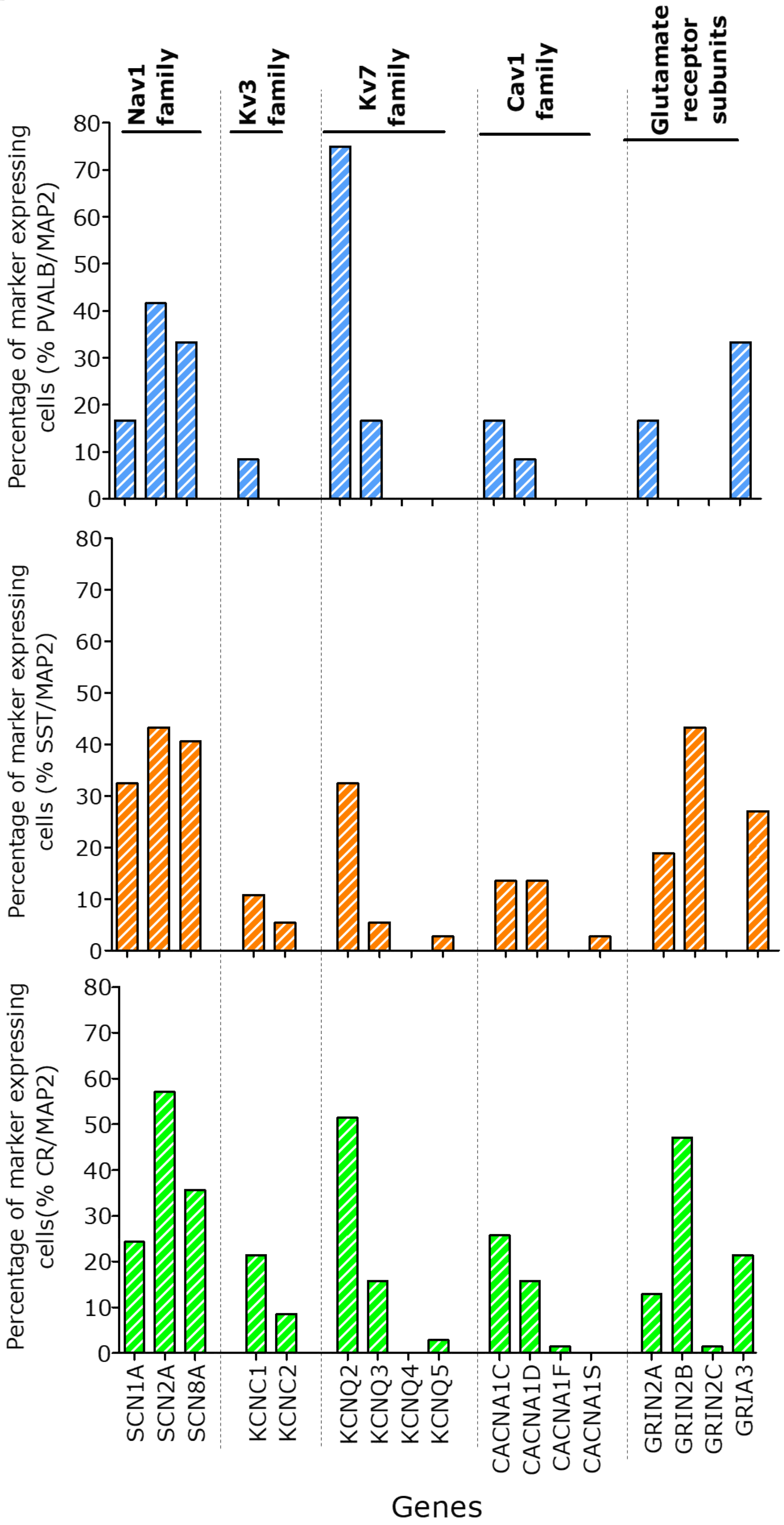

Nav1 family
Kv3 family
Kv7 family
Cav1 family
Glutamate receptor subunits
Percentage of marker expressing cells (% PVALB/MAP2)
Percentage of marker expressing cells (% SST/MAP2)
Percentage of marker expressing cells(% CR/MAP2)
0
10
20
30
40
50
60
70
80
SCN1A
SCN2A
SCN8A
KCNC1
KCNC2
KCNQ2
KCNQ3
KCNQ4
KCNQ5
CACNA1C
CACNA1D
CACNA1F
CACNA1S
GRIN2A
GRIN2B
GRIN2C
GRIA3
Genes

Figure 6.2 Longitudinal trajectories of molecular markers associated with neuronal physiology, obtained from single-cell RNA sequencing of hESC-derived interneurons.

A-B) Quantification of the expression level and proportion of marker positive cells for selected genes using single-cell RNA sequencing data (see Chapter 4). 6.2A) Markers indirectly linked to neuronal physiology, ITPR1, ATP1B3, ATP1A1, CAP2, TPM4 and vGAT. 6.2B) Ion channels, SCN2A, KCNQ2, CACNA1C, SCN1A, CLCN3 and KCNK1. Gene expression level data has been presented as TPM, with each symbol plotted representing the expression value of that marker for a single cell. The line graphs plotted show the proportion of cells expressing each named marker at each time point along the hESC-derived interneuron differentiation protocol. The marker positive cells have been counted if the cells also co-expressing MAP2, to restrict analysis to post-mitotic neurons. Total MAP2 expressing cells at day 20=132 cells, day 35=177, day 50=124 cells and day 65 = 129 cells. Biological replicates (independent differentiations) n=2.

C) Expression pattern of Na+ ion channels (SCN2A, SCN1A, SCN8A), K+ ion channels (Kv3 KCNC1, KCNC2, and Kv7 KCNQ2, KCNQ3, KCNQ4, KCNQ5) and Ca2+ ion channels (CACNA1S, CACNA1C, CACNA1D, CACNA1F), and ionotropic glutamate receptor subunits (GRIA3, GRIN2C, GRIN2A, GRIA3) by SST or PVALB or CR positive interneurons (all cells included in analysis were MAP2 positive). The graphs plotted show the proportion of cells expressing each named marker for each interneuron subtype (CR, SST, PVALB). Total MAP2 expressing cells co-expressing SST = 35 cells, CR= 67 cells and PVALB= 12 cells. Biological replicates (independent differentiations) n=2.

TPM, transcripts per million. vGAT, vesicular GABA transporter. ATP1A1, ATPase Na+/K+ transporting subunit alpha 1. ATP1B3, ATPase Na+/K+ transporting subunit beta 3. CACNA1C, voltage-gated L-Type calcium channel Cav1.2 alpha 1 subunit. CACNA1, voltage-gated L-Type calcium channel Cav1.3 alpha 1 subunit. CACNA1F, voltage-gated L-Type calcium channel Cav1.4 Alpha 1 Subunit. CACNA1S, voltage-gated L-Type calcium channel Cav1.1 alpha 1 subunit. CAP2, cyclase associated actin cytoskeleton regulatory protein 2. CLCN3, chloride voltage-gated channel 3. CR, calretinin. GRIA3, glutamate ionotropic receptor AMPA type subunit 3. GRIN2A, glutamate ionotropic receptor NMDA type subunit 2A. GRIN2B, glutamate ionotropic receptor NMDA type subunit 2B. GRIN2C, glutamate ionotropic receptor NMDA type subunit 2C. ITPR1, inositol 1,4,5-trisphosphate receptor type 1. KCNC1, potassium voltage-gated channel subfamily C member 1. KCNC2, potassium voltage-gated channel subfamily C member 2. KCNK1, potassium two pore domain channel subfamily K member 1. KCNQ2, voltage-gated potassium channel subunit Kv7.2. KCNQ3, voltage-gated potassium channel subunit Kv7.3. KCNQ4, voltage-gated potassium channel subunit Kv7.4. KCNQ5, voltage-gated potassium channel subunit Kv7.5. MAP2, microtubule associated protein 2. PVALB, parvalbumin . SCN1A, voltage-gated sodium channel subunit alpha Nav1.1. SCN2A, voltage-gated sodium channel subunit alpha Nav1.2. SCN8A, voltage-gated sodium channel subunit alpha Nav1.8. SST, somatostatin . TPM4, tropomyosin 4.

6.3.4 Lentiviral transduction efficiently generated hESC-derived starter neuron population

Through the longitudinal molecular and functional characterisation described above, the hPSC model of interneuron differentiation would appear to give rise to functionally immature interneuron-like cells. However, whether the hESC-derived neurons underwent the stage of synaptogenesis to form viable neuronal circuits in vitro, is unknown. The pseudotyped EnvA ΔG ChR2-mCherry viral tracer system was used to selectively identify and manipulate the circuits containing starter and traced neurons. The generation of starter neuron population required lentiviral transduction to induce the inheritable expression of TVA, rabies glycoprotein and histone-GFP under the human synapsin promoter (Figure 6.3A and Appendix folder 5 figures). The nuclear GFP reporter allowed titering of the lentivirus (number of transducing units (TU)/ml) through the proportion of GFP positive neurons in the culture.

Lentiviral transduction was carried out as either a single or double administration of the lentivirus, and a range of different lentiviral dilutions were trialled to allow titering (Figure 6.3B-6.3C). After 7-10 days post lentiviral transduction the hESC-derived interneuron cultures were checked by independent approaches for GFP positive cells. First, by eye under a fluorescent microscope to gauge an estimate of GFP expression in live cells. Second, by flow cytometry to determine the proportion of GFP positive cells (see Appendix folder 5 figures). Both these approaches demonstrated a variation in GFP fluorescent intensity between GFP positive starter neurons. The consequence of which for the viral tracer system is unclear. Third, immunohistochemistry staining of the hESC-derived cultures for GFP. No GFP positive cells were detected in the negative control (virus-free media exchange on transduction days) in any of the detection methods (Figure 6.3B).

Lentiviral treatment successfully generated GFP positive cells, the proportion of which was dependent on the dilution factor used (Figure 6.3B and Appendix folder 5 figures). A progressive decrease in GFP positive population occurred with the increase in dilution factor (more dilute lentiviral stock used), which became significant by 1:750 dilution factor (single hit 1:250, 45% to 30% $^{**}p< 0.01$, double hit 1:250, 75% to 30% $^{***}p< 0.001$ Figure 6.3B-6.3C).

The repeated transduction with the lentiviral vector resulted in a significantly greater proportion of GFP positive cell relative to the equivalent single hit treatment for 1:250 (single hit 45% vs double hit 75% ***$p < 0.001$) and 1:500 dilution factors (single hit 40% vs double hit 60% **$p < 0.01$) (Figure 6.3C). In summary, these approaches led to a calculated lentiviral titer of 4.5×10^6 TU/ml, and indicated that a double hit was not necessary to efficiently generate a 30-40% abundant starter cell population (Figure 6.3B-6.3C).

Figure 6.3A

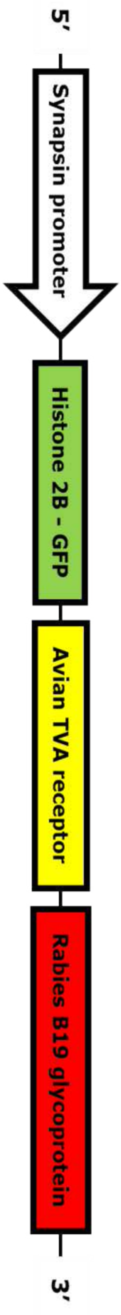

5'
Synapsin promoter
Histone 2B - GFP
Avian TVA receptor
Rabies B19 glycoprotein
3'

Figure 6.3B

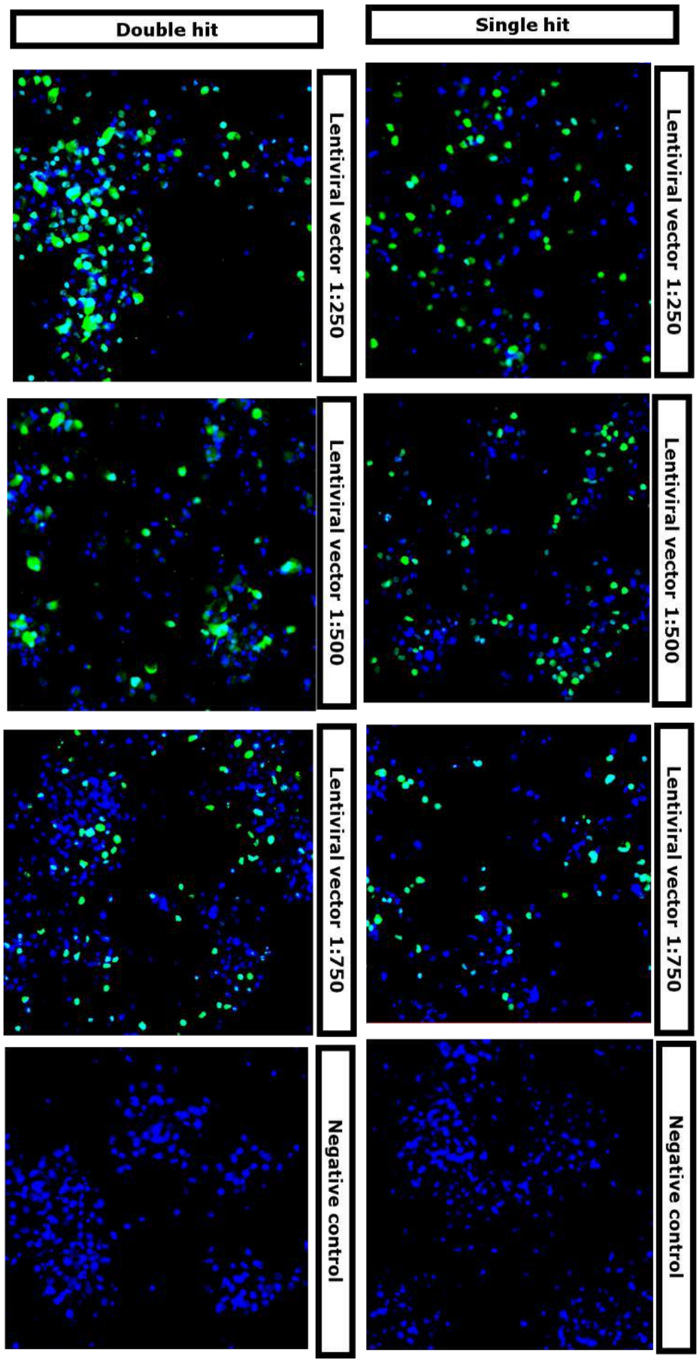

Double hit
Single hit
Lentiviral vector 1:250
Lentiviral vector 1:250
Lentiviral vector 1:500
Lentiviral vector 1:500
Lentiviral vector 1:750
Lentiviral vector 1:750
Negative control
Negative control

Figure 6.3C

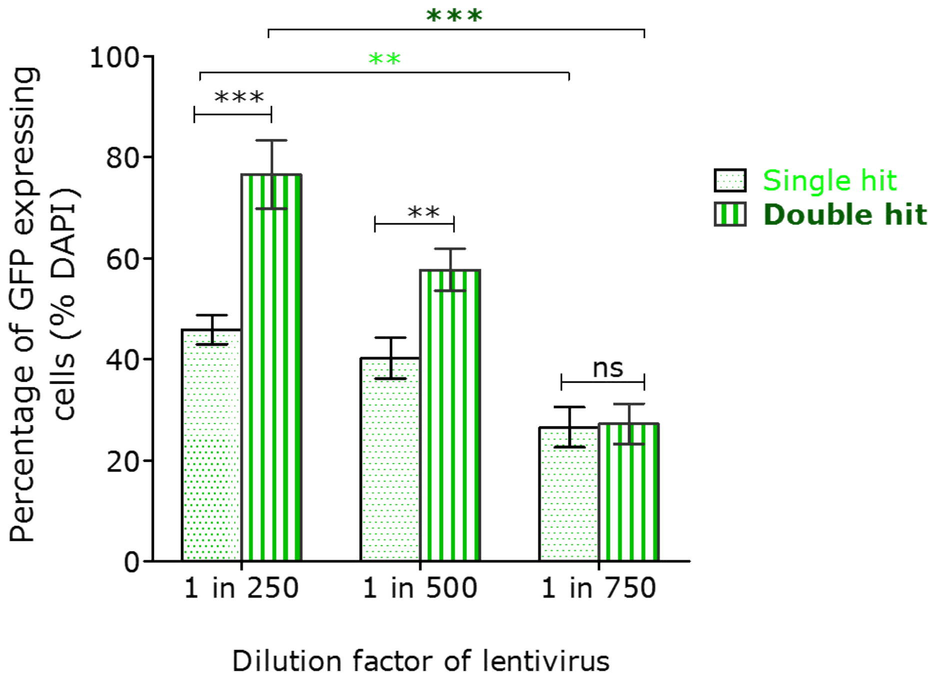


Figure 6.3 hESC-derived starter neurons: lentiviral transduction and titering.

A) Linearised schematic of lentiviral vector. Human synapsin promoter-driven expression of the avian TVA receptor, the B19 rabies glycoprotein and histone 2B-green fluorescent protein (H2B.GFP) fusion protein. The three proteins are kept as three separate entities by 2A self-cleaving peptide linker sequence joining each of the proteins.

B) Transduction of hESC-derived interneurons with lentiviral vector to generate starter neuron population for viral tracer system. The data has been presented in the form of a panel of immunohistochemistry deconvoluted images, with different lentiviral transduction treatments. As described in detail in Chapter 2, hESC-derived interneurons were first transfected at approx. day 35 with either lentiviral stock diluted to 1:250, 1:500 and 1:750. A proportion of these cultures were then transduced a second time (double hit) 7 days later with the same dilution factor as used previously. At day 50 all cultures were fixed for immunohistochemistry staining. For both single and double hit treatments, a negative control was used (virus-free media added). Biological replicates (independent differentiations) n=3. Minimum technical replicates n=6.

C) Percentage of GFP expressing cells out of the total cell population (total cell count measured using DAPI count) for each of the lentiviral transduction treatments. The graph plotted shows the mean percentage of GFP expressing cells for each treatment, along with +/- SEM. Biological replicates (independent differentiations) n=3. Minimum technical replicates n=6. The average total number of cells counted per technical replicate can be found in Appendix folder 7.

Statistics performed to determine significant differences in percentage of GFP expressing cells, within each treatment condition (single hit or double hit) relative to the highest dilution factor (1:250 either single or double hit) using nonparametric, Kruskal-Wallis test and Dunn's Multiple Comparison Test (*p<0.05, **p<0.01,

***p<0.001, light green text for single hit, dark green text for double hit). Mann-Whitney test was performed to determine significant differences in percentage of GFP expressing cells between corresponding single and double hit treatments, (e.g. 1:250 single hit vs 1:250 double hit) with ns p>0.05, *p<0.05, **p<0.01, ***p<0.001 (black text). GFP, green fluorescent protein. H2B, histone-2B. hESC, human embryonic stem cell. SEM, standard error of the mean.

6.3.5 EnvA pseudotyped ΔG ChR2-mCherry viral tracer showed restricted uptake to TVA expressing cells

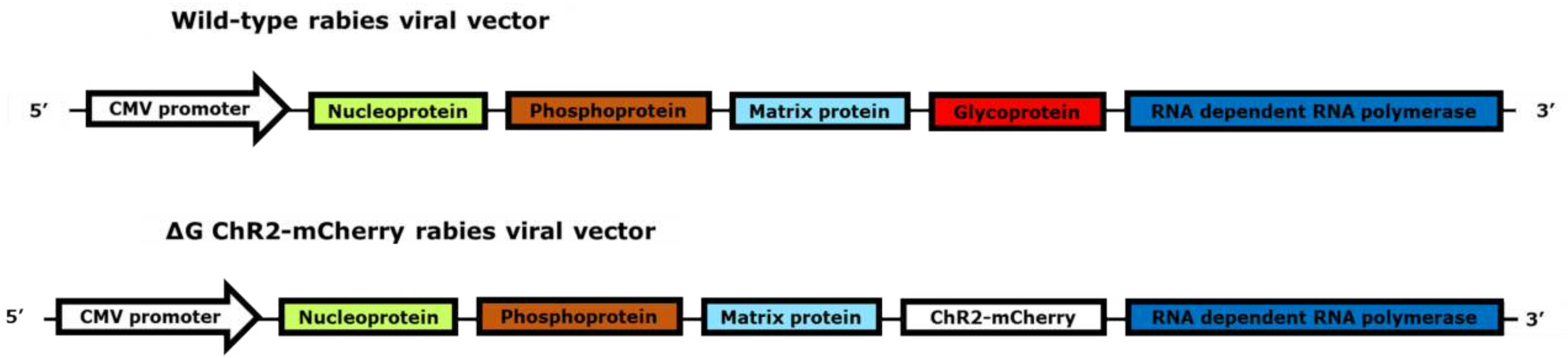
The starter cells had enforced expression of TVA to match the complementary envelope protein (EnvA) displayed on the surface of the EnvA pseudotyped ΔG ChR2-mCherry rabies tracer. The generation of the EnvA pseudotyped ΔG ChR2-mCherry rabies virus required the ΔG ChR2-mCherry rabies virus to be taken up and amplified by BHK-EnvA packaging cells (see Chapter 2). However, as the unpseudotyped ΔG ChR2-mCherry rabies virus was the precursor to the pseudotyped viral tracer, there could be contamination from the unpseudotyped viral tracer in the viral harvest collected. As the unpseudotyped virus can enter all cells (TVA expressing or not) it was necessary to determine whether there was any contaminating unpseudotyped viral tracer in the pseudotyped viral tracer stock. This was achieved by determining the transduction efficiency of the unpseudotyped and EnvA pseudotyped ΔG ChR2-mCherry rabies virus in HEK cell line with and without TVA receptor expression (Figure 6.4A-6.4C).

Investigations began with determining the proportion of mCherry positive HEK TVA expressing cells after 7-10 days post-infection with the pseudotyped and unpseudotyped viral tracer, over a range of dilution factors (Figure 6.4B-6.4C). Immunohistochemistry staining of HEK TVA cultures for mCherry confirmed the efficient transduction of both viruses in this cell line, which was also seen by eye (under the fluorescent microscope) and using flow cytometry (Figure 6.4B, and Appendix folder 5 figures). The proportion of mCherry positive HEK TVA cells significantly decreased with progressively higher dilution factors for the EnvA pseudotyped ΔG ChR2-mCherry rabies virus (35% to 15% 1:100 vs 1:250 *$p<0.05$, 35% to 10% 1:100 vs 1:750 **$p<0.01$, Figure 6.4C). Although a similar trend was seen for the unpseudotyped viral tracer, the proportion of mCherry positive cells was higher for 1:500 dilution factor (15% vs 30% **$p<0.01$), and only showed a significant decrease relative to 1:100 dilution only for 1:750 dilution (35% to 10%, 1:100 vs 1:750 ***$p<0.001$, Figure 6.4C). The higher proportion of mCherry positive HEK TVA cells for the unpseudotyped ΔG ChR2-mCherry rabies viral treatments meant that the titer for unpseudotyped viral tracer

(3.2 x10^6 TU/ml) was higher relative to the EnvA pseudotyped viral tracer (2.8 x10^6 TU/ml) (as would be expected, see Chapter 2).

Finally, to determine whether the EnvA pseudotyped or unpseudotyped ΔG ChR2-mCherry viral tracer uptake was selective to TVA positive cells, the viral tracer treatments were repeated in HEK cells. A similar trend in proportion of mCherry positive cells across dilution factors was reported for unpseudotyped rabies viral tracer transductions of HEK cells (as seen in HEK TVA cell line). This result was predicted, as unpseudotyped viral tracer has no restrictions on cell entry. Therefore the titer calculated was similar across the two cell lines (3.2 - 3.6 x10^6 TU/ml) (Figure 6.4B-6.4E). In contrast, for every dilution factor of EnvA pseudotyped ΔG ChR2-mCherry viral tracer treatment (from 1:100 to 1:750) no mCherry positive cells were detected by eye, flow cytometry or immunohistochemistry staining (Figure 6.4D-6.4E, and Appendix folder 5 figures). This would indicate that pseudotyping of the viral tracer was successful, with no contaminating unpseudotyped viral tracer present, but efficient transduction required a high dosage (1:100, Figure 6.4C).

Figure 6.4A



Wild-type rabies viral vector

ΔG ChR2-mCherry rabies viral vector

5′
CMV promoter
Nucleoprotein
Phosphoprotein
Matrix protein
Glycoprotein
RNA dependent RNA polymerase
3′
5′
CMV promoter
Nucleoprotein
Phosphoprotein
Matrix protein
ChR2-mCherry
RNA dependent RNA polymerase
3′

Figure 6.4B

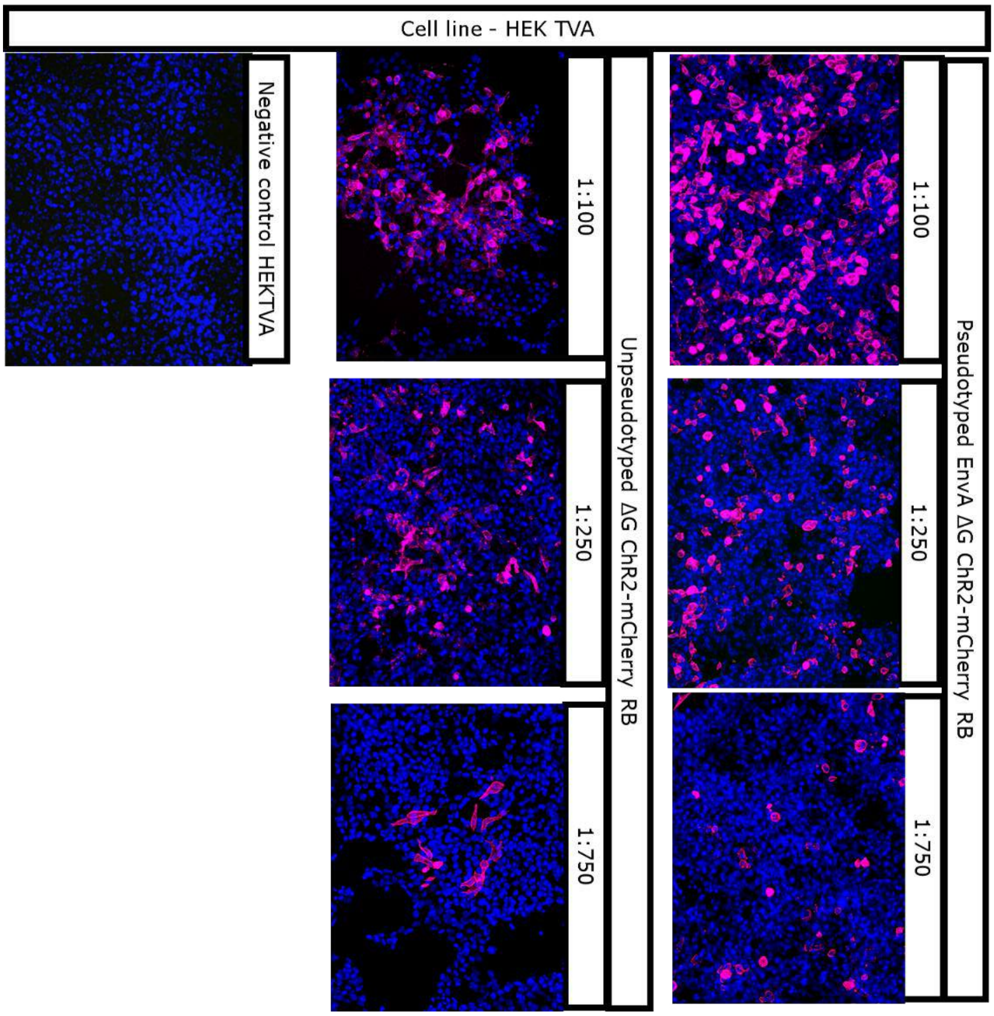

Cell line - HEK TVA
Negative control HEKTVA
Unpseudotyped ΔG ChR2-mCherry RB
1:100
1:250
1:750
Pseudotyped EnvA ΔG ChR2-mCherry RB
1:100
1:250
1:750

Figure 6.4C

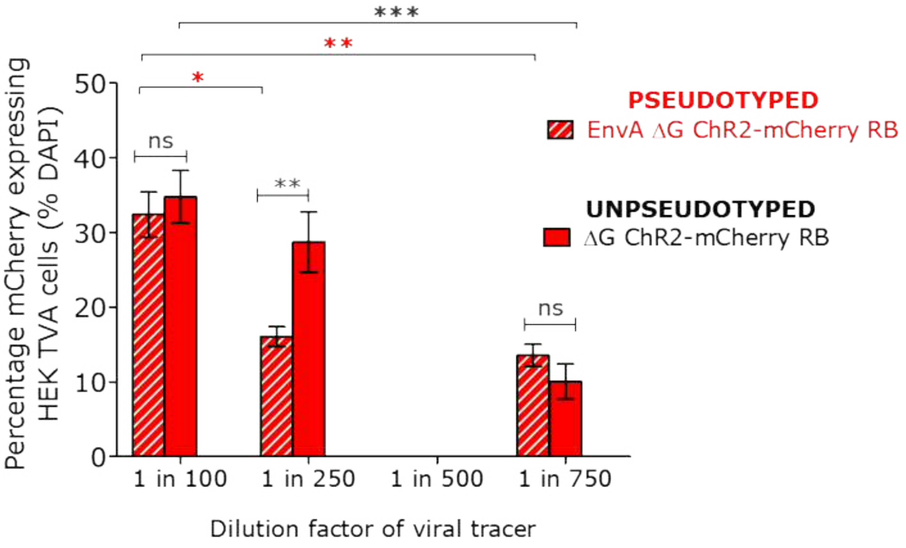


Figure 6.4D

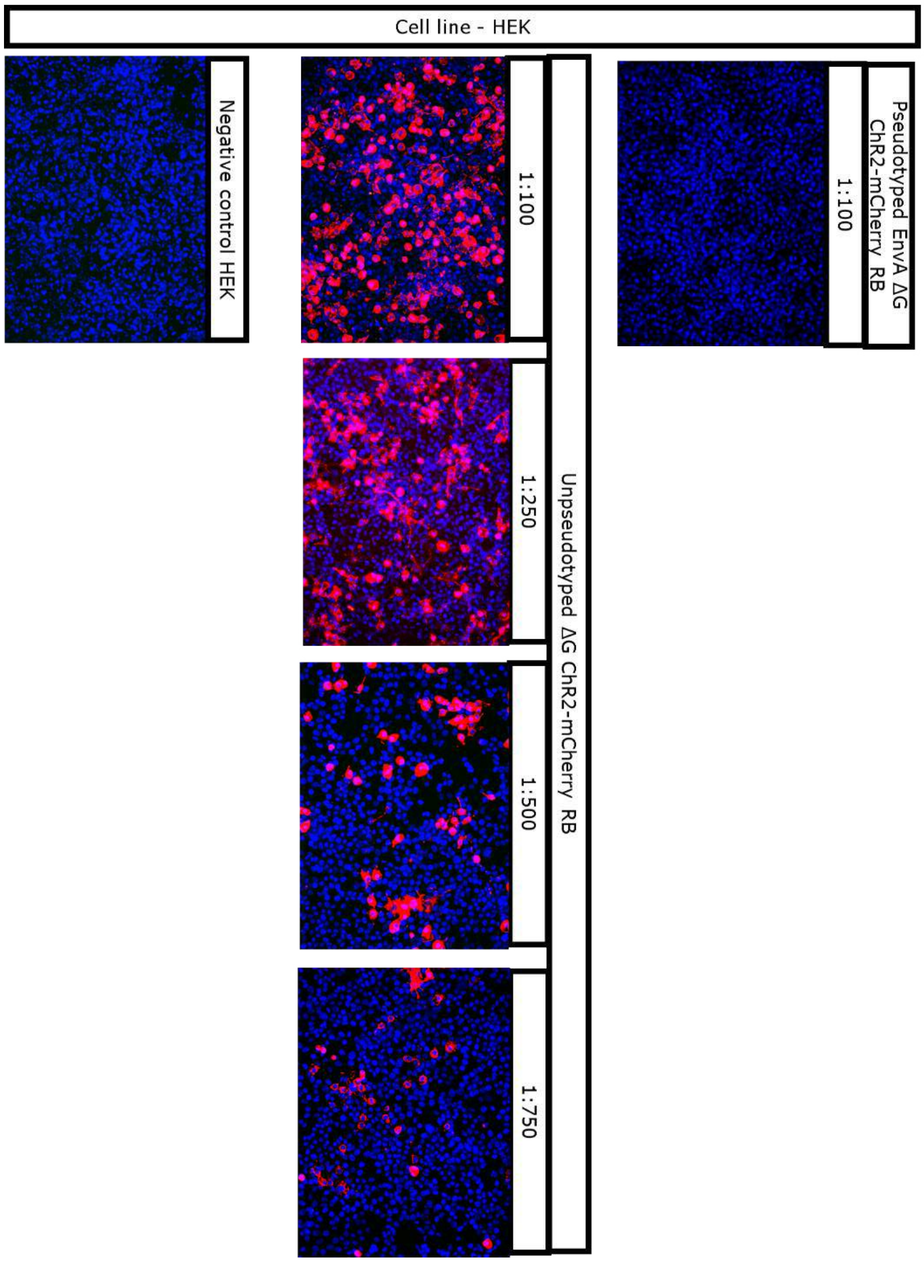

Cell line - HEK
Negative control HEK
1:100
1:250
1:500
1:750
Unpseudotyped ΔG ChR2-mCherry RB
1:100
Pseudotyped EnvA ΔG ChR2-mCherry RB

Figure 6.4E

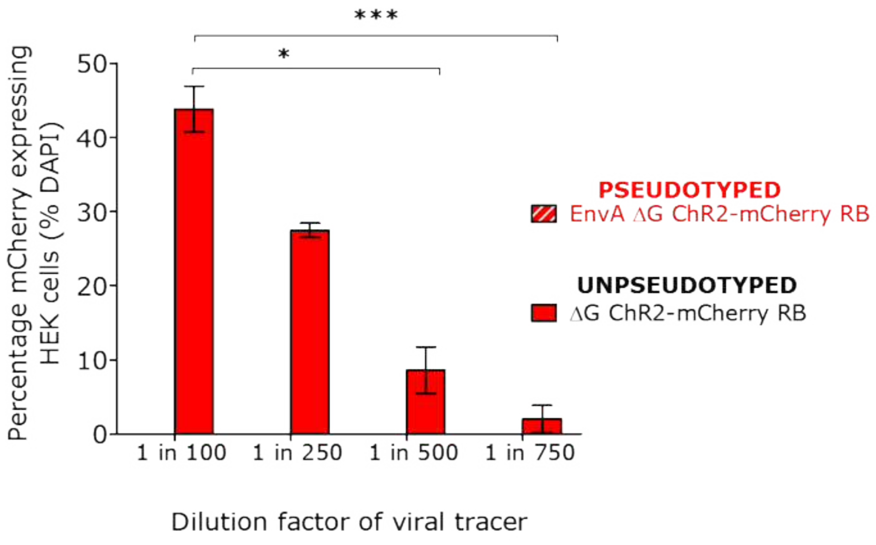


Figure 6.4 The EnvA pseudotyped ΔG ChR2-mCherry rabies viral tracer: viral transduction and titering.

A) Linearised schematic of the rabies viral vector, the first showing the five genes encoded by the wild-type RB and the second showing the modified ΔG rabies tracer with glycoprotein gene sequence replaced with sequence encoding for channelrhodopsin 2 (ChR2) and fluorescent protein mCherry fusion protein (ChR2-mCherry).

B) Transduction of HEK TVA cells with unpseudotyped and EnvA pseudotyped ΔG ChR2-mCherry RB. The data has been presented in the form of a panel of immunohistochemistry deconvoluted images, with different pseudotyped/unpseudotyped ΔG ChR2-mCherry RB transduction treatments labelled. As described in detail in Chapter 2, HEK TVA cells were either transduced with stock diluted in 1:100, 1:250, 1:500 and 1:750 unpseudotyped or EnvA pseudotyped ΔG ChR2-mCherry RB. After 7-14 days all cultures were fixed for immunohistochemistry staining. All experiment had a corresponding negative control (virus-free media added). Biological replicates (independent differentiations) n=3. Minimum technical replicates n=6.

C) Percentage of mCherry expressing HEK TVA cells out of the total cell population (total cell count measured using DAPI count) for each of the unpseudotyped and EnvA pseudotyped ΔG ChR2-mCherry RB transduction treatments. The graph plotted shows the mean percentage of mCherry expressing cells for each treatment, along with +/- SEM. Note 1:500 treatment is missing due to detachment of cells on fixing. Biological replicates (independent differentiations) n=3. Minimum technical replicates n=6. The average total number of cells counted per technical replicate can be found in Appendix folder 7.

D) Transduction of HEK cells with unpseudotyped and EnvA pseudotyped ΔG ChR2-mCherry RB. The data has been presented in the form of a panel of immunohistochemistry deconvoluted images, with different ΔG ChR2-mCherry RB transduction treatments labelled. As described in detail in Chapter 2, HEK cells were either transduced with stock diluted in 1:100, 1:250, 1:500 and 1:750 unpseudotyped or EnvA pseudotyped ΔG ChR2-mCherry RB. As no mCherry positive cells were detected in any of the pseudotyped ΔG ChR2-mCherry RB treatment, 1:100 dilution factor example has been shown only. After 7-14 days all cultures were fixed for immunohistochemistry staining. All experiment had a corresponding negative control (virus-free media added). Biological replicates (independent differentiations) n=3. Minimum technical replicates n=6. The average total number of cells counted per technical replicate can be found in Appendix folder 7.

E) Percentage of mCherry expressing HEK cells out of the total cell population (total cell count measured using DAPI count) for each of the unpseudotyped and EnvA pseudotyped ΔG ChR2-mCherry rabies virus transduction treatments. The graph plotted shows the mean percentage of mCherry expressing HEK cells for each treatment, along with +/- SEM. Biological replicates (independent differentiations) n=3. Minimum technical replicates n=6. The average total number of cells counted per technical replicate can be found in Appendix folder 7.

Statistics performed to determine significant differences in percentage of mCherry expressing cells between each treatment condition relative to the highest dilution factor (1:100 either pseudotyped or non-pseudotyped viral tracer) using nonparametric, Kruskal-Wallis test and Dunn's Multiple Comparison Test (*$p<0.05$, **$p<0.01$, ***$p<0.001$, red text pseudotyped, black text unpseudotyped). Mann-Whitney test was performed to determine significant differences in percentage of mCherry expressing cells between corresponding pseudotyped or non-pseudotyped ΔG ChR2-mCherry RB treatments, (e.g. 1:100 unpseudotyped vs 1:100

pseudotyped) with ns $p>0.05$, *$p<0.05$, **$p<0.01$, ***$p<0.001$. Channelrhodopsin 2, ChR2. RB, rabies virus. HEK, human embryonic kidney. Glycoprotein deficient, ΔG. SEM, standard error of the mean.

6.3.6 Monosynaptic viral tracer revealed the existence of small neuronal circuits within hESC-derived interneuronal cultures

With the monosynaptic viral tracer system generated, the EnvA pseudotyped ΔG ChR2-mCherry rabies tracer and hESC-derived starter neurons, the next stage involved the application of the system to hESC-derived interneuron cultures. The hESC-derived interneurons were transduced with the lentiviral vector, which led to 40-50% abundance of starter neurons by day 50 in vitro (Figure 6.5A). The culture was then infected with the pseudotyped viral tracer. By eye the hESC-derived interneuron cultures appeared healthy up to 14 days post-infection, with mCherry positive cells detected from 5 days post-infection onwards. All starter neurons were identified as MAP2 and GFP double positive on immunostaining (Figure 6.3A and 6.5A). The restricted expression of *GFP* to MAP2 positive cells confirmed the fidelity of the lentiviral vector which encoded the genes (TVA, histone-GFP, glycoprotein) under human synapsin promoter (Figure 6.3A and 6.5A). The synaptic inputs onto starter neurons were detected by immunostaining for GFP and mCherry, with infected starter neurons (GFP/mCherry double positive) found to have a relatively small number of inputs (presynaptic partners), typically between 3-5 inputs per starter neuron (Figure 6.5A). A representative immunostaining has been shown in Figure 6.5A, which clearly demonstrates the greater proportion of mCherry positive cells relative to GFP/mCherry double positive cells. Moreover, all mCherry positive traced neurons were identified in close proximity to GFP/mCherry double positive cells, which supports the monosynaptic nature of the tracer (Figure 6.5A). Finally, the restricted uptake of EnvA pseudotyped rabies tracer was confirmed in the hESC-derived interneuron culture using negative control with no starter cells (no lentiviral transduction step), but treated with pseudotyped rabies tracer (Figure 6.5A). The negative control had no GFP or mCherry positive cells, hence the EnvA pseudotyped ΔG ChR2-mCherry rabies tracer could not transduce TVA negative hESC-derived neurons (Figure 6.5A).

6.3.7 The monosynaptic ChR2-mCherry viral tracer system had no consequence on the intrinsic physiology of hESC-derived interneurons

To evaluate the physiology of the traced circuits, the monosynaptic viral tracer system was used in combination with optogenetics to exploit the *ChR2-mCherry* co-expression in rabies tracer infected neurons. However, this ChR2-mCherry viral tracer system has not been previously used in hPSC neuronal cultures. Therefore, preliminary investigations were carried out to check whether either lentiviral transduction to generate starter neurons or infection by the rabies viral tracer had any negative effects on the physiology of the neuronal culture (Figure 6.5B). This was determined by measuring the intrinsic parameters of the patched neurons from the viral free culture (non-transduced neurons), starter neurons (GFP positive) and traced neurons (mCherry positive). As shown in Figure 6.5B, there was no significant difference between any of the groups in their intrinsic physiology. Hence, the resting membrane potential (-19.19mV non-transduced, -22.74mV starter, -17.41mV traced $p>0.05$ Figure 6.5B), membrane time constant (86ms non-transduced, 71ms starter, 75ms traced $p>0.05$ Figure 6.5B), and capacitance (35pF non-transduced, 25pF starter, 37pF traced $p>0.05$ Figure 6.5B) were unchanged by using the viral tracing platform. In addition, a high proportion of the patched neurons (independent of identity) displayed spontaneous activity, both spontaneous mini-excitatory postsynaptic potentials and, less commonly, spontaneous action potentials (see Appendix folder 6 figures). This would support the conclusion that hESC-derived interneurons were successfully forming viable synaptic connections.

6.3.8 Photostimulation efficiently and consistently evoked action potential response in hESC-derived traced neurons

Optogenetics was performed on hESC-derived interneurons 7-14 days post-infection with the pseudotyped ΔG ChR2-mCherry rabies tracer. The traced (mCherry positive) neurons were selected as the recording population, due to their illustrated ability to form synaptic connections (Figure 6.5A). In order to evaluate the efficiency of the ChR2 driven response, traced neurons were stimulated through both current injection and blue light (to activate ChR2) (Figure 6.5C-6.5D). The photostimulated response in traced neurons was found to be equally efficient as current injection, in terms of action potential amplitude, half-width and threshold (Figure 6.5D). This would demonstrate the efficient expression of ChR2 in the traced neurons. However, the photostimulated evoked responses appeared to be more consistent across patched neurons, with a greater variation observed in the current injection responses (Figure 6.5D). The reason for the smaller between-cell variation to photostimulation is unclear. However, a possible explanation may relate to the cell-wide distribution and activation of ChR2, generating a consistent change in membrane potential over the entire cell membrane when exposed to blue light. Whereas, current injection would generate an immediate response in voltage-gated ion channels closest to the patch site, that then propagates over the remaining cell body. Therefore, depending on the ion channel composition closest to this patched site, this may result in a different strength of response, thus causing between-cell variation in spike parameters.

Overall, the monosynaptic viral tracer was successfully applied to hESC-derived interneuron culture. By the general health of the culture, and by the physiology of the neurons, the viral tracer system was concluded to be non-toxic to these cells over 14 day period post-infection with the pseudotyped viral tracer.

Figure 6.5A

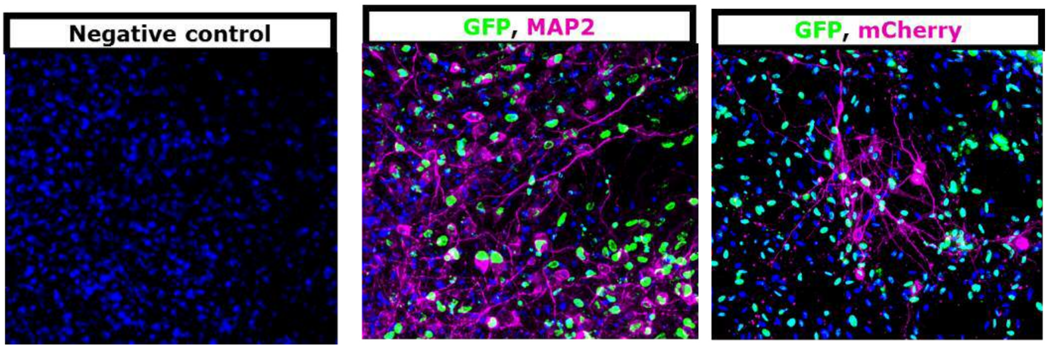

Negative control
GFP, MAP2
GFP, mCherry

Figure 6.5B

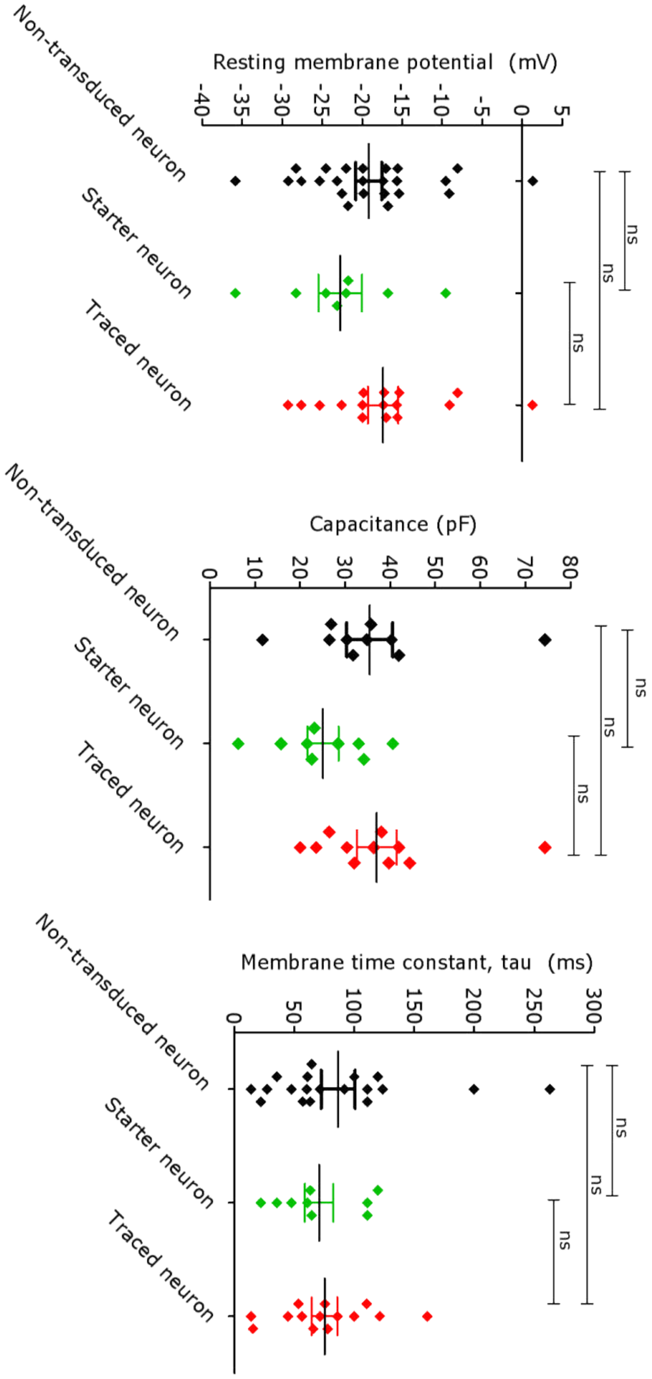
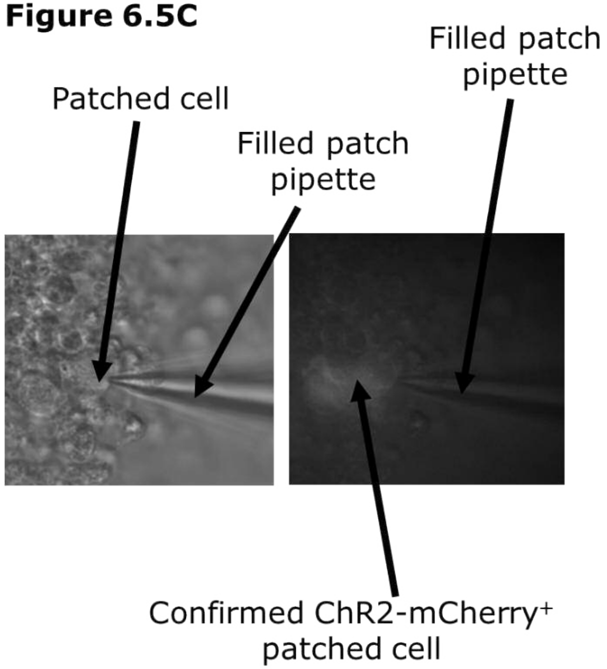
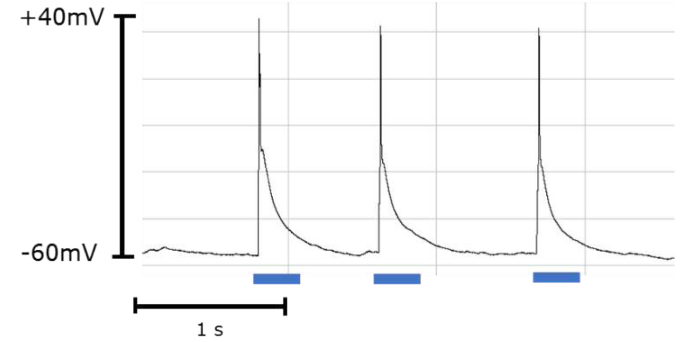
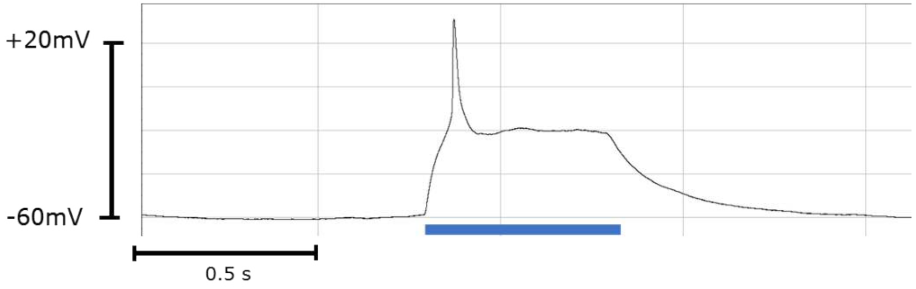


Figure 6.5C


Patched cell
Filled patch pipette
Filled patch pipette
Confirmed ChR2-mCherry⁺ patched cell



490nm blue light pulses (approx. 250ms)



490nm blue light exposure (approx. 500ms)

Figure 6.5D

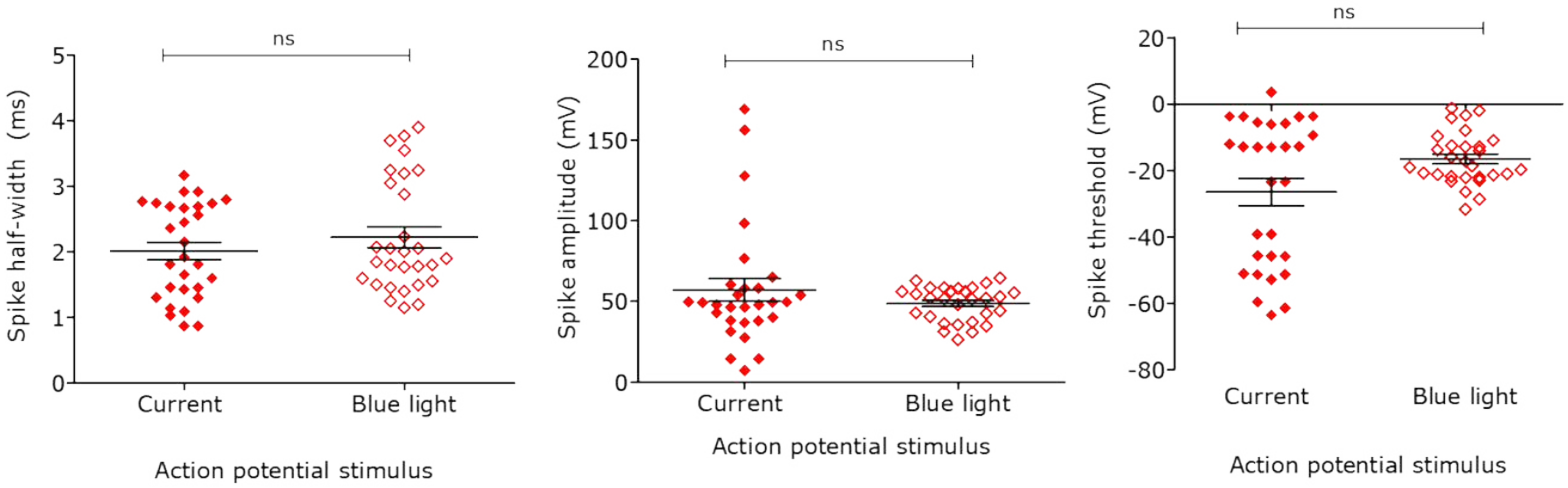

Spike half-width (ms)
0
1
2
3
4
5
ns
Current
Blue light
Action potential stimulus
Spike amplitude (mV)
0
50
100
150
200
ns
Current
Blue light
Action potential stimulus
Spike threshold (mV)
-80
-60
-40
-20
0
20
ns
Current
Blue light
Action potential stimulus

Figure 6.5 Application of EnvA pseudotyped ΔG ChR2-mCherry rabies viral tracer system on hESC-derived interneurons.

A) Day 65 hESC-derived interneurons complete with starter neuron population and up to 14 days post-transduction with EnvA pseudotyped ΔG ChR2-mCherry RB. The data has been presented in the form of a panel of immunohistochemistry deconvoluted images. First starter neuron free culture treated with EnvA pseudotyped ΔG ChR2-mCherry RB (negative control). Second GFP and MAP2 immunohistochemistry staining to confirm restricted co-expression of GFP to MAP2 positive population (due to lentiviral vector driven by synapsin promoter). Third, GFP and mCherry staining to identify non-infected starter neurons (GFP positive), infected starter neurons (GFP/mCherry double positive) and traced neurons (mCherry positive).

B) Intrinsic physiology of patched neurons identified as GFP and mCherry negative (non-transduced population), GFP positive (starter neurons) and mCherry positive (traced neurons). These parameters measured included the immediate resting potential of the neuron (taken within first 30sec of breaking into neuron), membrane time constant (tau) and capacitance (capacitance (pF) = tau (ms) / input resistance (GΩ)). The graph has been plotted with each symbol on the plot representing the value of a single patched neuron, the mean +/- SEM have been shown for each time point. Holding potential, -60mV. Recording conducted from 7 to 14 days post-transduction with EnvA pseudotyped ΔG ChR2-mCherry RB. Biological replicates (independent differentiations) n=3. Technical replicates per group n= 8.

C) Representative example of blue-light evoked action potentials, along with labelled brightfield and fluorescent images of a mCherry positive patched neuron. Holding potential, -60mV. Recording conducted from 7 to 14 days post-transduction with EnvA pseudotyped ΔG ChR2-mCherry RB.

D) Comparison of blue-light and current evoked action potential properties of patched neurons, including spike threshold (mV), spike amplitude (mV) and spike half-width (ms). The graph has been plotted with each symbol on the plot representing the value of a single patched neuron, the mean +/- SEM have been shown for each time point. Holding potential, -60mV. Recording conducted from 7 to 14 days post-transduction with EnvA pseudotyped ΔG ChR2-mCherry RB. Biological replicates (independent differentiations) n=3. Technical replicates per group n= 25.

Statistics performed to determine significant differences in physiological responses of patched cells over time using nonparametric, Kruskal-Wallis test and Dunn's Multiple Comparison Test (*p<0.05, **p<0.01, ***p<0.001). Mann-Whitney test was performed to determine significant differences in evoked responses of patched cells to current injection or blue-light photostimulation (ns p>0.05, *p<0.05, **p<0.01, ***p<0.001). channelrhodopsin 2, ChR2. RB, rabies virus. GFP, green fluorescent protein. hESC, human embryonic stem cell. Glycoprotein deficient, ΔG. SEM, standard error of the mean.

6.4 Discussion

This Chapter has provided an understanding of the functional properties of hESC-derived interneurons cultured under SCM1 treatment (Telezhkin *et al.* 2016; Telezhkin *et al.* 2018). A longitudinal assessment of the intrinsic and evoked physiology of hESC-derived interneurons was determined, along with preliminary evidence of neuronal networks in the in vitro culture using a retrograde monosynaptically spreading viral tracer.

6.4.1 Comparative assessment of the physiology of hPSC-derived interneurons

When considered in relation to similar interneuron differentiation protocols using an extrinsic cue strategy to drive differentiation, the physiology of hPSC-derived interneurons in this study was broadly consistent with previous reports (Cambray *et al.* 2012; Bardy *et al.* 2016; Close *et al.* 2017; Xiang *et al.* 2017). Overall, whole-cell patch clamp recordings have demonstrated the limited ability of cell cycle inhibitors and neurotrophic/neurogenic factors to generate highly functional hESC-derived interneurons in vitro (Cambray *et al.* 2012; Bardy *et al.* 2016; Close *et al.* 2017; Xiang *et al.* 2017). Instead hPSC-derived interneuronal cultures remain abundant in neuron-like cells that were weakly hyperpolarised, had low complexity branching pattern (and thus, low capacitance), and incapable of generating high frequency trains of action potentials (Cambray *et al.* 2012; Bardy *et al.* 2016; Close *et al.* 2017; Xiang *et al.* 2017). Moreover, by the end of the paradigm there was still a proportion of neuron-like cells that remained passive to stimulation, as reported previously (Xiang *et al.* 2017). Hence, these findings reported in this Chapter and independently, emphasise the need for alternative measures to promote functional maturation.

Previous work has demonstrated the successful use of co-cultures of primary rodent derived glial or cortical glutamatergic neurons in generating interneuron-like cells with physiology closer to authentic counterparts (Maroof *et al.* 2013; Sun *et al.* 2016; Noakes *et al.* 2019). Although, arguably the most effective strategy has been in vivo transplantation into the rodent brain (typically neocortex) (Goulburn *et al.* 2011; Nicholas *et al.* 2013). However, completing interneuron differentiation in vivo often extends the duration of the experiment (>5 months), requires additional skill sets and expenses (for animal husbandry and surgery) (Goulburn *et al.* 2011; Nicholas *et al.* 2013; Yuan *et al.* 2018). With clear pros and cons of each strategy, compromises may have to be made when working with PSC-derived neuronal models.

6.4.2 Molecular and functional similarities of hPSC-derived interneurons and primary foetal tissue

Although hESC-derived interneuron physiology described here did not recapitulate that of the endogenous interneurons in the adult brain, how did the physiology profile compare to foetal derived interneurons? Ex vivo recordings have been performed on foetal cortical brain slices across 9WPC-26WPC developmental period (Moore *et al.* 2009; Moore *et al.* 2011; Zhong *et al.* 2018). Overall, these findings reported intrinsic and evoked physiology parameters characteristic of immature neurons, with patch neurons having high input resistance, weakly hyperpolarised resting membrane potentials and weak/absent evoked responses to current injection (Moore *et al.* 2009; Moore *et al.* 2011; Zhong *et al.* 2018). Moreover, at this stage of development cortical neurons had very few neuronal partners (5 or less) and low capacitance (20-30pF) (Moore *et al.* 2009; Moore *et al.* 2011). Thus, foetal derived neurons, from the early second trimester period, had a connectivity and physiological profile similar to those reported for in vitro derived interneurons in this study, and reported independently (Moore *et al.* 2009; Moore *et al.* 2011; Cambray *et al.* 2012; Bardy *et al.* 2016; Close *et al.* 2017; Xiang *et al.* 2017).

6.4.3 Bridging the gap between neuron molecular and functional profiles

The shared phenotypic properties of hESC-derived interneurons and human foetal derived interneurons (from the early second trimester) reflects the similar expression profile of these two sources of interneurons, reported in Chapter 5 and independently in the literature (Close *et al.* 2017; Xiang *et al.* 2017; Yuan *et al.* 2018). The functional immaturity of neurons can be observed in the limited range and low expression level of ion channels, synaptic proteins and neurotransmitter receptors (Close *et al.* 2017; Mayer *et al.* 2019). This was especially true for PVALB interneurons generated in this paradigm, emphasising the difficulty of authentically generating this subtype in vitro (Close *et al.* 2017; Xiang *et al.* 2017; Yuan *et al.* 2018).

Although this study was able to assess the hESC-derived interneuron profile at the functional and molecular level, a more sophisticated approach would have been to carry out patch-seq analysis (Bardy *et al.* 2016; Cadwell *et al.* 2016; Földy *et al.* 2016; Fuzik *et al.* 2015; Cadwell *et al.* 2017; Tripathy *et al.* 2017; van den Hurk *et al.* 2018). The advantage of the patch-seq protocol is that the single-cell physiology data, collected in whole-cell patch clamp recordings, can be directly traced to the single-cell RNA sequencing data of the corresponding cell (van den Hurk *et al.* 2018). With a large enough sample, reliable and reproducible predictions can be made about the neuron physiology profile from the expression of selected genes (Tripathy *et al.* 2017).

The expression profile of hESC-derived interneurons in this study included genes previously named through patch-seq analysis as significantly correlated within a phenotypic trait (Fuzik *et al.* 2015; Tripathy *et al.* 2017). For example, ATPase Na^+/K^+ transporter subunits encoded by *ATP1A1* and *ATP1B3*, voltage-gated Cl^- ion channel subunit (*CLCN3*), inositol 1,4,5-trisphosphate receptor type 1 (*ITPR1*) and K^+ ion two pore domain channel subfamily K member 1 (*KCNK1*) were all reported as negatively correlated with resting membrane potential (Fuzik *et al.* 2015; Tripathy *et al.* 2017). Although this study detected a small, but significant hyperpolarisation of the resting membrane potential for hESC-derived interneurons over time, this was not accompanied by the expression pattern changes in *CLCN3, KCNK1, ATP1A1* and *ATP1B3* predicted by the previous patch-seq data (Fuzik *et al.* 2015; Tripathy *et al.* 2017). Moreover, patch-seq analysis would suggest *CAP2* and *ATP1A1* expression are unrelated predictors of neuronal physiology, which leaves it unclear why *CAP2* and *ATP1A1* had high co-expression, leading to the mirrored temporal changes in abundance of *CAP2* and *ATP1A1* expressing cells in hESC-derived interneurons. The inability of the work described in this Chapter to reproduce previous patch-seq analysis findings may be due to species-differences, with these previous studies being conducted on mouse primary tissue (Fuzik *et al.* 2015; Tripathy *et al.* 2017). Moreover, Fuzik *et al.* study was carried out on a selected interneuron subtype

(CCK interneurons), which were not abundant in the in vitro culture analysed here (Fuzik *et al.* 2015). Nevertheless, as this was not a like-for-like comparison between mouse and human data, and only a small sample size, a conclusive reason cannot be stated for the differences.

The expression of glutamate ionotropic receptor subunits in neuronal populations also provides a predictive marker of neuronal maturity. AMPA, kainate and NMDA receptor subunits in human cortical foetal tissue (8WPC - 22WPC) show an age-dependent expression pattern (Mayer *et al.* 2019). During earlier ages of foetal development, cortical neurons were found to express *GRIN2A*, which becomes downregulated later in development, with the reciprocal rise in the expression of *GRIN2B* (Mayer *et al.* 2019). hESC-derived interneuronal culture generated in this paradigm showed heterogeneity in the expression of *GRIN2A* and *GRIN2B*, suggesting a mixed maturity of the neuron-like cells, corresponding to early-mid second trimester (Mayer *et al.* 2019). Furthermore, this heterogeneity could in part be explained by subtype-dependent differences, with a higher proportion of SST and CR interneuron-like cells co-expressing *GRIN2B.* Whereas, the PVALB interneuron-like cells failed to express this marker. This could be a genuine subtype-dependent difference, or reflect the relative immaturity of the PVALB interneuron-like cells made in vitro (compared to SST and CR interneuron-like cells).

In addition, hESC-derived interneurons had a high abundance of AMPA receptor subunit *GRIA3* expressing cells, which had a subtype-dependent expression pattern, biased to *PVALB* interneurons. Supporting evidence for this finding has come from calcium imaging responses of ex vivo human foetal cortical tissue (14WPC-22WPC) to AMPA mediated-stimulation (Mayer *et al.* 2019). Mayer *et al.* reported a strong AMPA signalling response by MGE-derived cortical interneurons compared to CGE-derived cortical interneurons at the same age (Mayer *et al.* 2019). Moreover, the level of *GRIA3* expression has been positively correlated with firing frequency (e.g. fast spiking PVALB interneurons) (Földy *et al.* 2016).

Additional molecular markers used to predict neuronal physiology are voltage-gated ion channel subunits. For inhibitory interneurons, voltage-gated ion channels show both an age-dependent and subtype-dependent expression pattern (Földy *et al.* 2016; Close *et al.* 2017; Telezhkin *et al.* 2018; Mayer *et al.* 2019). The Kv3 family of ion channels, in particular Kv3.1 (*KCNC1*) and Kv3.2 (*KCNC2*), which are predominantly expressed by interneurons (rather than excitatory neurons), are both positively correlated with high frequency spike patterns typical of PVALB interneurons (Martina *et al.* 1998; Földy *et al.* 2016; Muñoz-Manchado *et al.* 2018). The low abundance of *PVALB, SST* and *CR* hESC-derived interneurons co-expressing *KCNC1* and *KCNC2* likely accounts (in part) for the patched neurons having high action potential half-widths and absence of high frequency trains in response to current injections. Whereas, Kv7 family of ion channels show region-selective expression, with only Kv7.2 (*KCNQ2*) and Kv7.3 (*KCNQ3)* being abundant in the brain (Telezhkin *et al.* 2018). This study reported a similar biased expression pattern, with hESC-derived neurons preferentially abundant in *KCNQ2* and *KCNQ3* expressing neurons, with much lower/absent expression of *KCNQ4* and *KCNQ5*. Moreover, the single-cell RNA sequencing data presented here would indicate a PVALB interneuron subtype bias in *KCNQ2* expression. Although this is a novel report, conditional (*Nkx2.1::Cre* driven) *Kcnq2* KO mice have inappropriately excitable PVALB interneurons, with SST interneuron function left intact (Soh *et al.* 2018). The PVALB dysfunction led to a higher susceptibility to seizures in the conditional *Nkx2.1::Cre* and *Pvalb::Cre* driven *Kcnq2* KO mice (Soh *et al.* 2018). This mouse data would provide supporting evidence for the biased expression of *KCNQ2* in hESC-derived PVALB interneurons reported in this study (Soh *et al.* 2018).

The absence/low abundance of Cav1.1 (*CACNA1S*) and Cav1.4 (*CACNA1F*) expressing hESC-derived interneurons would be predicted from rodent work demonstrating relatively low neuronal expression of *CACNA1S* (which is mainly restricted to skeletal muscle) and largely retinal restricted expression of *CACNA1F* (Heyes *et al.* 2015). Whereas, Cav1.2 (*CACNA1C*) and Cav1.3 (*CACNA1D*) were equally co-expressed among *PVALB, CR* and *SST* expressing hESC-derived interneurons (although at a low proportion). This subtype independent expression pattern was consistent with rodent primary brain tissue immunostaining illustrating a brain-wide distribution of these genes (Heyes *et al.* 2015). Furthermore, L-type Cav1.2/1.3 ion channels have been implicated in the regulation of interneuronal growth and maturation (Jiang &

Swann 2005). On agonist-mediated activation of Cav1.2/1.3 ion channel activity in ex vivo mouse hippocampal slice cultures, PVALB interneurons displayed upregulated expression of *Vgat* and *Gad1*, and grew more extensive dendritic branches (Jiang & Swann 2005). The reciprocal was true when Cav1.2/1.3 ion channel activity was dampened using selective Cav1.2/1.3 antagonists (Jiang & Swann 2005). Therefore, the low abundance of *CACNA1C* and *CACNA1D* expressing hESC-derived interneurons may account for the small proportion of *vGAT* expressing cells, and absence of complex branch patterns and connectivity.

6.4.4 Viral tracers and insight into in vitro network activity and connectivity

The application of the EnvA pseudotyped ΔG ChR2-mCherry rabies tracer system on hESC-derived interneurons provided novel proof-of-principle data demonstrating selective and efficient uptake and monosynaptic transmission of the viral tracer within this in vitro culture. Previous reports have suggested a restricted timeframe for analysis of the rabies viral tracer of 5-7 days post-transduction, but this data demonstrated the non-toxicity of the viral tracer up to 14 days (Wickersham *et al.* 2007; Osakada *et al.* 2011; Osakada & Callaway 2013). Although this may be cell line and/or neuronal-type specific. Moreover, the traced circuits were limited to infected starter neurons and their presynaptic partners, with a higher abundance of traced neurons to infected starter neurons in the population. Although multiple synaptic connections were made onto starter neurons, the overall morphology and branch pattern observed in traced neurons typically lacked the complexity of mature interneurons. This branching pattern and connectivity may have been enhanced by co-culturing the hESC-derived interneurons with endogenous partners, such as hESC-derived cortical glutamatergic neurons (Sun *et al.* 2016). In addition, such experiments exploring the network connectivity of hESC-derived interneurons with native synaptic partners would provide a more translatable and authentic measure of the connectivity of hESC-derived interneurons.

Although all traced neurons displayed evoked responses to both current and blue light stimulation, the experiments did not directly demonstrate the viable nature of the synaptic connection made between traced and infected starter neurons. Previous work with rabies viral tracers has shown that trans-synaptic spread of the viral tracer does not require the synapse to be viable (Wickersham *et al.* 2007). Furthermore, the exact mechanism of viral trans-synaptic transmission is still unclear, with unanswered concerns over whether strength/activity of the synapse affects the efficiency of trans-synaptic transmission (Wickersham *et al.* 2007; Ginger *et al.* 2013). Hence, a future requirement would be to conduct paired recordings stimulating (through current injection/blue light) traced neurons and recording from monosynaptically connected infected starter neurons. Only if a parallel response was detected in the downstream starter neuron could the synapse be confirmed as viable. Although direct evidence of functional connectivity is a critical accompaniment to viral tracing, the viable nature of the connections would be predicted due to the detection of spontaneous activity (an indication of synaptic stimulation) in the majority of patched neurons from day 50 onwards.

6.4.5 Conclusion

In conclusion, post-mitotic hESC-derived interneurons closely recapitulate both the molecular and functional profile of early second trimester (15WPC) human foetal derived interneurons (Close *et al.* 2017; Xiang *et al.* 2017; Yuan *et al.* 2018). This highlights a shortcoming of hPSC-derived neurons as a representative model of adult brain neuronal activity, which may be improved by co-culturing with native neuronal partners. Nonetheless, hPSC-derived interneurons hold promise as a neurodevelopmental model, mimicking many of the molecular and functional changes observed in early human foetal development.

Chapter 7 - General Discussion

7.1 Summary

The molecular steps underpinning interneuron diversification into the spectrum of subtypes present in the adult human brain has proved challenging to define. Within this thesis a hPSC model depiction of human interneuron development has been described. The study focused on exploring the utility of hESC-derived interneurons as a neurodevelopment model, with the overall aim to determine whether the molecular and functional transition that marks interneuron differentiation trajectory can be recapitulated in vitro.

The evaluation began with a longitudinal characterisation of the hESC-derived interneurons (Chapter 3-4). In order to tease apart the molecular heterogeneity within and across populations, longitudinal single-cell RNA sequencing was performed. Overall, these findings were promising, demonstrating the ability to generate an enriched culture of MGE-like progenitors that gave rise to a diverse set of interneuron subtypes in vitro (Chapter 4). Moreover, temporal modifications to the paradigm may provide a means to further refine the hESC-derived interneuron subtype composition (Chapter 3).

However, with valid concerns over how true the in vitro representation of interneuron differentiation is, a novel cross-comparison of the model was made to their authentic counterparts in the human MGE (Chapter 5). The foetal tissue was obtained across the late first trimester to early second trimester (10WPC to 15WPC), a developmental period which encompassed the emergence of subtype committed interneuron precursors in the MGE. The single-cell RNA sequencing analysis enabled novel age and subtype-dependent expression patterns to be detected, leading to candidate lineage determining genes to be named. Moreover, encouraging similarities were found between the molecular profiles of hESC-derived interneurons and human MGE. However, clear distinctions were also detected, which marked the shortcomings of the in vitro model.

To expand upon the longitudinal molecular characterisation, the in vitro model was also evaluated based on the physiology and connectivity of the post-mitotic neuronal cultures (Chapter 6). Through whole-cell patch clamp recordings, functional characterisation of the interneuron-like cells was achieved. The molecular heterogeneity in hESC-derived interneurons was also reflected in the diversity of physiology profiles recorded. Although, the intrinsic and evoked neuronal responses of hESC-derived interneurons poorly recapitulated that of their counterparts in the adult brain. Instead, the differentiation trajectory of the interneuron-like cells was marked by the transition to a modestly higher functional maturity, closely resembling that of native foetal cortical interneurons. In addition, hESC-derived interneurons were capable of forming small neuronal networks, as demonstrated using the monosynaptic viral tracer system. Although the connectivity was not extensive, the ability of interneuron-like cells to form synaptic connections marked the successful recapitulation of a late stage of interneuron development in vitro.

In summary, the interneuron differentiation of hESC has potential as a platform for neurodevelopmental research. However, certain refinements to the paradigm would be required in the future to help close the gap between hESC-derived interneurons and their native counterparts. This work could be guided by the novel cross-comparison reported between the in vitro model and the human MGE.

7.2 Questions, limitations and future studies

7.2.1 Exploring the changes driven by serial expansions

Sequential serial expansions of hESC-derived interneuron progenitor cultures were explored as a means to recapitulate the generation of temporally distinct newborn progenitors in the MGE over development. Fate-mapping studies of grafts derived from different aged MGE and CGE would indicate that early and late-born progenitors have different lineage commitment biases (Butt *et al.* 2005; Miyoshi *et al.* 2010; Inan *et al.* 2012). Therefore, this manipulation could enable a wider spectrum of interneuron subtypes to be generated in vitro. As described in Chapter 3, changes in the molecular profile of both the progenitor and terminally differentiated cultures were identified. Ultimately, this led to an elevation of SST:CR interneuron subtype composition in the terminally differentiated cultures. However, the experimental design used could not distinguish between this shift being caused by the change in the lineage commitment of the progenitors (e.g. from CR to SST lineage), or the selective survival of SST interneuron precursors in the culture (Figure 7.1). To answer this question the experimental design would need to be modified to allow tracing of the progeny from the same clonal expansion (progenitors born from divisions originating from the same ancestral progenitor) across sequential cultures.

7.2.1.1 Combining lineage tracing and single-cell RNA sequencing

Retroviral barcoding is the ideal tool for clonal lineage studies, providing a high throughput method to follow the differentiation trajectory of a parent cell and all its progeny (Yao *et al.* 2016; Biddy *et al.* 2018). The barcodes are short sequences of nucleotides that become integrated into the cell genome using lentiviral vectors, and are therefore heritable across generations (Biddy *et al.* 2018; Kester & van Oudenaarden 2018). By using a variety of different barcode sequences, the progeny derived from different ancestral progenitors can be independently traced. Furthermore, use of barcodes that are transcribed (e.g. sequence placed under an ubiquitous promoter), lineage tracing and single-cell RNA sequencing can be combined to characterise the molecular identity of the population and understand how cells are clonally related to one another (Kester & van Oudenaarden 2018). By implementing this combinational approach during the serial expansion experiment in Chapter 3, by first retrovirally labelling the day 20 hESC-derived interneuron progenitors and then performing single-cell RNA sequencing analysis on each subsequent culture generated by serial expansion, one could tease apart the possible causes of the elevated SST:CR (Figure 7.1).

Figure 7.1

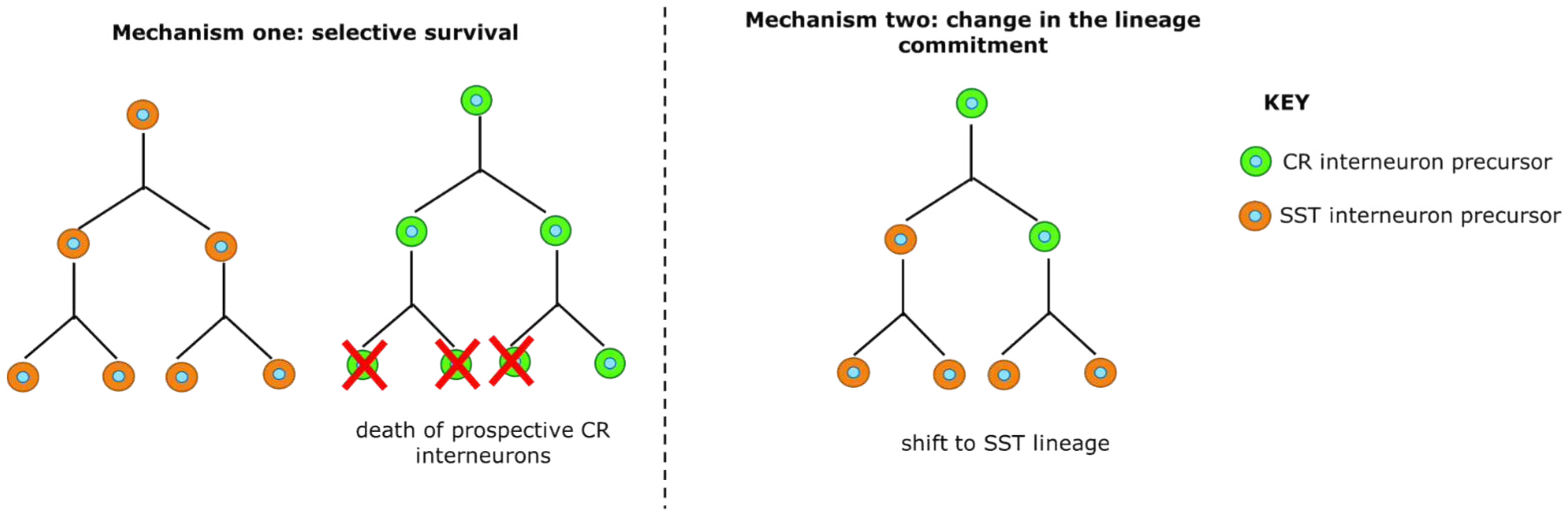

Mechanism one: selective survival
death of prospective CR interneurons
Mechanism two: change in the lineage commitment
shift to SST lineage
KEY
CR interneuron precursor
SST interneuron precursor

Figure 7.1 Clonal lineage and fate restrictions.

Questions still remain over how restricted the lineage commitment of interneuron precursors is, and if there is flexibility in the interneuron fate, what factors can influence this? In an attempt to recapitulate the generation of early and late-born progenitors in the developing CGE/MGE, serial expansion of hESC-derived interneuron progenitors was performed (see Chapter 3). On characterising the terminally differentiated cultures from each serial expansion, a change in subtype composition was detected, an elevation in SST:CR in later expansions. This could be explained by alternative mechanisms, which have been depicted in left and right-hand schematics in Figure 7.1. **Mechanism one,** interneuron precursors from the same clonal lineage are restricted to one fate, and SST:CR shift was simply due to the death of CR interneuron precursors in later serial cultures. **Mechanism two,** interneuron precursors derived from the same ancestor have the potential to become different interneuron subtypes (e.g. SST and CR), and the serial expansions triggered this change in favour of SST lineage.

SST, somatostatin. CR, calretinin. CGE, caudal ganglionic eminence. MGE, medial ganglionic eminence. hESC, human embryonic stem cells.

7.2.1.2 Dual reporter hESC lines

Another question raised through molecular characterisation of each serial expansion of the interneuron-like progenitors was how well the expansion led to newborn progenitors being generated. By later serial expansions it was generally concluded that the progenitors were failing to divide, and were transitioning to a post-mitotic state even under SHH treatment (Chapter 3). Using dual reporter hPSC line that differentially labelled progenitors from immature neurons would be an ideal and simple method to selectively enrich the next progenitor culture with only proliferating progenitors. A prime example of a dual reporter PSC line used to distinguish between interneurons at different stages of differentiation was the mESC dual reporter line *Nkx2.1::mCherry, Lhx6::GFP* (Tyson *et al.* 2015). In this case the mESC line was exploited to specifically distinguish between earlier-born (older) MGE-like immature neurons (*Lhx6::GFP* positive) from later-born (younger) MGE-like progenitors (*Nkx2.1::mCherry* positive) (Tyson *et al.* 2015). However, this exact reporter strategy may not be ideal for the hESC-derived interneuron culture generated in this study, due to the greater mix of MGE-like and CGE-like progenitors. Therefore, ideally hPSC line with region of origin independent markers, which also display temporal dynamics across interneuron development, would be used to drive the reporter expression (potentially *SOX2* and *DLX5*).

7.2.2 Exploring the means to better bridge the gap between transcriptome and physiology

The longitudinal single-cell RNA sequencing performed on hESC-derived interneurons was presented and analysed in Chapter 4. The molecular characterisation was then subsequently examined in parallel to the physiology of hESC-derived interneurons determined using whole-cell patch clamping (Chapter 6). However, the independent collection of the molecular and functional characterisation data raised several caveats. One issue of particular relevance was the inability to directly pair the gene expression profile of the neuron to the physiology of the same neuron (Bardy *et al.* 2016; Cadwell *et al.* 2016; Tripathy *et al.* 2017; van den Hurk *et al.* 2018). This restricted our understanding of how the molecular heterogeneity of post-mitotic interneuron-like cells shaped the diverse electrophysiology profiles recorded in the population.

7.2.2.1 Patch-seq

To achieve a more refined understanding of the link between the expression profile and neuronal physiology, a strategy combining whole-cell patch clamp recordings with single-cell RNA sequencing has been developed (Bardy *et al.* 2016; Cadwell *et al.* 2016; van den Hurk *et al.* 2018). As introduced in Chapter 6, this methodology is named patch-seq (Bardy *et al.* 2016; Cadwell *et al.* 2016; van den Hurk *et al.* 2018). In contrast to the C1 Fluidigm microfluidic system, patch-seq protocols rely on following up whole-cell patch clamp recordings with the careful manual isolation of the neuron (Figure 7.2). However, once the cells are captured the reaction chemistry for generating single-cell cDNA is conducted in a similar manner, with parallel reactions carried out on the isolated cells in the chosen reaction chemistry (e.g. SMART-seq). Moreover, the addition of unique barcodes (short nucleotide sequences) to each reaction chemistry acts to tag the cDNA fragments (Figure 7.2). The barcode is used to correctly assign the single-cell RNA sequencing data to the electrophysiology data collected from the same cell (van den Hurk *et al.* 2018). By identifying any relationship between genes and physiology across the single-cell samples, prediction models can be generated that allows a particular neuronal trait (e.g. firing frequency) to be predicted by the expression level of a specific gene (Figure 7.2). However, caution should be placed over such prediction models with too small sample sizes (Cadwell *et al.* 2016).

In summary, patch-seq data has the potential to provide an enhanced, novel understanding of the link between the transcriptome and physiology of hESC-derived interneurons. Moreover, patch-seq data compared between different neuronal classes from adult mouse primary tissue would indicate that the gene-physiology patterns are well conserved across neuronal lineages (Tripathy *et al.* 2017). Therefore, this work could be applied beyond interneurons. Overall, this line of work would be worth pursuing in the future. Although, fine-tuning and optimisation of the patch-seq protocol would be required to ensure that the single-cell libraries were at the same high standards as achieved with the Fluidigm system.

Figure 7.2

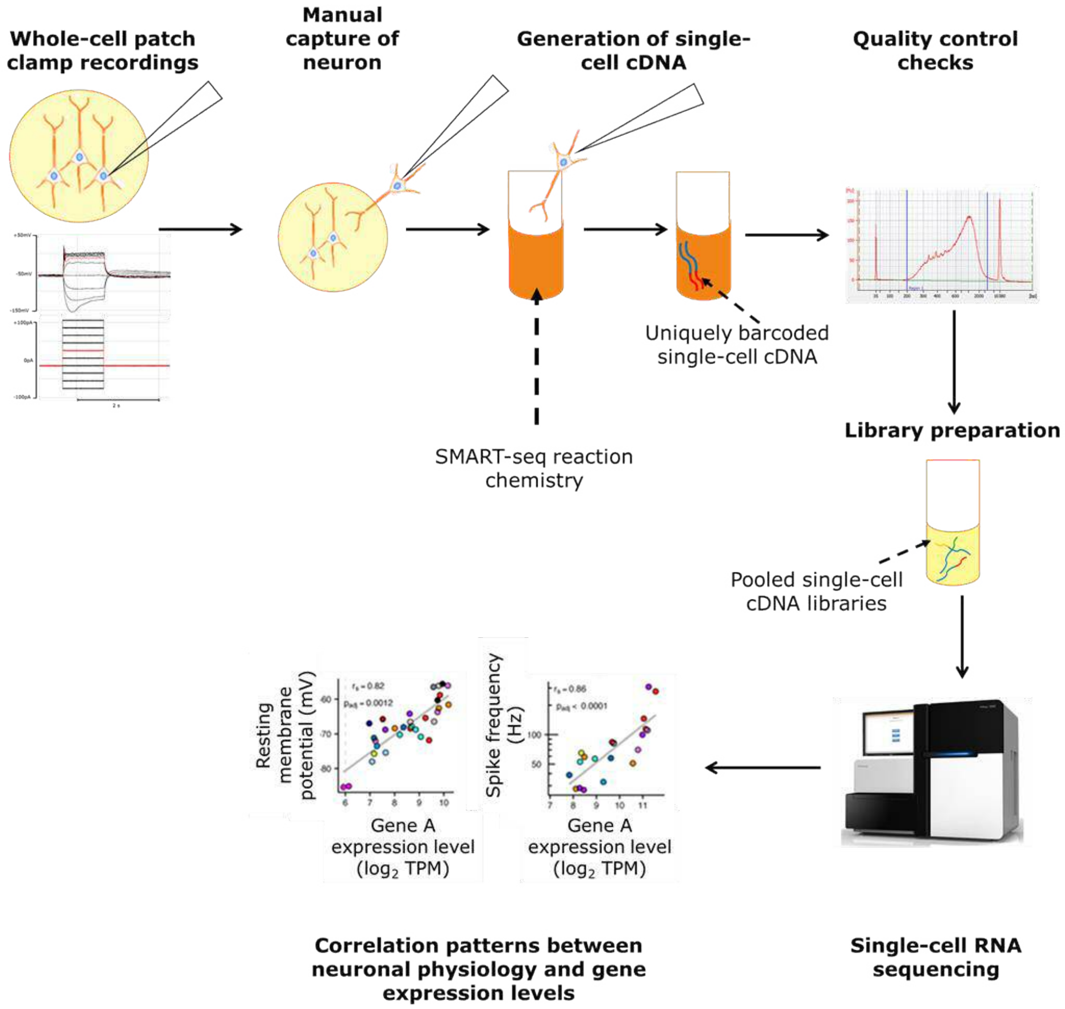

Whole-cell patch clamp recordings
Manual capture of neuron
Generation of single-cell cDNA
Quality control checks
Uniquely barcoded single-cell cDNA
SMART-seq reaction chemistry
Library preparation
Pooled single-cell cDNA libraries
Resting membrane potential (mV)
r_s = 0.82
p_{adj} = 0.0012
-60
-70
-80
6
7
8
9
10
Gene A expression level ($\log_2$ TPM)
Spike frequency (Hz)
r_s = 0.86
p_{adj} < 0.0001
100
50
7
8
9
10
11
Gene A expression level ($\log_2$ TPM)
Correlation patterns between neuronal physiology and gene expression levels
Single-cell RNA sequencing

Figure 7.2 The principle of patch-seq.

The basic methodology of patch-seq involves carrying out whole-cell patch clamp recordings, and on completing recordings, manually picking the patched neuron (using positive pressure through the patch-pipette). The neuron is then immediately transferred into an isolated tube/well with the appropriate reaction chemistry (e.g. SMART-seq) that marks the start of the single-cell RNA sequencing steps. These steps involved lysing the cell to extract RNA, which is then reverse transcribed to generate cDNA that is uniquely barcoded and preamplified ready for downstream application (e.g. quality control checks and single-cell library preparation) (Bardy *et al.* 2016; Cadwell *et al.* 2016; Fuzik *et al.* 2016; Cadwell *et al.* 2017; van den Hurk *et al.* 2018). On repeating the steps described above, a large number of single-cell cDNA samples can be generated, each with their own unique barcode. The barcode is then used to back-trace the single cell expression data to the electrophysiology profile of the same cell. The final step involves performing a cross-comparison between the physiology parameters and expression profile for each cell, any comparisons with a significant gene-physiology correlation across the single cell samples can be used as a means to make predictions about the physiology of a neuron from the expression profile (and visa versa). Figures were adapted from Tripathy *et al.* 2017.

TPM, transcripts per million.

7.2.3 Exploring the reasons behind the MGE and in vitro model discrepancies

The first account of the longitudinal molecular characterisation of human MGE at single-cell resolution was presented in Chapter 5. The human MGE single-cell profiles were used as the gold-standard to measure the hESC-derived interneuron single-cell profiles to. Although this did demonstrate many common features between the in vivo and in vitro derived cell expression profiles, in other respects the profiles did not match. Why?

7.2.3.1 MGE developmental period

There are many factors that could have contributed to the mismatch between the MGE and hESC-derived interneurons. One possible factor that could have augmented the differences was the limited range of MGE ages selected for comparison (10WPC to 15WPC). During human development, the ganglionic eminence becomes established in the ventral telencephalon at 7WPC, and then spatial patterning divides the ganglionic eminence into the three regions (MGE, CGE, LGE) by 8WPC (Zecevic *et al.* 2011; Onorati *et al.* 2014). The MGE continues to generate newborn interneurons into the late second trimester. Therefore, to encompass the full period of human MGE development would be challenging, both in terms of labour, cost and time. However, with an incomplete profile of MGE development, it is hard to fully assess the overall similarity of hESC-derived interneurons to the native counterparts in the MGE.

To help narrow down how large an expanded developmental period would be necessary, previous foetal primary tissue to PSC-derived neuronal model cross-comparisons can be used. These independently reported comparisons shared the same conclusion that at the molecular level post-mitotic hPSC-derived neurons most closely resembled their counterparts in early second trimester cortical foetal tissue (14WPC – 16WPC) (Close *et al.* 2017; Xiang *et al.* 2017; Yuan *et al.* 2018). Therefore, extending the developmental period beyond 15WPC may not be of high enough value to justify the additional expenditure. Whereas, evaluation of the hESC-derived interneuron progenitors (which was poorly matched to all MGE ages sampled) may benefit from collecting earlier MGE ages (from 8WPC) when the ventricular/subventricular zones were more extensive in the MGE (Hansen *et al.* 2013).

7.2.3.2 MGE vs CGE

An additional question of significance was whether a comparison between hESC-derived interneurons and human MGE alone could fairly evaluate the in vitro cultures. As demonstrated at protein and RNA level, the hPSC interneuronal differentiation paradigm utilised here gave rise to a mix of MGE-like and CGE-like progenitors (Chapter 3 and 4). Although the culture was significantly biased towards NKX2.1 positive MGE-like progenitors, the most abundant subtypes in the terminally differentiated cultures were CR, SST and RELN. These are classically viewed as more dorsal MGE and CGE-derived interneurons (Flames *et al.* 2007; Wonders *et al.* 2008; Miyoshi *et al.* 2010). Therefore, future work carrying out single-cell RNA sequencing analysis independently on human dorsal and ventral MGE and CGE would be encouraged, and has already proved fruitful in mouse (Mayer *et al.* 2018; Mi *et al.* 2018). From this work there is the potential of understanding the molecular mechanism of interneuron diversification for a much wider spectrum of interneuron subtypes. Hence, this data would be the optimal comparison to the mixed CGE and MGE-like cells obtained in vitro, and may provide a better guide for refinements to the in vitro protocol.

7.2.4 Exploring ways to improve hESC-derived interneuron functional maturity

A principle aim of Chapter 6 was to monitor the developmental trajectory of the interneuron-like cells at the functional level, both as a population (network connectivity) and on a single-cell basis (physiology parameters). Although early spatial and temporal factors, within the region of origin, influence the lineage commitment of interneurons, the final moulding of the interneuron molecular and physiological profile is likely subject to extrinsic cues in the region of destination (Kessaris *et al.* 2014; Ishino *et al.* 2017; Quattrocolo *et al.* 2017; Wamsley & Fishell 2017; Mayer *et al.* 2018). However, in light of the immature functional status of the interneuron-like cells demonstrated in Chapter 6, to successfully model this late stage of interneuron development in vitro would require refinements to the paradigm used.

7.2.4.1 Co-cultures of native partners

One proposed strategy to generate more functionally mature PSC-derived neurons in vitro has been to co-culture neurons with their native synaptic partners (Maroof *et al.* 2013; Fantuzzo *et al.* 2017; Sarkar *et al.* 2018). In addition to this co-culturing potentially enhancing the functional maturity of the hESC-derived interneurons, the platform also opens up the avenue to explore how co-culturing early post-mitotic hESC-derived interneurons with different synaptic partners (typical of different brain regions) may affect the differentiation potential of the interneuronal culture (Bele *et al.* 1995; Rolando *et al.* 2010; Quattrocolo *et al.* 2017). Within the adult brain there are brain region-dependent differences in interneuron subtype diversity and abundance, which also coincides with differences in synaptic partners and extracellular infrastructure (e.g. perineuronal nets) (Zeisel *et al.* 2015; Lake *et al.* 2016; Muñoz-Manchado *et al.* 2018). However, questions still remain over whether the region-dependent differences in interneurons was already intrinsically programmed in the interneuron precursor or driven by region-dependent extrinsic cues (Bele *et al.* 1995; Rolando *et al.* 2010; Ishino *et al.* 2017; Quattrocolo *et al.* 2017).

The malleable nature of in vitro co-cultures would make the system ideal to explore the influence of extrinsic cues, in the form of different synaptic partners. With efficient hPSC differentiation paradigms for multiple interneuron synaptic partners, including cortical pyramidal cells, hippocampal pyramidal cells and striatal MSNs, multiple region-specific circuits can be reproduced in vitro (Arber *et al.* 2015; Sarkar *et al.* 2018). Moreover, novel microfluidic cell culture platforms can be used to restrict co-cultured populations to contact only, with no shared media (Fantuzzo *et al.* 2017; Sarkar *et al.* 2018; Virlogeux *et al.* 2018). This is achieved by the co-culture plate consisting of two adjacent and independent chambers connected by short (450μm long) grooves that allow neuronal processes to pass between chambers. This design enables synaptic connections to be made, between the two otherwise isolated neuronal cultures (Fantuzzo *et al.* 2017; Sarkar *et al.* 2018; Virlogeux *et al.* 2018). Therefore, the effect of different synaptic inputs/outputs alone can be monitored. Overall, this tool has expanded the manipulation possible in neuronal co-cultures.

7.3 Wider application

7.3.1 Understanding the vulnerability of interneurons in neurodevelopmental disorders

The longitudinal single-cell RNA sequencing analysis of human MGE has a wide scope of application, beyond what was explored here. Although not the primary focus of this study the data could be analysed in relation to known neuropsychiatric risk genes, a line of investigation successfully explored using single-cell RNA sequencing of mouse foetal cortical tissue (Loo *et al.* 2019). Interneurons have been implicated in neuropathologies underpinned by the proposed excitation-inhibition imbalance theory (Rutter *et al.* 2009; Hirano *et al.* 2008; Haenschel *et al.* 2009). However, whether interneuron dysfunction/loss was a primary cause of the pathology, or a downstream effect remains uncertain. If known risk genes were expressed during native interneuron development, this would give support for the primary role of interneuron dysfunction/deficit in the disorder, and encourage interneuron targeted therapeutic approaches. Furthermore, the reason for the selective vulnerability of interneuron subtypes in neuropathology is unknown (Robbins *et al.* 1991; Reynolds *et al.* 2002; Lewis *et al.* 2005). By exploring the risk gene expression pattern in the human MGE at single-cell resolution, evidence for subtype-selective expression may arise, therefore explaining the phenomenon.

7.3.2 Refining the interneuron differentiation of hPSC

The comparative assessment of hESC-derived interneurons to the human MGE single-cell profiles in Chapter 5 brought potential avenues to refine and optimise the hPSC interneuron differentiation paradigm to light. The longitudinal single-cell RNA sequencing analysis named signalling pathways active in the human MGE. Therefore, activation of these named pathways, including insulin-like growth factor signalling and secretogranin V neuroendocrine signalling, during the in vitro differentiation paradigm, may provide a more complete recapitulation of the native MGE environment, thereby giving rise to more authentic interneuron-like cells. Confidence in the candidate signalling pathways named was drawn from the inclusion of activin and BMP4 signalling inhibition, which has been previously demonstrated in mouse to promote interneuron differentiation (especially towards SST fate) (Mukhopadhyay *et al.* 2009).

Alternatively, an intrinsic cue approach could be taken to enforce the expression of transcription factors associated with interneuron fate, but were poorly expressed in the hESC-derived interneuron culture, such as *MEG3, DLX5, POU3F4, NFI B* (Chapter 5). These transcription factors have been repeatedly named across mouse CGE and MGE single-cell RNA sequencing analysis, and in the human MGE data presented here (Chen *et al.* 2017; Mayer *et al.* 2018; Mi *et al.* 2018). Although these genes were not shown to be selective to one particular interneuron lineage, these genes may overall enhance the authenticity of the interneuron-like cells generated. On the other hand, *SKP1* was detected as a gene specifically associated with *NPY* expressing subpopulation in the MGE. To determine whether the association was causal, future experiments manipulating the expression of *SKP1* in interneuron-like progenitors would be required.

One means to perform targeted manipulation of gene expression is through the use of the dead Cas9 (dCas9)– transcriptional effector system (Figure 7.3). This sophisticated system takes advantage of the precise and high fidelity DNA targeting possible with endonuclease Cas9- guide RNA complex, which has been extensively used as a genomic editing tool (Anton *et al.* 2018). Through modifications to the Cas9-complex to inactive the endonuclease activity, and to include a transcriptional activator/repressor protein fused to the dCas9, a versatile tool has been generated to modify the expression level of one or more genes of interest (Qi *et al.* 2013; Kearns *et al.* 2014; Chavez *et al.* 2015; Anton *et al.* 2018). This approach has been successfully applied to reprogram hPSC to a neuronal fate by targeting the promoter regions of *NGN2* and neurogenic differentiation factor 1 (*NEUROD1*) in hPSC using dCas9-transcriptional activator system (Chavez *et al.* 2015). By applying the dCas9- transcriptional effector system to upregulate candidate lineage determining genes named in this study, this may provide a means to generate interneuron subtype-selective cultures. Research pursuing the establishment of differentiation paradigms that give rise to a more defined interneuron subtype composition would be of high value for both research and regenerative medicine. This is largely due to the subtype-dependent implication of interneurons in different neuropathologies (Robbins *et al.* 1991; Reynolds *et al.* 2002; Lewis *et al.* 2005; Ito-Ishida *et al.* 2015).

Figure 7.3

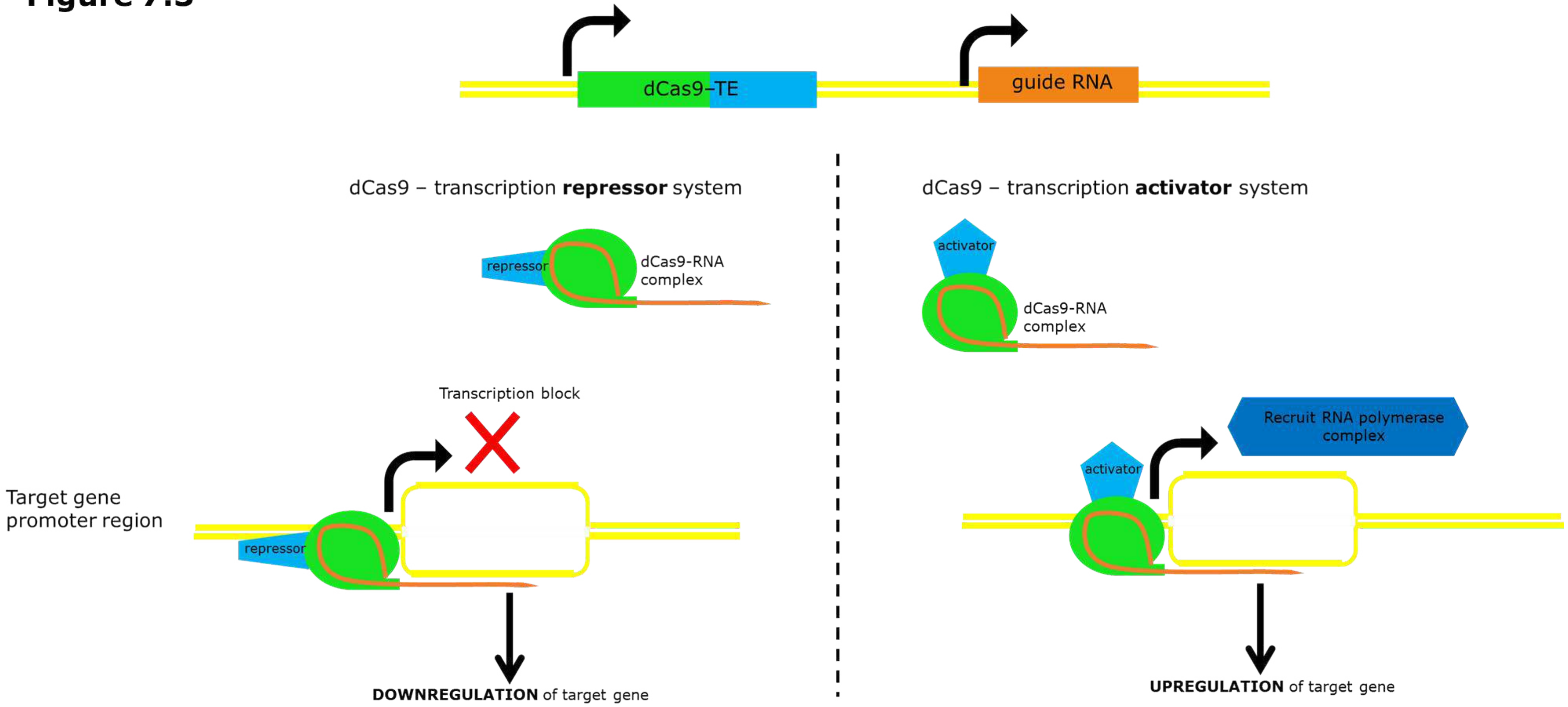

dCas9–TE
guide RNA
dCas9 – transcription **repressor** system
repressor
dCas9-RNA complex
Transcription block
Target gene promoter region
repressor
DOWNREGULATION of target gene
dCas9 – transcription **activator** system
activator
dCas9-RNA complex
Recruit RNA polymerase complex
activator
UPREGULATION of target gene

Figure 7.3 Application of the dCas9 – transcriptional effector system for targeted manipulation of gene expression.

The catalytically inactive endonuclease, dCas9, complex can target regions of the genome using complementary guide RNA sequences. By modifying the dCas9 gene to encode for dCas9 – transcription effector fusion protein has provided an elegant means to modify the expression of targeted genes. The target gene expression levels can be driven up or down depending on the choice of transcriptional effector (repressor or activator respectively). This is achieved by the repressor and activators acting to block or recruit the assembly of the RNA polymerase complex, respectively. Examples of repressor domains successfully used in the system are Krüppel-asociated box or four-linked mSin3 interaction domains (Qi *et al.* 2013). Examples of activator domains are VP64 or p65 (Qi *et al.* 2013). The use of the dCas9 – transcriptional effector system provides an ideal means to directly test the effect of candidate lineage determining transcription factors, and other genes of interest, on the differentiation potential of hESC-derived interneuron precursors.

hESC, human embryonic stem cell. dCas9, endonuclease inactive/dead Cas9. TE, transcriptional effector.

7.3.3 In vitro disease modelling and drug development

Although a purely neurodevelopmental focused application of hPSC-derived interneurons was used here, in vitro interneuronal models have a much wider scope of applications (Figure 7.4). In particular hPSC-derived interneurons have been an invaluable tool in investigating the interneuron-specific deficits in different neurodevelopmental disorders, including Rett syndrome, ASD and schizophrenia (Li *et al.* 2013; Patriarchi *et al.* 2016; Shao *et al.* 2019). This is particularly true since hPSC platform expanded to include induced pluripotent stem cells (iPSC). iPSC are generated from somatic cells (typically fibroblasts) taken from patients or healthy controls, and these specialised cells are reprogrammed using a specific set of transcription factors ("Yamanaka factors") to convert the somatic cells back into PSC (Yu *et al.* 2007). Through the use of patient iPSC derived or mutant hESC derived interneurons an understanding of how unknown and known genetic insults affect different stages of the neuronal differentiation has been achieved (Li *et al.* 2013; Patriarchi *et al.* 2016; Shao *et al.* 2019) (Figure 7.4). Once deficits in the disease model are determined, different tools can be used to determine the molecular mechanism underpinning the mutant phenotype, leading to the identification of novel, candidate therapeutic targets (Shao *et al.* 2019). Genetic and pharmacological approaches can then be applied in attempt to correct the phenotype (Shao *et al.* 2019). In summary, the nature of hPSC-derived neuronal models enables high throughput screening of candidate therapeutic treatments, thus making the platform a valuable asset in drug development.

7.3.4 Regenerative medicine

Finally, hPSC-derived interneuron progenitors hold promise within regenerative medicine (Figure 7.4). Cell replacement therapy has been proposed as a therapeutic strategy in patients with traumatic brain injury (e.g. stroke) or neurodegenerative disorders (e.g. Parkinson's disease) were pharmacological approaches are not available and/or have undesirable side-effects. Classically, cell replacement therapy was coined as a means to deliver the appropriate neuronal progenitors into the atrophied region of the host brain, which could then integrate into the host circuits to rebuild the deteriorated circuits. Over time, cell replacement therapy benefits have expanded beyond this initial principle, by using grafts as a vehicle to deliver neurotrophic, neurogenic and anti-inflammatory small molecule factors, which can help reduce the rate of further deterioration in the brain (Ebert *et al.* 2008; Akhtar *et al.* 2018). Although the use of hPSC-derived interneuron progenitors in regenerative medicine has not reached clinical use, current findings in the literature have demonstrated the promise of such a use in the future, with hPSC-derived interneuronal grafts being a potential therapeutic strategy for refractory epilepsies and traumatic brain injury (Cunningham *et al.* 2014; Tornero *et al.* 2017). With the help of tools that allow the identification and manipulation of circuits formed between host and graft, such as viral tracer systems, a better evaluation of graft efficacy and potential side-effects at the behavioural level can be obtained in rodent models. Thus, helping to guide improvements to aid progress in making hPSC-derived interneuronal grafts a viable therapeutic strategy (Figure 7.4).

Figure 7.4

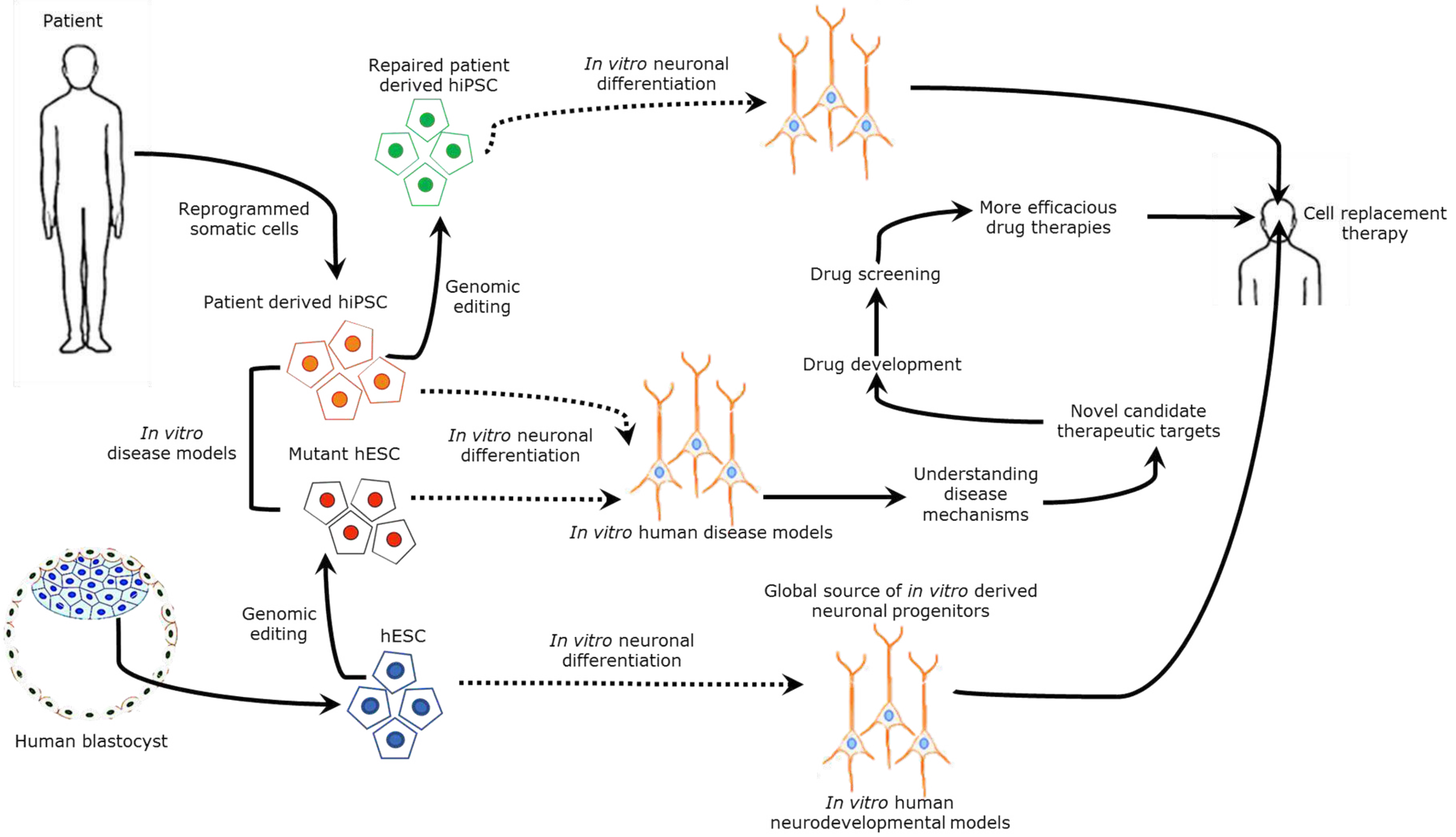

Patient
Reprogrammed somatic cells
Repaired patient derived hiPSC
In vitro neuronal differentiation
Patient-specific source of *in vitro* derived neuronal progenitors
Genomic editing
Patient derived hiPSC
In vitro disease models
Mutant hESC
In vitro neuronal differentiation
More efficacious drug therapies
Cell replacement therapy
Drug screening
Drug development
Novel candidate therapeutic targets
Understanding disease mechanisms
In vitro human disease models
Human blastocyst
Genomic editing
hESC
In vitro neuronal differentiation
Global source of *in vitro* derived neuronal progenitors
In vitro human neurodevelopmental models

Figure 7.4 The application of hPSC-derived neurons.

iPSC are derived from somatic cells that have been reprogrammed to revert back to pluripotent state. Patient derived iPSC can be differentiated in vitro into various neuronal lineages to understand the neurodevelopmental potential of patient derived neuronal progenitors, and the functional deficits in the neurons generated. This can lead to the understanding of disease mechanisms, and in turn flag potential therapeutic targets. Refer to Shao *et al.* for example with schizophrenia patient derived iPSC lines (Shao *et al.* 2019). Alternatively, hiPSC lines can be generated for the investigation into monogenetic diseases, where the patient derived hiPSC can be genetically edited to the non-pathogenic (healthy control) variant. Please see Xu *et al.* for example on Huntington's disease patient derived iPSC line (Xu *et al.* 2017). The patient derived hiPSC are now referred to as "repaired", and can be directed in vitro into neuronal committed progenitors that can be used for personalised cell replacement therapies. hESC are derived from human blastocyst inner mass cells and are not used for patient-specific in vitro disease models, but can be genomically modified to mimic the disease state. These cell lines are referred to as "edited" hESC, Li *et al.* presented an example of hESC-derived Rett syndrome model (Li *et al.* 2013). Together hPSC-derived neuronal models provide a global in vitro model of neurodevelopment, and a promising source of neuronal progenitors for regenerative medicine. Cartoon illustrations of human figure, blastocyst and neurons were collected from https://www.motifolio.com (Motifolio 2018).

hESC, human embryonic stem cells. iPSC, induced pluripotent stem cells. hPSC, human pluripotent stem cells.

7.4 Concluding remarks

In the research presented in this thesis complementary approaches have been taken to characterise both the molecular and functional basis of interneuron differentiation trajectory using hPSC-derived interneurons. The work presented has both answered and raised new questions, which will act to guide future experiments. Ultimately, by combining inter-disciplinary and inter-species studies, the multi-faceted features of this neuronal class can be teased apart, which will be of great benefit to both research and regenerative medicine.

References

Agilent Technologies. (2017). *Agilent High Sensitivity DNA Kit Guide.* Available at: https://www.agilent.com/cs/library/usermanuals/public/G2938-90322_HighSensitivityDNAKit_QSG.pdf [Accessed:25.09.2019].

Akhtar, A. A., Gowing, G., Kobritz, N., … Svendsen, C. N. (2018). Inducible Expression of GDNF in Transplanted iPSC-Derived Neural Progenitor Cells. *Stem Cell Reports*, **10**(6), 1696–1704.

Al-Jaberi, N., Lindsay, S., Sarma, S., Bayatti, N., & Clowry, G. J. (2015). The Early Fetal Development of Human Neocortical GABAergic Interneurons. *Cerebral Cortex*, **25**(3), 631–645.

Anderson, N., Chen, C., & Grabel, L. P. P. of I. I. in V. and in V. from P. S. C. (2016). Hedgehog Promotes Production of Inhibitory Interneurons in Vivo and in Vitro from Pluripotent Stem Cells. *Journal of Developmental Biology*, **4**(3), 26.

Anton, T., Karg, E., & Bultmann, S. (2018). Applications of the CRISPR/Cas system beyond gene editing. *Biology Methods and Protocols*, **3**(1). doi:10.1093/biomethods/bpy002

Arber, C., & Li, M. (2013). Cortical interneurons from human pluripotent stem cells: prospects for neurological and psychiatric disease. *Frontiers in Cellular Neuroscience*, **7**(March), 10.

Arber, C., Precious, S. V., Cambray, S., … Li, M. (2015). Activin A directs striatal projection neuron differentiation of human pluripotent stem cells. *Development*, **142**(7), 1375–1386.

Arendt, D., Musser, J. M., Baker, C. V. H., … Wagner, G. P. (2016). The origin and evolution of cell types. *Nature Reviews Genetics*, **17**(12), 744–757.

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**(43), 16707–12.

Azizi, E., Carr, A. J., Plitas, G., … Pe’er, D. (2018). Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell*. doi:10.1016/j.cell.2018.05.060

Backman, M., Machon, O., Mygland, L., … Krauss, S. (2005). Effects of canonical Wnt signaling on dorso-ventral specification of the mouse telencephalon. *Developmental Biology*, **279**(1), 155–168.

Bagur, R., & Hajnóczky, G. (2017). Intracellular Ca 2+ Sensing: Its Role in Calcium Homeostasis and Signaling. *Molecular Cell*, **66**(6), 780–788.

Bain, G., Kitchens, D., Yao, M., Huettner, J. E., & Gottlieb, D. I. (1995). Embryonic Stem Cells Express Neuronal Properties in Vitro. *Developmental Biology*, **168**(2), 342–357.

Bakken, T. E., Miller, J. A., Ding, S.-L., … Lein, E. S. (2016). A comprehensive transcriptional map of primate brain development. *Nature*, **535**(7612), 367–375.

Bandler, R. C., Mayer, C., & Fishell, G. (2017). Cortical interneuron specification: the juncture of genes, time and geometry. *Current Opinion in Neurobiology*, **42**, 17–24.

Bardy, C., van den Hurk, M., Kakaradov, B., … Gage, F. H. (2016). Predicting the functional states of human iPSC-derived neurons with single-cell RNA-seq and electrophysiology. *Molecular Psychiatry*, **21**(11), 1573–1588.

Barker, R. A., Parmar, M., Studer, L., & Takahashi, J. (2017). Human Trials of Stem Cell-Derived Dopamine Neurons for Parkinson’s Disease: Dawn of a New Era. *Cell Stem Cell*, **21**(5), 569–573.

Batista-Brito, R., Machold, R., Klein, C., & Fishell, G. (2008). Gene Expression in Cortical Interneuron Precursors is Prescient of their Mature Function. *Cerebral Cortex*, **18**(10), 2306–2317.

Bele, S., Kiessling, M., & Gass, P. (1995). Embryonic cortical neurons differentiate into various types of interneurons when heterotopically transplanted into the adult rat brain. *Brain Research*, **704**(2), 210–217.

Biddy, B. A., Kong, W., Kamimoto, K., … Morris, S. A. (2018). Single-cell mapping of lineage and identity in direct reprogramming. *Nature*, **564**(7735), 219–224.

Boldog, E., Bakken, T. E., Hodge, R. D., … Tamás, G. (2018). Transcriptomic and morphophysiological evidence for a specialized human cortical GABAergic cell type. *Nature Neuroscience*, **21**(9), 1185–1195.

Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G., & Deisseroth, K. (2005). Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, **8**(9), 1263–1268.

Brandão, J. A., & Romcy-Pereira, R. N. (2015). Interplay of environmental signals and progenitor diversity on fate specification of cortical GABAergic neurons. *Frontiers in Cellular Neuroscience*, **9**. doi:10.3389/fncel.2015.00149

Brennand, K. J., Simone, A., Jou, J., … Gage, F. H. (2011). Modelling schizophrenia using human induced pluripotent stem cells. *Nature*, **473**(7346), 221–225.

Brinkkoetter, P. T., Pippin, J. W., & Shankland, S. J. (2010). Cyclin I-Cdk5 governs survival in post-mitotic cells. *Cell Cycle*, **9**(9), 1729–1731.

Butt, S. J. B., Fuccillo, M., Nery, S., … Fishell, G. (2005). The Temporal and Spatial Origins of Cortical Interneurons Predict Their Physiological Subtype. *Neuron*, **48**(4), 591–604.

Butt, S. J. B., Sousa, V. H., Fuccillo, M. V., … Fishell, G. (2008). The Requirement of Nkx2-1 in the Temporal Specification of Cortical Interneuron Subtypes. *Neuron*, **59**(5), 722–732.

Cadwell, C. R., Palasantza, A., Jiang, X., … Tolias, A. S. (2016). Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. *Nature Biotechnology*, **34**(2), 199–203.

Cadwell, C. R., Scala, F., Li, S., … Tolias, A. S. (2017). Multimodal profiling of single-cell morphology, electrophysiology, and gene expression using Patch-seq. *Nature Protocols*, **12**(12), 2531–2553.

Calvigioni, D., Máté, Z., Fuzik, J., … Harkany, T. (2016). Functional Differentiation of Cholecystokinin-Containing Interneurons Destined for the Cerebral Cortex. *Cerebral Cortex*, bhw094.

Cambray, S., Arber, C., Little, G., … Rodríguez, T. a. (2012). Activin induces cortical interneuron identity and differentiation in embryonic stem cell-derived telencephalic neural precursors. *Nature Communications*, **3**(1), 841.

Camp, J. G., Badsha, F., Florio, M., … Treutlein, B. (2015). Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(51), 15672–7.

Carpenter, A. E., Jones, T. R., Lamprecht, M. R., … Sabatini, D. M. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology*, **7**(10), R100.

Carri, A. D., Onorati, M., Lelos, M. J., … Cattaneo, E. (2013). Developmentally coordinated extrinsic signals drive human pluripotent stem cell differentiation toward authentic DARPP-32+ medium-sized spiny neurons. *Development*, **140**(2), 301–312.

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

Chatzi, C., Brade, T., & Duester, G. (2011). Retinoic Acid Functions as a Key GABAergic Differentiation Signal in the Basal Ganglia. *PLoS Biology*, **9**(4), e1000609.

Chavez, A., Scheiman, J., Vora, S., … Church, G. M. (2015). Highly efficient Cas9-mediated transcriptional programming. *Nature Methods*, **12**(4), 326–328.

Chen, Y.-J. J., Friedman, B. A., Ha, C., … Modrusan, Z. (2017). Single-cell RNA sequencing identifies distinct mouse medial ganglionic eminence cell types. *Scientific Reports*, **7**(1), 45656.

Chen, Y.-J. J., Vogt, D., Wang, Y., … Rubenstein, J. L. R. (2013). Use of “MGE Enhancers” for Labeling and Selection of Embryonic Stem Cell-Derived Medial Ganglionic Eminence (MGE) Progenitors and Neurons. *PLoS ONE*, **8**(5), e61956.

Chung, H. K., Jacobs, C. L., Huo, Y., … Lin, M. Z. (2015). Tunable and reversible drug control of protein production via a self-excising degron. *Nature Chemical Biology*, **11**(9), 713–720.

Clontech Laboratories. (2018). *SMART-Seq v4 Ultra Low Input RNA Kit for the Fluidigm C1 System, IFCs User Manual.* Available at: https://www.takarabio.com/products/next-generation-sequencing/single-cell-rna-and-dna-seq/smart-seq-v4-for-mrna-seq-with-the-fluidigm-c1 [Accessed:25.09.2019].

Close, J. L., Yao, Z., Levi, B. P., … Lein, E. (2017). Single-Cell Profiling of an In Vitro Model of Human Interneuron Development Reveals Temporal Dynamics of Cell Type Production and Maturation. *Neuron*, **93**(5), 1035-1048.e5.

Colasante, G., Lignani, G., Rubio, A., … Broccoli, V. (2015). Rapid Conversion of Fibroblasts into Functional Forebrain GABAergic Interneurons by Direct Genetic Reprogramming. *Cell Stem Cell*, **17**(6), 719–734.

Cole, A. E., Murray, S. S., & Xiao, J. (2016). Bone Morphogenetic Protein 4 Signalling in Neural Stem and Progenitor Cells during Development and after Injury. *Stem Cells International*, **2016**, 1–16.

Cunningham, M., Cho, J.-H., Leung, A., … Chung, S. (2014). hPSC-Derived Maturing GABAergic Interneurons Ameliorate Seizures and Abnormal Behavior in Epileptic Mice. *Cell Stem Cell*, **15**(5), 559–573.

Danjo, T., Eiraku, M., Muguruma, K., … 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**(5), 1919–1933.

Darmanis, S., Sloan, S. A., Zhang, Y., … Quake, S. R. (2015). A survey of human brain transcriptome diversity at the single cell level. *Proceedings of the National Academy of Sciences*, **112**(23), 7285–7290.

de Bruijn, D. R. H., Kater-Baats, E., Eleveld, M., Merkx, G., & Geurts van Kessel, A. (2001). Mapping and characterization of the mouse and human SS18 genes, two human SS18-like genes and a mouse Ss18 pseudogene. *Cytogenetic and Genome Research*, **92**(3–4), 310–319.

DeFelipe, J., López-Cruz, P. L., Benavides-Piccione, R., … Ascoli, G. a. (2013). New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nature Reviews. Neuroscience*, **14**(3), 202–16.

Dehorter, N., Ciceri, G., Bartolini, G., Lim, L., del Pino, I., & Marín, O. (2015). Tuning of fast-spiking interneuron properties by an activity-dependent transcriptional switch. *Science (New York, N.Y.)*, **349**(6253), 1216–20.

Dobin, A., Davis, C. A., Schlesinger, F., … Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**(1), 15–21.

Doerr, J., Schwarz, M. K., Wiedermann, D., … Brüstle, O. (2017). Whole-brain 3D mapping of human neural transplant innervation. *Nature Communications*, **8**, 14162.

Drissen, R., Buza-Vidas, N., Woll, P., … Nerlov, C. (2016). Distinct myeloid progenitor–differentiation pathways identified through single-cell RNA sequencing. *Nature Immunology*, **17**(6), 666–676.

Ebert, A. D., Beres, A. J., Barber, A. E., & Svendsen, C. N. (2008). Human neural progenitor cells over-expressing IGF-1 protect dopamine neurons and restore function in a rat model of Parkinson’s disease. *Experimental Neurology*, **209**(1), 213–223.

Elbert, A., Vogt, D., Watson, A., … Bérubé, N. G. (2019). CTCF Governs the Identity and Migration of MGE-Derived Cortical Interneurons. *The Journal of Neuroscience*, **39**(1), 177–192.

Elias, L. A. B., Potter, G. B., & Kriegstein, A. R. (2008). A Time and a Place for Nkx2-1 in Interneuron Specification and Migration. *Neuron*, **59**(5), 679–682.

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

Fan, X., Dong, J., Zhong, S., … Tang, F. (2018). Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis. *Cell Research*, **28**(7), 730–745.

Fantuzzo, J. A., De Filippis, L., McGowan, H., … Pang, Z. P. (2017). μNeurocircuitry: Establishing in vitro models of neurocircuits with human neurons. *TECHNOLOGY*, **05**(02), 87–97.

Fink, C. C., Bayer, K.-U., Myers, J. W., Ferrell, J. E., Schulman, H., & Meyer, T. (2003). Selective Regulation of Neurite Extension and Synapse Formation by the β but not the α Isoform of CaMKII. *Neuron*, **39**(2), 283–297.

Flames, N., & Marín, O. (2005). Developmental Mechanisms Underlying the Generation of Cortical Interneuron Diversity. *Neuron*, **46**(3), 377–381.

Flames, N., Pla, R., Gelman, D. M., Rubenstein, J. L. R., Puelles, L., & Marin, O. (2007). Delineation of Multiple Subpallial Progenitor Domains by the Combinatorial Expression of Transcriptional Codes. *Journal of Neuroscience*, **27**(36), 9682–9695.

Fluidigm. (2015). *Gene expression with the 96.96 IFC using delta gene assays on preamplified samples.* Available at: https://www.fluidigm.com/binaries/content/documents/fluidigm/resources/96.96-ge-delta-gene-qr-100-9792/96.96-ge-delta-gene-qr-100-9792/fluidigm%3Afile [Accessed:25.09.2019].

Fluidigm. (2018a). *Biomark HD, Biomark & EP1 Software , Fluidigm real-time PCR analysis software*. Available at: https://www.fluidigm.com/software. [Accessed:25.09.2019].

Fluidigm. (2018b). *C1, Singular Analysis Toolset Software.* Available at: https://www.fluidigm.com/software [Accessed:25.09.2019].

Földy, C., Darmanis, S., Aoto, J., Malenka, R. C., Quake, S. R., & Südhof, T. C. (2016). Single-cell RNAseq reveals cell adhesion molecule profiles in electrophysiologically defined neurons. *Proceedings of the National Academy of Sciences*, **113**(35), E5222–E5231.

Frazer, S., Prados, J., Niquille, M., … Dayer, A. (2017). Transcriptomic and anatomic parcellation of 5-HT3AR expressing cortical interneuron subtypes revealed by single-cell RNA sequencing. *Nature Communications*, **8**, 14219.

Fuccillo, M., Joyner, A. L., & Fishell, G. (2006). Morphogen to mitogen: the multiple roles of hedgehog signalling in vertebrate neural development. *Nature Reviews Neuroscience*, **7**(10), 772–783.

Fuzik, J., Zeisel, A., Máté, Z., … Harkany, T. (2015). Integration of electrophysiological recordings with single-cell RNA-seq data identifies neuronal subtypes. *Nature Biotechnology*, **34**(2), 175–183.

Galan, J.-M., & Peter, M. (1999). Ubiquitin-dependent degradation of multiple F-box proteins by an autocatalytic mechanism. *Proceedings of the National Academy of Sciences*, **96**(16), 9124–9129.

Garas, F. N., Kormann, E., Shah, R. S., … Sharott, A. (2018). Structural and molecular heterogeneity of calretinin-expressing interneurons in the rodent and primate striatum. *Journal of Comparative Neurology*, **526**(5), 877–898.

Gilani, A. I., Chohan, M. O., Inan, M., … Moore, H. (2014). Interneuron precursor transplants in adult hippocampus reverse psychosis-relevant features in a mouse model of hippocampal disinhibition. *Proceedings of the National Academy of Sciences*, **111**(20), 7450–7455.

Ginger, M., Haberl, M., Conzelmann, K.-K., Schwarz, M. K., & Frick, A. (2013). Revealing the secrets of neuronal circuits with recombinant rabies virus technology. *Frontiers in Neural Circuits*, **7**(January), 2.

Glickstein, S. B., Moore, H., Slowinska, B., … Ross, M. E. (2007). Selective cortical interneuron and GABA deficits in cyclin D2-null mice. *Development*, **134**(22), 4083–4093.

Gokce, O., Stanley, G. M., Treutlein, B., … Quake, S. R. (2016). Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq. *Cell Reports*, **16**(4), 1126–1137.

Goulburn, A. L., Alden, D., Davis, R. P., … Elefanty, A. G. (2011). A Targeted NKX2.1 Human Embryonic Stem Cell Reporter Line Enables Identification of Human Basal Forebrain Derivatives. *STEM CELLS*, **29**(3), 462–473.

GraphPad Software. (1995-2017). *Graphpad Prism 5*. Available at: www.graphpad.com [Accessed:25.09.2019].

Grealish, S., Heuer, A., Cardoso, T., … Parmar, M. (2015). Monosynaptic Tracing using Modified Rabies Virus Reveals Early and Extensive Circuit Integration of Human Embryonic Stem Cell-Derived Neurons. *Stem Cell Reports*, **4**(6), 975–983.

Grimaldi, P., Carletti, B., Magrassi, L., & Rossi, F. (2005). Fate restriction and developmental potential of cerebellar progenitors. *In Progress in Brain Research,* pp. 57–68.

GSL Biotech. (2019). *SnapGene.* Available at: https://www.snapgene.com [Accessed:25.09.2019].

Haenschel, C., Bittner, R. A., Waltz, J., … Rodriguez, E. (2009). Cortical Oscillatory Activity Is Critical for Working Memory as Revealed by Deficits in Early-Onset Schizophrenia. *Journal of Neuroscience*, **29**(30), 9481–9489.

Hansen, D. V, Lui, J. H., Flandin, P., … Kriegstein, A. R. (2013). Non-epithelial stem cells and cortical interneuron production in the human ganglionic eminences. *Nature Neuroscience*, **16**(11), 1576–1587.

Hashimshony, T., Wagner, F., Sher, N., & Yanai, I. (2012). CEL-Seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cell Reports*, **2**(3), 666–673.

Hedrich, U. B. S., Liautard, C., Kirschenbaum, D., … Lerche, H. (2014). Impaired Action Potential Initiation in GABAergic Interneurons Causes Hyperexcitable Networks in an Epileptic Mouse Model Carrying a Human NaV1.1 Mutation. *Journal of Neuroscience*, **34**(45), 14874–14889.

Heyes, S., Pratt, W. S., Rees, E., … Dolphin, A. C. (2015). Genetic disruption of voltage-gated calcium channels in psychiatric and neurological disorders. *Progress in Neurobiology*, **134**, 36–54.

Hirano, S., Hirano, Y., Maekawa, T., … Onitsuka, T. (2008). Abnormal Neural Oscillatory Activity to Speech Sounds in Schizophrenia: A Magnetoencephalography Study. *Journal of Neuroscience*, **28**(19), 4897–4903.

Hu, J. S., Vogt, D., Lindtner, S., Sandberg, M., Silberberg, S. N., & Rubenstein, J. L. R. (2017a). Coup-TF1 and Coup-TF2 control subtype and laminar identity of MGE-derived neocortical interneurons. *Development*, **144**(15), 2837–2851.

Hu, J. S., Vogt, D., Sandberg, M., & Rubenstein, J. L. (2017b). Cortical interneuron development: a tale of time and space. *Development*, **144**(21), 3867–3878.

Huard, J. M., Forster, C. C., Carter, M. L., Sicinski, P., & Ross, M. E. (1999). Cerebellar histogenesis is disturbed in mice lacking cyclin D2. *Development (Cambridge, England)*, **126**(9), 1927–35.

Inan, M., Welagen, J., & Anderson, S. A. (2012). Spatial and Temporal Bias in the Mitotic Origins of Somatostatin- and Parvalbumin-Expressing Interneuron Subgroups and the Chandelier Subtype in the Medial Ganglionic Eminence. *Cerebral Cortex*, **22**(4), 820–827.

Invitrogen. (2008). *Quant-iT PicoGreen dsDNA Reagent and Kits.* Available at: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/mp07581.pdf [Accessed:25.09.2019].

Invitrogen. (2013). *Lipofectamine 2000 DNA Transfection Reagent Protocol.* Available at: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/Lipofectamine_2000_Reag_protocol.pdf?icid=cvc-lipofectamine-c1m2 [Accessed:25.09.2019].

Invitrogen. (2014). *Lipofectamine 3000 Reagent Protocol.* Available at: https://www.thermofisher.com/content/dam/LifeTech/Documents/PDFs/lipofectamine3000_protocol.pdf [Accessed:25.09.2019].

Invitrogen. (2015). *Qubit dsDNA HS assay kits.* Available at: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/Qubit_dsDNA_HS_Assay_UG.pdf [Accessed:25.09.2019].

Ishino, Y., Yetman, M. J., Sossi, S. M., Steinecke, A., Hayano, Y., & Taniguchi, H. (2017). Regional Cellular Environment Shapes Phenotypic Variations of Hippocampal and Neocortical Chandelier Cells. *The Journal of Neuroscience*, **37**(41), 9901–9916.

Ito-Ishida, A., Ure, K., Chen, H., Swann, J. W., & Zoghbi, H. Y. (2015). Loss of MeCP2 in Parvalbumin-and Somatostatin-Expressing Neurons in Mice Leads to Distinct Rett Syndrome-like Phenotypes. *Neuron*, **88**(4), 651–658.

Jiang, M., & Swann, J. W. (2005). A role for L-type calcium channels in the maturation of parvalbumin-containing hippocampal interneurons. *Neuroscience*, **135**(3), 839–850.

Kang, H. J., Kawasawa, Y. I., Cheng, F., … Šestan, N. (2011). Spatio-temporal transcriptome of the human brain. *Nature*, **478**(7370), 483–489.

Kataoka, Y., Kalanithi, P. S. A., Grantz, H., … Vaccarino, F. M. (2010). Decreased number of parvalbumin and cholinergic interneurons in the striatum of individuals with Tourette syndrome. *The Journal of Comparative Neurology*, **518**(3), 277–291.

Kearns, N. A., Genga, R. M. J., Enuameh, M. S., Garber, M., Wolfe, S. A., & Maehr, R. (2014). Cas9 effector-mediated regulation of transcription and differentiation in human pluripotent stem cells. *Development*, **141**(1), 219–223.

Kelava, I., & Lancaster, M. A. (2016). Stem Cell Models of Human Brain Development. *Cell Stem Cell*, **18**(6), 736–48.

Kemp, P. J., Rushton, D. J., Yarova, P. L., … Telezhkin, V. (2016). Improving and accelerating the differentiation and functional maturation of human stem cell-derived neurons: role of extracellular calcium and GABA. *The Journal of Physiology*, **594**(22), 6583–6594.

Kepecs, A., & Fishell, G. (2014). Interneuron cell types are fit to function. *Nature*, **505**(7483), 318–326.

Kessaris, N., Magno, L., Rubin, A. N., & Oliveira, M. G. (2014). Genetic programs controlling cortical interneuron fate. *Current Opinion in Neurobiology*, **26**, 79–87.

Kester, L., & van Oudenaarden, A. (2018). Single-Cell Transcriptomics Meets Lineage Tracing. *Cell Stem Cell*. doi:10.1016/j.stem.2018.04.014

Kim, T.-G., Yao, R., Monnell, T., … Chung, S. (2014). Efficient Specification of Interneurons from Human Pluripotent Stem Cells by Dorsoventral and Rostrocaudal Modulation. *STEM CELLS*, **32**(7), 1789–1804.

Koornstra, G., Wattel, E., & Exalto, N. (1990). Crown-rump length measurements revisited. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, **35**(2–3), 131–138.

Lake, B. B., Ai, R., Kaeser, G. E., … Zhang, K. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science*, **352**(6293), 1586–1590.

Lanciego, J. L., & Wouterlood, F. G. (2011). A half century of experimental neuroanatomical tracing. *Journal of Chemical Neuroanatomy*, **42**(3), 157–183.

Lee, A. T., Cunniff, M. M., See, J. Z., … Sohal, V. S. (2019). VIP Interneurons Contribute to Avoidance Behavior by Regulating Information Flow across Hippocampal-Prefrontal Networks. *Neuron*. doi:10.1016/j.neuron.2019.04.001

Lewis, D. A., Hashimoto, T., & Volk, D. W. (2005). Cortical inhibitory neurons and schizophrenia. *Nature Reviews Neuroscience*, **6**(4), 312–324.

Li, X.-J., Zhang, X., Johnson, M. A., Wang, Z.-B., LaVaute, T., & Zhang, S.-C. (2009). Coordination of sonic hedgehog and Wnt signaling determines ventral and dorsal telencephalic neuron types from human embryonic stem cells. *Development*, **136**(23), 4055–4063.

Li, Y., Wang, H., Muffat, J., … Jaenisch, R. (2013). Global Transcriptional and Translational Repression in Human-Embryonic-Stem-Cell-Derived Rett Syndrome Neurons. *Cell Stem Cell*, **13**(4), 446–458.

Lima Caldeira, G., Peça, J., & Carvalho, A. L. (2019). New insights on synaptic dysfunction in neuropsychiatric disorders. *Current Opinion in Neurobiology*, **57**, 62–70.

Lin, L.-C., & Sibille, E. (2013). Reduced brain somatostatin in mood disorders: a common pathophysiological substrate and drug target? *Frontiers in Pharmacology*, **4**. doi:10.3389/fphar.2013.00110

Liu, S., & Trapnell, C. (2016). Single-cell transcriptome sequencing: recent advances and remaining challenges. *F1000Research*, **5**, 182.

Liu, Y., Liu, H., Sauvey, C., Yao, L., Zarnowska, E. D., & Zhang, S.-C. (2013). Directed differentiation of forebrain GABA interneurons from human pluripotent stem cells. *Nature Protocols*, **8**(9), 1670–1679.

Loo, L., Simon, J. M., Xing, L., … Zylka, M. J. (2019). Single-cell transcriptomic analysis of mouse neocortical development. *Nature Communications*, **10**(1), 134.

Loughna, P., Chitty, L., Evans, T., & Chudleigh, T. (2009). Fetal Size and Dating: Charts Recommended for Clinical Obstetric Practice. *Ultrasound*, **17**(3), 160–166.

Marín, O. (2012). Interneuron dysfunction in psychiatric disorders. *Nature Reviews. Neuroscience*, **13**(2), 107–20.

Marín, O., Anderson, S. A., & Rubenstein, J. L. (2000). Origin and molecular specification of striatal interneurons. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, **20**(16), 6063–76.

Maroof, A. M., Keros, S., Tyson, J. A., … Studer, L. (2013). Directed Differentiation and Functional Maturation of Cortical Interneurons from Human Embryonic Stem Cells. *Stem Cell*, **12**(5), 559–572.

Martina, M., Schultz, J. H., Ehmke, H., Monyer, H., & Jonas, P. (1998). Functional and Molecular Differences between Voltage-Gated K + Channels of Fast-Spiking Interneurons and Pyramidal Neurons of Rat Hippocampus. *The Journal of Neuroscience*, **18**(20), 8111–8125.

Mayer, C., Hafemeister, C., Bandler, R. C., … Satija, R. (2018). Developmental diversification of cortical inhibitory interneurons. *Nature*. doi:10.1038/nature25999

Mayer, S., Chen, J., Velmeshev, D., … Kriegstein, A. R. (2019). Multimodal Single-Cell Analysis Reveals Physiological Maturation in the Developing Human Neocortex. *Neuron*. doi:10.1016/j.neuron.2019.01.027

Meganathan, K., Lewis, E. M. A., Gontarz, P., … Kroll, K. L. (2017). Regulatory networks specifying cortical interneurons from human embryonic stem cells reveal roles for CHD2 in interneuron development. *Proceedings of the National Academy of Sciences*, **114**(52), E11180–E11189.

Mi, D., Li, Z., Lim, L., … Marín, O. (2018). Early emergence of cortical interneuron diversity in the mouse embryo. *Science*, **360**(6384), 81–85.

Mickelsen, L. E., Bolisetty, M., Chimileski, B. R., … Jackson, A. C. (2019). Single-cell transcriptomic analysis of the lateral hypothalamic area reveals molecularly distinct populations of inhibitory and excitatory neurons. *Nature Neuroscience*. doi:10.1038/s41593-019-0349-8

Mitsuhashi, T., Aoki, Y., Eksioglu, Y. Z., … Caviness, V. S. (2001). Overexpression of p27Kip1 lengthens the G1 phase in a mouse model that targets inducible gene expression to central nervous system progenitor cells. *Proceedings of the National Academy of Sciences*, **98**(11), 6435–6440.

Miyoshi, G., Hjerling-Leffler, J., Karayannis, T., … Fishell, G. (2010). Genetic Fate Mapping Reveals That the Caudal Ganglionic Eminence Produces a Large and Diverse Population of Superficial Cortical Interneurons. *Journal of Neuroscience*, **30**(5), 1582–1594.

Miyoshi, G., Young, A., Petros, T., … Fishell, G. (2015). Prox1 Regulates the Subtype-Specific Development of Caudal Ganglionic Eminence-Derived GABAergic Cortical Interneurons. *Journal of Neuroscience*, **35**(37), 12869–12889.

Mo, Z., Moore, A. R., Filipovic, R., … Zecevic, N. (2007). Human Cortical Neurons Originate from Radial Glia and Neuron-Restricted Progenitors. *Journal of Neuroscience*, **27**(15), 4132–4145.

Moignard, V., & Göttgens, B. (2016). Dissecting stem cell differentiation using single cell expression profiling. *Current Opinion in Cell Biology*, **43**, 78–86.

Molecular Devices. (2019). *Clampfit 10.6.* Available at: http://mdc.custhelp.com/app/answers/detail/a_id/18779/~/axon%E2%84%A2pclamp%E2%84%A2-10-electrophysiology-data-acquisition-%26-analysis-software-download [Accessed:25.09.2019].

Moore, A. R., Filipovic, R., Mo, Z., Rasband, M. N., Zecevic, N., & Antic, S. D. (2009). Electrical Excitability of Early Neurons in the Human Cerebral Cortex during the Second Trimester of Gestation. *Cerebral Cortex*, **19**(8), 1795–1805.

Moore, A. R., Zhou, W.-L., Jakovcevski, I., Zecevic, N., & Antic, S. D. (2011). Spontaneous Electrical Activity in the Human Fetal Cortex In Vitro. *Journal of Neuroscience*, **31**(7), 2391–2398.

Motifolio. (2018). *Scientific Illustration Toolkits.* Available at: https://www.motifolio.com/index.html. [Accessed:25.09.2019].

Mukhopadhyay, A., McGuire, T., Peng, C.-Y., & Kessler, J. A. (2009). Differential effects of BMP signaling on parvalbumin and somatostatin interneuron differentiation. *Development*, **136**(15), 2633–2642.

Muñoz-Manchado, A. B., Bengtsson Gonzales, C., Zeisel, A., … Hjerling-Leffler, J. (2018). Diversity of Interneurons in the Dorsal Striatum Revealed by Single-Cell RNA Sequencing and PatchSeq. *Cell Reports*, **24**(8), 2179-2190.e7.

Nagel, G., Szellas, T., Huhn, W., … Bamberg, E. (2003). Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proceedings of the National Academy of Sciences*, **100**(24), 13940–13945.

National Center for Biotechnology Information. (2019). *Primer-BLAST.* Available at: https://www.ncbi.nlm.nih.gov/tools/primer-blast/ [Accessed:25.09.2019].

Nicholas, C. R., Chen, J., Tang, Y., … Kriegstein, A. R. (2013). Functional Maturation of hPSC-Derived Forebrain Interneurons Requires an Extended Timeline and Mimics Human Neural Development. *Cell Stem Cell*, **12**(5), 573–586.

Nicoleau, C., Varela, C., Bonnefond, C., … 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**(9), 1763–1774.

Noakes, Z., Keefe, F., Tamburini, C., … Li, M. (2019). Human Pluripotent Stem Cell-Derived Striatal Interneurons: Differentiation and Maturation In Vitro and in the Rat Brain. *Stem Cell Reports*. doi:10.1016/j.stemcr.2018.12.014

Nóbrega-Pereira, S., Kessaris, N., Du, T., Kimura, S., Anderson, S. A., & Marín, O. (2008). Postmitotic Nkx2-1 Controls the Migration of Telencephalic Interneurons by Direct Repression of Guidance Receptors. *Neuron*, **59**(5), 733–745.

Nowakowski, T. J., Bhaduri, A., Pollen, A. A., … Kriegstein, A. R. (2017). Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science*, **358**(6368), 1318–1323.

Ofengeim, D., Giagtzoglou, N., Huh, D., Zou, C., & Yuan, J. (2017). Single-Cell RNA Sequencing: Unraveling the Brain One Cell at a Time. *Trends in Molecular Medicine*, **23**(6), 563–576.

Okabe, S., Forsberg-Nilsson, K., Spiro, A. C., Segal, M., & McKay, R. D. G. (1996). Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro. *Mechanisms of Development*, **59**(1), 89–102.

Onorati, M., Castiglioni, V., Biasci, D., … Cattaneo, E. (2014). Molecular and functional definition of the developing human striatum. *Nature Neuroscience*, **17**(12), 1804–1815.

Osakada, F., & Callaway, E. M. (2013). Design and generation of recombinant rabies virus vectors. *Nature Protocols*, **8**(8), 1583–1601.

Osakada, F., Mori, T., Cetin, A. H., Marshel, J. H., Virgen, B., & Callaway, E. M. (2011). New rabies virus variants for monitoring and manipulating activity and gene expression in defined neural circuits. *Neuron*, **71**(4), 617–631.

Pan, G., Yang, J.-M., Hu, X.-Y., & Li, X.-M. (2016). Postnatal development of the electrophysiological properties of somatostatin interneurons in the anterior cingulate cortex of mice. *Scientific Reports*, **6**(1), 28137.

Patriarchi, T., Amabile, S., Frullanti, E., … Meloni, I. (2016). Imbalance of excitatory/inhibitory synaptic protein expression in iPSC-derived neurons from FOXG1 +/− patients and in foxg1 +/− mice. *European Journal of Human Genetics*, **24**(6), 871–880.

Poulin, J., Tasic, B., Hjerling-Leffler, J., Trimarchi, J. M., & Awatramani, R. (2016). Disentangling neural cell diversity using single-cell transcriptomics. *Nature Neuroscience*, **19**(9), 1131–41.

Powell, E. M., Campbell, D. B., Stanwood, G. D., Davis, C., Noebels, J. L., & Levitt, P. (2003). Genetic Disruption of Cortical Interneuron Development Causes Region- and GABA Cell Type-Specific Deficits, Epilepsy, and Behavioral Dysfunction. *The Journal of Neuroscience*, **23**(2), 622–631.

Qi, L. S., Larson, M. H., Gilbert, L. A., … Lim, W. A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell*, **152**(5), 1173–1183.

QIAGEN. (2012).*QIAGEN Plasmid Purification Handbook, QIAGEN Plasmid DNA Purification using QIAGEN Plasmid Midi and Maxi Kits.* Available at: https://www.scribd.com/document/319129565/QIAGEN-Plasmid-Purification-Handbook-April-2012 [Accessed:25.09.2019].

Quattrocolo, G., Fishell, G., & Petros, T. J. (2017). Heterotopic Transplantations Reveal Environmental Influences on Interneuron Diversity and Maturation. *Cell Reports*, **21**(3), 721–731.

R Core Team. (2013). *R: A language and environment for statistical computing, version 3.3.3.* Available at: http://www.R-project.org/ [Accessed:25.09.2019].

Rakić, S., Kanatani, S., Hunt, D., … Parnavelas, J. G. (2015). Cdk5 Phosphorylation of ErbB4 is Required for Tangential Migration of Cortical Interneurons. *Cerebral Cortex*, **25**(4), 991–1003.

Ramsköld, D., Luo, S., Wang, Y.-C., … Sandberg, R. (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nature Biotechnology*, **30**(8), 777–782.

Rasband, W.S. (1997-2018). *ImageJ.* Available at: https://imagej.nih.gov/ij/ [Accessed:25.09.2019].

Reinchisi, G., Ijichi, K., Glidden, N., Jakovcevski, I., & Zecevic, N. (2012). COUP-TFII Expressing Interneurons in Human Fetal Forebrain. *Cerebral Cortex*, **22**(12), 2820–2830.

Reynolds, G. P., Beasley, C. L., & Zhang, Z. J. (2002). Understanding the neurotransmitter pathology of schizophrenia: selective deficits of subtypes of cortical GABAergic neurons. *Journal of Neural Transmission*, **109**(5–6), 881–889.

Robbins, R. J., Brines, M. L., Kim, J. H., … Spencer, D. D. (1991). A selective loss of somatostatin in the hippocampus of patients with temporal lobe epilepsy. *Annals of Neurology*, **29**(3), 325–332.

Roche Molecular Systems. (2017). *EvoScript Universal cDNA Master, Instructions for Use - Version 04.* Available at: https://lifescience.roche.com/en_gb/products/evoscript-universal-cdna-master.html#documents [Accessed:25.09.2019].

Rolando, C., Gribaudo, S., Yoshikawa, K., Leto, K., De Marchis, S., & Rossi, F. (2010). Extracerebellar progenitors grafted to the neurogenic milieu of the postnatal rat cerebellum adapt to the host environment but fail to acquire cerebellar identities. *European Journal of Neuroscience*, **31**(8), 1340–1351.

Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, **2**(5), 255–267.

Rutter, L., Carver, F. W., Holroyd, T., … Coppola, R. (2009). Magnetoencephalographic gamma power reduction in patients with schizophrenia during resting condition. *Human Brain Mapping*, **30**(10), 3254–3264.

Sahakyan, A., Yang, Y., & Plath, K. (2018). The Role of Xist in X-Chromosome Dosage Compensation. *Trends in Cell Biology*, **28**(12), 999–1013.

Sarkar, A., Mei, A., Paquola, A. C. M., … Gage, F. H. (2018). Efficient Generation of CA3 Neurons from Human Pluripotent Stem Cells Enables Modeling of Hippocampal Connectivity In Vitro. *Cell Stem Cell*, **22**(5), 684-697.e9.

Shah, K., & Lahiri, D. K. (2017). A Tale of the Good and Bad: Remodeling of the Microtubule Network in the Brain by Cdk5. *Molecular Neurobiology*, **54**(3), 2255–2268.

Shao, Z., Noh, H., Bin Kim, W., … Chung, S. (2019). Dysregulated protocadherin-pathway activity as an intrinsic defect in induced pluripotent stem cell–derived cortical interneurons from subjects with schizophrenia. *Nature Neuroscience*, **22**(2), 229–242.

Sigma-Aldrich. (2019). *TRI Reagent Protocol.* Available at: https://www.sigmaaldrich.com/technical-documents/protocols/biology/tri-reagent.html [Accessed:25.09.2019].

Soh, H., Park, S., Ryan, K., Springer, K., Maheshwari, A., & Tzingounis, A. V. (2018). Deletion of KCNQ2/3 potassium channels from PV+ interneurons leads to homeostatic potentiation of excitatory transmission. *ELife*, **7**. doi:10.7554/eLife.38617

Sohal, V. S., Zhang, F., Yizhar, O., & Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature*, **459**(7247), 698–702.

Sousa, A. M. M., Zhu, Y., Raghanti, M. A., … Sestan, N. (2017). Molecular and cellular reorganization of neural circuits in the human lineage. *Science*, **358**(6366), 1027–1032.

Southwell, D. G., Paredes, M. F., Galvao, R. P., … Alvarez-Buylla, A. (2012). Intrinsically determined cell death of developing cortical interneurons. *Nature*, **491**(7422), 109–113.

Spira, M. E., & Hai, A. (2013). Multi-electrode array technologies for neuroscience and cardiology. *Nature Nanotechnology*, **8**(2), 83–94.

Südhof, T. C. (2008). Neuroligins and neurexins link synaptic function to cognitive disease. *Nature*, **455**(7215), 903–911.

Sun, A. X., Yuan, Q., Tan, S., … Je, H. S. (2016). Direct Induction and Functional Maturation of Forebrain GABAergic Neurons from Human Pluripotent Stem Cells. *Cell Reports*, **16**(7), 1942–1953.

Sun, Y., Gao, Y., Tidei, J. J., … Zhao, X. (2019). Loss of MeCP2 in immature neurons leads to impaired network integration. *Human Molecular Genetics*, **28**(2), 245–257.

Tang, F., Barbacioru, C., Wang, Y., … Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods*, **6**(5), 377–382.

Telezhkin, V., Schnell, C., Yarova, P., … Kemp, P. J. (2016). Forced cell cycle exit and modulation of GABA A , CREB, and GSK3β signaling promote functional maturation of induced pluripotent stem cell-derived neurons. *American Journal of Physiology-Cell Physiology*, **310**(7), C520–C541.

Telezhkin, V., Straccia, M., Yarova, P., … Kemp, P. J. (2018). Kv7 channels are upregulated during striatal neuron development and promote maturation of human iPSC-derived neurons. *Pflügers Archiv - European Journal of Physiology*, **470**(9), 1359–1376.

ThermoScientific. (2016). *Product Information, DNase I, RNase-free.* Available at: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0012000_DNase_I_RNasefree_1UuL_UG.pdf [Accessed:25.09.2019].

Tian, J., Huang, R., Cohen, J. Y., … Watabe-Uchida, M. (2016). Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons. *Neuron*, **91**(6), 1374–1389.

Tornero, D., Tsupykov, O., Granmo, M., … Kokaia, Z. (2017). Synaptic inputs from stroke-injured brain to grafted human stem cell-derived neurons activated by sensory stimuli. *Brain*, aww347.

Tripathy, S. J., Toker, L., Li, B., … Pavlidis, P. (2017). Transcriptomic correlates of neuron electrophysiological diversity. *PLOS Computational Biology*, **13**(10), e1005814.

Tyson, J. a, Goldberg, E. M., Maroof, A. M., Xu, Q., Petros, T. J., & Anderson, S. a. (2015). Duration of culture and sonic hedgehog signaling differentially specify PV versus SST cortical interneuron fates from embryonic stem cells. *Development (Cambridge, England)*, **142**(7), 1267–78.

van den Hurk, M., Erwin, J. A., Yeo, G. W., Gage, F. H., & Bardy, C. (2018). Patch-Seq Protocol to Analyze the Electrophysiology, Morphology and Transcriptome of Whole Single Neurons Derived From Human Pluripotent Stem Cells. *Frontiers in Molecular Neuroscience*, **11**. doi:10.3389/fnmol.2018.00261

Varga, C., Tamas, G., Barzo, P., Olah, S., & Somogyi, P. (2015). Molecular and Electrophysiological Characterization of GABAergic Interneurons Expressing the Transcription Factor COUP-TFII in the Adult Human Temporal Cortex. *Cerebral Cortex*, **25**(11), 4430–4449.

Virlogeux, A., Moutaux, E., Christaller, W., … Saudou, F. (2018). Reconstituting Corticostriatal Network on-a-Chip Reveals the Contribution of the Presynaptic Compartment to Huntington's Disease. *Cell Reports*, **22**(1), 110–122.

Vogt, D., Hunt, R. F., Mandal, S., … Rubenstein, J. L. R. (2014). Lhx6 Directly Regulates Arx and CXCR7 to Determine Cortical Interneuron Fate and Laminar Position. *Neuron*, **82**(2), 350–364.

Wall, N., DeLaParra, M., Callaway, E., & Kreitzer, A. (2013). Differential innervation of direct- and indirect-pathway striatal projection neurons. *Neuron*, **79**(2), 347–360.

Wamsley, B., & Fishell, G. (2017). Genetic and activity-dependent mechanisms underlying interneuron diversity. *Nature Reviews Neuroscience*, **18**(5), 299–309.

Wayman, G. A., Lee, Y.-S., Tokumitsu, H., Silva, A., & Soderling, T. R. (2008). Calmodulin-Kinases: Modulators of Neuronal Development and Plasticity. *Neuron*, **59**(6), 914–931.

Wertz, A., Trenholm, S., Yonehara, K., … Roska, B. (2015). Single-cell-initiated monosynaptic tracing reveals layer-specific cortical network modules. *Science*, **349**(6243), 70–74.

Wester, J. C., Mahadevan, V., Rhodes, C. T., … McBain, C. J. (2019). Neocortical Projection Neurons Instruct Inhibitory Interneuron Circuit Development in a Lineage-Dependent Manner. *Neuron*. doi:10.1016/j.neuron.2019.03.036

Wickersham, I. R., Lyon, D. C., Barnard, R. J. O., … Callaway, E. M. (2007). Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons. *Neuron*, **53**(5), 639–647.

Wills, Q. F., Livak, K. J., Tipping, A. J., … Holmes, C. (2013). Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments. *Nature Biotechnology*, **31**(8), 748–52.

Wolff, S. B. E., Gründemann, J., Tovote, P., … Lüthi, A. (2014). Amygdala interneuron subtypes control fear learning through disinhibition. *Nature*, **509**(7501), 453–458.

Wonders, C. P., Taylor, L., Welagen, J., Mbata, I. C., Xiang, J. Z., & Anderson, S. A. (2008). A spatial bias for the origins of interneuron subgroups within the medial ganglionic eminence. *Developmental Biology*, **314**(1), 127–136.

Wu, Y., & Parent, A. (2000). Striatal interneurons expressing calretinin, parvalbumin or NADPH-diaphorase: a comparative study in the rat, monkey and human. *Brain Research*, **863**(1–2), 182–191.

Xiang, Y., Tanaka, Y., Patterson, B., … Park, I.-H. (2017). Fusion of Regionally Specified hPSC-Derived Organoids Models Human Brain Development and Interneuron Migration. *Cell Stem Cell*. doi:10.1016/j.stem.2017.07.007

Xu Q, Cobos I, De La Cruz E, Rubenstein JL, A. S. (2004). Origins of Cortical Interneuron Subtypes. *Journal of Neuroscience*, **24**(11), 2612–2622.

Xu, Q., Guo, L., Moore, H., Waclaw, R. R., Campbell, K., & Anderson, S. A. (2010). Sonic Hedgehog Signaling Confers Ventral Telencephalic Progenitors with Distinct Cortical Interneuron Fates. *Neuron*, **65**(3), 328–340.

Xu, X., Tay, Y., Sim, B., … Pouladi, M. A. (2017). Reversal of Phenotypic Abnormalities by CRISPR/Cas9-Mediated Gene Correction in Huntington Disease Patient-Derived Induced Pluripotent Stem Cells. *Stem Cell Reports*. doi:10.1016/j.stemcr.2017.01.022

Yang, N., Chanda, S., Marro, S., … Wernig, M. (2017). Generation of pure GABAergic neurons by transcription factor programming. *Nature Methods*, **14**(6), 621–628.

Yao, Z., Mich, J. K., Ku, S., … Ramanathan, S. (2016). A Single-Cell Roadmap of Lineage Bifurcation in Human ESC Models of Embryonic Brain Development. *Cell Stem Cell*, 1–15.

Yizhar, O., Fenno, L. E., Prigge, M., … Deisseroth, K. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*, **477**(7363), 171–178.

Yu, J., Vodyanik, M. A., Smuga-Otto, K., … Thomson, J. A. (2007). Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science*, **318**(5858), 1917–1920.

Yuan, F., Chen, X., Fang, K.-H., … Liu, Y. (2018). Induction of human somatostatin and parvalbumin neurons by expressing a single transcription factor LIM homeobox 6. *ELife*, **7**. doi:10.7554/eLife.37382

Zecevic, N., Hu, F., & Jakovcevski, I. (2011). Interneurons in the developing human neocortex. *Developmental Neurobiology*, **71**(1), 18–33.

Zechel, S., Zajac, P., Lönnerberg, P., Ibáñez, C. F., & Linnarsson, S. (2014). Topographical transcriptome mapping of the mouse medial ganglionic eminence by spatially resolved RNA-seq. *Genome Biology*, **15**(10), 486.

Zeisel, A., Munoz-Manchado, A. B., Codeluppi, S., … Linnarsson, S. (2015). Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*, **347**(6226), 1138–1142.

Zhang, P., Yang, M., Zhang, Y., … Li, S. (2019). Dissecting the Single-Cell Transcriptome Network Underlying Gastric Premalignant Lesions and Early Gastric Cancer. *Cell Reports*, **27**(6), 1934-1947.e5.

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

Zheng, G. X. Y., Terry, J. M., Belgrader, P., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, **8**(1), 14049.

Zhong, S., Zhang, S., Fan, X., … Wang, X. (2018). A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature*, **555**(7697), 524–528.

Ziegenhain, C., Vieth, B., Parekh, S., … Enard, W. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Molecular Cell*, **65**(4), 631-643.e4.

Appendix

The appendix has been provided as online content organised according to the theme of the content. Provided below is the list of tables and figures within each Appendix folder to guide the reader to content of interest. The data in full has been provided electronically in the Dropbox folder named "PhD thesis - Appendix". Please open the following link to the shared Dropbox folders 1 to 8 below:

https://www.dropbox.com/sh/8ua9h2nzho924ss/AAAHL3LS-ddBAv_I89AXqAmJa?dl=0

FOLDER 1. Reagents list

Tables list:

Table 1. List of primary antibodies used in immunohistochemistry.

Table 2. List of secondary antibodies used in immunohistochemistry.

Table 3. List of primers used in QPCR reactions.

FOLDER 2. hESC-derived interneurons single-cell RNA sequencing and quality control checks

Tables list:

Table 4. Summary of the libraries sequenced on Illumina HiSeq for hESC-derived interneuron single-cell RNA sequencing study.

Table 5. List of differentially expressed genes generated by performing ANOVA pair-wise analysis between clusters on tSNE plot for each hESC-derived interneuron sample, figures 4.1C.

Table 6. List of differentially expressed genes generated by performing ANOVA pair-wise analysis between two hESC-derived interneuron samples.

Table 7. List of genes uniquely co-expressed by SST expressing hESC-derived neurons.

Table 8. List of genes uniquely co-expressed by CR expressing hESC-derived neurons.

Table 9. List of genes uniquely co-expressed by PVALB expressing hESC-derived. interneurons.

Figures list:

A) Schematic of the in vitro hESC interneuron differentiation protocol.

B) Representative examples of electropherograms generated from single-cell cDNA before and after library clean up.

C) Heatmap of relative gene expression of hESC-derived interneurons – at single-cell RNA level, day 20.

D) tSNE plot of the relative similarity in single-cell profiles of hESC-derived interneurons at day 20 (progenitors) and day 65 (post-mitotic neurons).

FOLDER 3. Human MGE single-cell RNA sequencing and quality control checks

Tables list:

Table 10. Summary of the libraries sequenced on Illumina HiSeq for human primary foetal MGE single-cell RNA sequencing study.

Table 11. List of differentially expressed genes generated by performing ANOVA pair-wise analysis between clusters on tSNE plot for each human MGE sample, figure 5.2C.

Table 12. List of differentially expressed genes generated by performing ANOVA pair-wise analysis between two human MGE samples.

Table 13. List of genes uniquely co-expressed by SST expressing human MGE-derived interneurons.

Table 14. List of genes uniquely co-expressed by CR expressing human MGE-derived interneurons.

Table 15. List of genes uniquely co-expressed by NPY expressing human MGE-derived interneurons.

Figures list:

E) Representative examples of electropherograms generated from single-cell cDNA before and after library clean up.

F) Heatmap of relative gene expression of primary human MGE– at single-cell RNA level, 10WPC and 15WPC.

G) tSNE plot of the relative similarity in single-cell profiles of primary human foetal MGE at 10WPC and 15WPC.

H) Relative gene expression values of known LGE, CGE and MGE enriched markers, in human MGE, CGE and LGE at 10WPC and 15WPC.

I) Longitudinal analysis of SST and LHX6 gene expression changes in LGE, CGE and MGE over 10WPC to 15WPC.

J) Molecular characterisation of primary human CGE ex vivo cultures– at protein level, immunostaining of 10WPC and 15WPC CGE.

K) Molecular characterisation of primary human LGE ex vivo cultures– at protein level, immunostaining of 10WPC and 15WPC LGE.

L) Molecular characterisation of primary human MGE ex vivo cultures– at protein level, immunostaining of 15WPC MGE.

FOLDER 4. In vivo vs in vitro single-cell RNA sequencing analysis

Tables list:

Table 16. List of differentially expressed genes generated by performing ANOVA pair-wise analysis between 10WPC human foetal MGE and hESC-derived interneuron samples.

Table 17. List of differentially expressed genes generated by performing ANOVA pair-wise analysis between 15WPC human foetal MGE and hESC-derived interneuron samples.

FOLDER 5. Plasmid digestions and viral transductions

Tables list:

Table 18. The entire list of plasmids used to generate lentiviral vector and digestion products.

Table 19. The entire list of plasmids used to generate the ChR2-mCherry glycoprotein-deficient rabies tracer and digestion products.

Figure list:

M) Gel electrophoresis displaying the products of the restriction enzyme digestion of the plasmids used to generate the lentiviral vector.

N) Gel electrophoresis displaying the products of the restriction enzyme digestion of the plasmids used to generate the viral tracer.

O) Flow cytometry measure of the proportion GFP positive neurons (starter neurons) in alive hESC-derived interneuron cultures treated with the lentiviral vector.

P) Flow cytometry measure of the proportion mCherry positive cells in alive HEK TVA cells treated with the unpseudotyped and EnvA pseudotyped ΔG ChR2-mCherry viral tracer.

Q) Flow cytometry measure of the proportion mCherry positive cells in alive HEK cells treated with the unpseudotyped and EnvA pseudotyped ΔG ChR2-mCherry viral tracer.

R) Flow cytometry measure of the proportion mCherry positive cells in alive HEK or HEK TVA cells in the negative controls (no viral treatment).

FOLDER 6. Electrophysiology traces

Figure list:

S-U) Representative illustrations of physiological responses of patched hESC-derived neurons to different whole-cell patch clamp paradigms, in current clamp mode.

V) Representative illustrations of spontaneous action potential recorded from a patched hESC-derived neuron in current clamp mode – at day 60.

W) Representative illustrations of spontaneous mini excitatory postsynaptic potentials recorded from a patched hESC-derived neuron – at day 60.

FOLDER 7. Cell counts

Figure list:

X) Average total cell count measured using DAPI count for interneuron differentiation of hESC analysed in Chapter 3, figure 3.1D-3.1E.

Y) Average total cell count measured using DAPI count for interneuron differentiation of hESC analysed in Chapter 3, figure 3.2B-3.2E.

Z) Average total cell count measured using DAPI count for interneuron differentiation of hESC analysed in Chapter 3, figure 3.3A-3.3E..

AA) Average total cell count measured using DAPI count for human foetal MGE, CGE and LGE analysed in Chapter 5, figure 5.1C.

AB) Average total cell count measured using DAPI count for the lentiviral transduced interneuron differentiation of hESC analysed analysed in Chapter 6, figure 6.3C.

AC) Average total cell count measured using DAPI count for viral tracer application HEK cells analysed analysed in Chapter 6, figure 6.4E.

AD) Average total cell count measured using DAPI count for viral tracer application to HEK-TVA cells analysed analysed in Chapter 6, figure 6.4C.

FOLDER 8. Academic publications

During the 4-year PhD programme at Cardiff University, the research I have carried out has contributed to the following publications: Noakes *et al.* 2019 (second author on research article) and Keefe & Li 2020 (pre-press review article, first author). Electronic copies of these articles have been provided in the Dropbox folder 8.