Single cell analysis of the adult neural stem cells of the lateral ventricle using a novel genetic marker



Thesis submitted for the degree of Doctor of Philosophy in Pharmacology at the University of Oxford

Arup Nath

New College

Trinity 2019

Declaration

I hereby declare that the work presented in this thesis is my own unless otherwise stated. This thesis has not been submitted to anywhere for any degrees or professional qualifications except to the University of Oxford for the degree of Doctor of Philosophy.

Arup Nath
Trinity 2019

Abstract

Canonically, adult neurogenesis occurs, in mammals, in two distinct regions of the brain—the subgranular zone of the dentate gyrus and the ventricular-subventricular zone (V-SVZ) of the lateral ventricle. The V-SVZ germinal niche is comprised, in part, by ependymal cells and the quiescent neural stem cells (NSCs) that lie near these cells. During neurogenesis, the quiescent NSCs change their transcriptional state and alter their cellular machinery to become activated NSCs. Together, quiescent and activated NSCs are also known as B-cells. B-cells generate transit amplifying cells, which then generate neuroblasts. Initially, it was believed that the function of V-SVZ neurogenesis was to populate the olfactory bulb with new neurons. However, recent work suggests that under disease and injury states it can lead to the formation of new striatal neurons.

Studying B-cells is important because although they were identified two decades ago, many important questions regarding their development, specification, and function remain unanswered. The fate of B-cells is specified at the time of their formation during embryonic development. This work suggested that there existed many different subpopulations of B-cells—each fated to generate different types of neurons in the adult mouse brain. However, the mechanism by which the B-cells are apportioned to the different subgroups still remains a mystery.

In this thesis, we demonstrate that a BAC-Penk-CRE transgenic line (which causes CRE expression in cells with an active pre-pro-Enkephalin promoter), generated in our lab and crossed to an appropriate reporter strain, labels a novel subgroup of B-cells, as well as previously identified subgroups of B-cells and neuroblasts. Using single-cell RNA sequencing, and other validation techniques, we show that this novel subgroup expresses the 5-Hydroxytryptamine Receptor 2C (5-HT_{2C} receptor) and the prolactin receptor—which have been functionally correlated with B-cells, but B-cells expressing these receptors have never been captured and sequenced before. We also show that this novel subgroup expresses the KiR7.1 channel and the folate receptor. These findings are novel and exciting and will allow for the further study of the heterogeneity of B-cells and to investigate the many unanswered questions.

Finally, not only have we been able to validate the RNAseq data, but we also show that in a partial progressive model of Parkinson's disease, cells expressing the BAC-Penk-CRE reporter increase their proliferation and can be found in the striatal parenchyma. Thereby demonstrating the functional importance of this group of V-SVZ stem cells.

Acknowledgements

I would like to first thank my supervisor Dr. Liliana Minichiello for her support both scientifically and personally throughout my doctorate. She has pushed me to become a better scientist and also a better person. Her words of encouragement and guidance have given me the confidence to pursue the work presented in this thesis. I will be forever grateful. I would also like to thank my co-supervisor Dr. Claus Nerlov for his insights and critical inputs during the design of the experiments and during data analysis.

The work presented in this thesis was the result of a massive collaborative effort and I would like thank several people in particular for their help and guidance. From the Weatherall Institute Molecular Medicine I would like to thank Dr. Emmanouela Repapi and Dr. Neil Ashley for their help with the RNA sequencing experiments. I would like to thank Dr. Paul Sopp and Kevin Clark for their help with flow cytometry. Dr. Roy Drissen provided immense help at many stages of my doctorate, without which this thesis would not be possible.

I would like to thank the members of the Minichiello lab throughout the years for their help and friendship. In particular, I would like to thank Dr. Barbara Zonta who has been both a friend and mentor.

Finally, I would like to thank the people I love—my family and my friends—for their unwavering support throughout my doctorate. Post-graduate life can be challenging and without their love and support, none of this would be possible.

Table of Contents

List of Abbreviations 1

Aims of the Thesis 3

Chapter 1 - Introduction

Adult neural stem cells and the ventricular germinal niche

1.1 Adult neural stem cells of the subventricular zone. 5
1.2 Markers of V-SVZ germinal niche cells. 6
1.3 The V-SVZ germinal niche. 9
1.4 The effects of ageing and sex on V-SVZ NSCs. 12
1.5 The development of the B-cells of the V-SVZ. 12
1.6 The dynamics of B-cell division. 13

An overview of the striatum

1.7 The striatum and its function. 15
1.8 Relevant cell types in the striatal parenchyma. 16
1.9 The IPNK mouse line. 17

Neurogenesis and neurodegenerative diseases

1.10 An overview of Parkinson's disease. 20
1.11 The role of V-SVZ neurogenesis in Parkinson's disease. 21

An overview of methods used to perform single cell analysis

1.12 Fluorescence-activated cell sorting. 24
1.13 High-throughput, microfluidic gene-expression analysis. 25
1.14 Single cell RNA-sequencing. 26

Chapter 2 - Materials and Methods

2.1 Mouse lines. 31
2.2 Solutions 33
2.3 Antibodies 34
2.4 Preparation and perfusion of animals for immunochemistry. 35
2.5 Cryosectioning. 35
2.6 Fluorescent immunochemistry. 35
2.7 Immunohistochemistry. 36
2.8 Fluorescence microscopy. 37
2.9 Penk-tdTomato^{+} cell counting and estimation. 37
2.10 Tissue processing for FACS. 38
2.11 Antibody incubation for FACS. 39
2.12 FACS protocol. 39

2.13 Cell processing for Smart-seq2. 40
2.14 Bioinformatics Analysis. 46
2.15 High-throughput gene expression analysis. 49
2.16 Tissue processing for wholemount lateral ventricle staining. 53
2.17 CLARITY tissue processing and imaging. 54
2.18 BrdU analysis. 55
2.19 Stereotactic injection of 6-OHDA. 55
2.20 Statistical analysis. 56

Chapter 3 – Results I: Using the Penk-tdTomato (IPNK) mouse line to isolate and sequence adult neural stem cells.

3.1 Penk-tdTomato expressing cells at the ventricular/sub-ventricular zone of the lateral ventricle. 57
3.2 Optimising the isolation protocol for single cells from acutely dissected striatal tissue. 62
3.3 Fluorescence-activated cell sorting-mediated isolation and capture of single Penk-tdTomato+ cells. 62
3.4 StriatumSeq1: Deep RNA sequencing of single Penk-tdTomato$^+$ cells. 66
3.5 Initial bioinformatics analysis of the RNA sequencing data. 67
3.6 Quality control analysis of the RNA sequencing data. 67
3.7 Unbiased clustering of Penk-tdTomato$^+$ cells. 73
3.8 Establishing the cellular identities of the clusters. 77
3.9 The Penk-tdTomato+ clusters display molecular signatures of V-SVZ neural stem cells at two stages of differentiation. 80
3.10 *In vivo* validation of the StriatumSeq1 clusters on IPNK brain sections. 85
3.11 Elimination of the microglial signature by FACS. 86
3.12 Validation of microglial signature elimination and M-Peak cells. 87
3.13 StriatumSeq2: Deep RNA sequencing of single Penk-tdTomato$^+$ cells of the V-SVZ. 90
3.14 Clustering of the Penk-tdTomato$^+$ cells. 90
3.15 The Penk-tdTomato+ reporter line captures V-SVZ NSCs at three stages of the differentiation arc. 93
3.16 Cell cycle analysis. 95
3.17 Analysing gene expression of the three clusters in StriatumSeq2. 97
3.18 Gene ontology analysis on the three clusters. 105
3.19 Comparing the StriatumSeq datasets to a published V-SVZ NSC single cell sequencing dataset. 108
3.20 Enkephalin gene expression in sequenced Penk-tdTomato$^+$ cells. 112

Chapter 4 – Results II: The *in vivo* validation of the bioinformatics data.

4.1 An immunocytochemical analysis of Penk-tdTomato+ cell types in the V-SVZ germinal niche. 114
4.2 *In vivo* validation of novel and specific markers of the Group1 cluster. 120

4.3 *En face* staining of the subventricular zone using wholemounts. 125
4.4 Validation of specific markers of the Group2 cluster. 129
4.5 Validation of specific markers of the Group3 cluster. 132
4.6 Three-dimensional visualisation of Penk-tdTomato$^+$ cells at the V-SVZ. 136
4.7 Using novel cell-surface markers to functionally isolate Group1 cells. 138

Chapter 5 – Results III: The *in vivo* functional analysis of Penk-tdTomato+ cells in a partial progressive model of Parkinson's disease.

5.1 Establishing the partial progressive model of Parkinson's disease. 143
5.2 The partial progressive model of Parkinson's disease causes a massive increase in cell proliferation in the affected striatum. 147
5.3 The partial progressive model of Parkinson's disease causes an increase in Penk-tdTomato$^+$ cell proliferation in the affected striatum. 149
5.4 The proliferating Penk-tdTomato$^+$ cells in the affected striatum do not significantly upregulate the expression of neuroblast markers. 151
5.5 Sham injection does not result in a specific increase of proliferating Penk-tdTomato$^+$ cells or of proliferating Penk-tdTomato+ cells expressing neurogenic markers. 155

Chapter 6 – Discussion

6.1 The relationship between Penk and tdTomato expression in IPNK mice. 157
6.2 The functional significance of Penk expression in neural stem cells. 158
6.3 Group1 cells likely represent a subset of B-cells. 160
6.4 Novel cell-surface markers for B-cells and their importance for the further characterization of B-cells. 162
6.5 The developmental arc of the three clusters. 164
6.6 Classifying the three clusters. 165
6.7 The isolation and sequencing of the Group2 cluster in StriatumSeq2. 166
6.8 Comparisons to previously published reports of sequenced V-SVZ cells. 168
6.9 Are Group1 cells choroid plexus cells? 169
6.10 Microglial signature in StriatumSeq1. 171
6.11 "M-Peak" cells display the signature of mature striatal neurons. 173
6.12 The use of whole-striatum lysates in place of micro-dissected V-SVZ tissues and alternative sequencing methods. 173
6.13 The use of a relatively small number of cells to perform bioinformatics analysis. 174
6.14 The partial progressive model of Parkinson's disease. 175
6.15 Experimental paradigm and revision. 179
6.16 Penk-tdTomato$^+$ cells proliferate in response to 6-OHDA injection into the dorsal striatum. 181
6.17 Penk-tdTomato$^+$ cells do not become neurogenic in response to 6-OHDA injection into the dorsal striatum. 183

Conclusions and Outlook 186

Bibliography 188

Appendix 199

List of Abbreviations

Abbreviation	*Term*
6-OHDA	6-hydroxydopamine
BrdU	Bromodeoxyuridine
ChAT	Choline acetyltransferase
CNS	Central Nervous system
CSF	Cerebrospinal fluid
D1R	Dopamine D1 receptor
D2R	Dopamine D2 receptor
DAB	3,3'-diaminobenzidine
DAPI	4′,6-diamidino-2-phenylindole
FACS	Fluorescence-activated cell sorting
FSC	Forward scatter
GO	Gene Ontology
GP	Globus pallidus
GPe	External segment of globus pallidus
GPi	Internal segment of globus pallidus
IPNK	BAC-Penk-Cre$^{tg/+}$;Rosa26-tdTomato$^{tg/+}$
LGE	Lateral ganglionic eminence
M-Peak	Middle-Peak
MCAO	Middle cerebral artery occlusion
MFB	Medial forebrain bundle
MGE	Medial ganglionic eminence

MPTP	1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine
MSN	Medium spiny neuron
PAGODA	Pathway and Gene Set Overdispersion Analysis
PB	Phosphate buffer
PBS	Phosphate buffered saline
PD	Parkinson's disease
Penk	Pre-pro-Enkephalin
PFA	Paraformaldehyde
RT-PCR	Reverse transcription polymerase chain reaction
SGZ	Subgranular zone
SN	Substantia nigra
SNc	Substantia nigra pars compacta
SNr	Substantia nigra pars reticulata
SSC	Side scatter
STAR	Spliced Transcripts Alignment to a Reference
STN	Subthalamic nucleus
V-SVZ	Ventricular-subventricular zone

Aims of the Thesis

We hypothesised that the BAC-Penk-Cre$^{tg/+}$;Rosa26-tdTomato$^{tg/+}$ (IPNK) line labeled stem cells in the ventricular-subventricular zone (V-SVZ) stem cell niche of the adult mouse, based on preliminary evidence (presented in the thesis). The IPNK line permanently labels any cells with an active pre-pro-Enkephalin promoter. Thus, the central aim of the thesis was to isolate and characterise the Penk-tdTomato^{+} cells of the V-SVZ. A secondary aim was to utilise our unique genetic tool to possibly identify markers, particularly cell-surface ones, that could specifically isolate B-cells in an unbiased manner. We also sought to investigate whether Penk-tdTomato^{+} cells can proliferate, migrate, and engage in neurogenesis in a model of neurodegeneration.

We arrived at our hypothesis after visual inspection of BAC-Penk-Cre$^{tg/+}$ mice crossed to several different reporter lines indicated that cells of the V-SVZ were being labeled, and thus expressed Penk at some point during their lifetime. To address the hypothesis, we sought to use single cell RNA sequencing. In order to obtain sequencing-ready cells, we utilised an isolation and sequencing procedure optimised in the Minichiello group (Nath *et al.,* 2017; Vatanashevanopakorn *et al.,* 2017). The sequencing data was then processed and analysed to determine what cell types were labeled with the IPNK reporter line.

In the second phase of the project, we sought to validate the RNA sequencing data. This would allow us to determine whether the cell types and markers identified by bioinformatics analysis were present in the brain of IPNK mice. We planned to do this by immunochemical analysis of mouse brain sections and three-dimensional imaging of the intact brain using the CLARITY protocol. We also aimed to validate the data in a more functional way. We reasoned that if the bioinformatics analysis did in fact reveal specific markers for Penk-tdTomato^{+} cells of the V-SVZ, then we should be able to use said markers to isolate those cells. To achieve this, we aimed to find cell-surface markers within the sequencing data that identified subgroups of Penk-

tdTomato$^+$ cells and obtain antibodies against those markers in order to be able to use FACS to perform the capture. We would then use high-throughput gene expression analysis to decide whether the target cell population was indeed captured.

The final phase of the project was to understand whether or not the hypothesized Penk-tdTomato$^+$ stem cells could be induced into proliferation and neurogenesis in a disease model. If Penk-tdTomato$^+$ cells of the V-SVZ were indeed stem cells, then we reasoned that they could become a means of therapeutic intervention. Therefore, to determine if they could be mobilised we used a model of partial progressive Parkinson's disease, induced in IPNK mice through the unilateral stereotactic injection of 6-OHDA.

Chapter 1 – Introduction

Adult neural stem cells and the ventricular germinal niche

1.1 Adult neural stem cells of the rodent subventricular zone.

Since the discovery of neurogenesis in adult brains several decades ago, much work has been done to probe the dynamics, utility, and identity of the cells involved in this process (Altman, 1969; Goldman and Nottebohm, 1983). It is now generally accepted that neurogenesis occurs in at least two distinct regions of the brain—the subgranular zone (SGZ) of the dentate gyrus in the hippocampus, and the epithelial structure that is the ventricular-subventricular zone (V-SVZ) of the lateral ventricles (Chaker *et al.,* 2016; Figure 1.2). Though hippocampal neurogenesis has been found to be important for the functioning of the mammalian brain in both rodents and humans—several recent reports in humans notwithstanding—this thesis and chapter will only address neurogenesis in the V-SVZ (Boldrini *et al.,* 2018; Sorrells *et al.,* 2018). The cells generated in the V-SVZ were initially understood to mainly be interneurons that would migrate immense distances (3-8 mm in mice) along the rostral migratory stream (RMS) to populate particular layers of the olfactory bulb (Merkle *et al.,* 2007; Merkle *et al.,* 2014; Obernier and Alvarez-Buylla, 2019; Platel *et al.,* 2019). Some astrocytes and oligodendrocytes are also reported to be generated by the stem cells of the V-SVZ (Lim and Alvarez-Buylla, 2014). However, disease and injury have been shown to promote the generation of new neurons and interneurons that go on to populate the striatal parenchyma (Arvidsson *et al.,* 2002; Ernst *et al.,* 2014).

The neural stem cells (NSCs) of the V-SVZ were initially identified to be in the astroglial lineage, and are also referred to as B-cells (Doetsch *et al.,* 1997; Doetsch *et al.,* 1999). Recently, it has been reported that the NSCs can be found in two states—quiescent or active. Active NSCs are derived from the quiescent NSCs, a process which involves, among other things, the

upregulation of Epidermal Growth Factor signalling and clearing of protein aggregates in lysosomes (Codega *et al.,* 2014; Llorens-Bobadilla *et al.,* 2015; Dulken *et al.,* 2017; Leeman *et al.,* 2018). Quiescent NSCs can also be differentiated by their specific activation of particular cellular pathways, such as those for cell-cell adhesion, extracellular matrix responses, cell communication, signaling receptors, transmembrane transporters, and ion channels (Chaker *et al.,* 2016). Activated NSCs on the other hand, upregulate processes related to cell cycle, DNA repair, and proteasomes (Chaker *et al.,* 2016; Leeman *et al.,* 2018). The two B-cell types also differ in how they perform energy metabolism. Quiescent NSCs upregulate processes related to glycolysis and fatty acid metabolism, whereas their activated counterparts favour oxidative phosphorylation (Lim and Alvarez-Buylla, 2014; Chaker *et al.,* 2016).

Activated NSCs give rise to transit amplifying cells (C-cells) which divide symmetrically approximately three times to generate clusters of neuroblasts (A-cells; Lim and Alvarez-Buylla, 2014). The neuroblasts can then divide up to one or two more times as they migrate along the RMS towards the olfactory bulb (Ponti *et al.,* 2013). It is estimated that the V-SVZ generates approximately 10,000 migrating progenitor cells every day in the young adult mouse (Lois *et al.,* 1996).

1.2 Markers of V-SVZ germinal niche cells.

B-cells were first identified as Glial Fibrillary Acidic Protein (GFAP)-expressing astroglial cells, and this protein remains the most commonly used marker of quiescent NSCs to this day (Doestch *et al.,* 1999). However, GFAP is not an exclusive marker for B-cells, as it is also a widely-used marker for astrocytes. Recently, many groups have also used combinations of antibodies to isolate what they claim are B-cells. To this end, cells expressing both CD133 and GLAST (Glutamate Aspartate Transporter) have been isolated and sequenced as B-cells (Llorens-Bobadilla *et al.,* 2015). Once again, it must be noted that GLAST is also a widely-used marker for

astrocytes. Other groups have used the dual expression of GFAP and CD133 to perform similar analyses (Dulken *et al.,* 2017). The use of CD133 has been controversial due to the fact that it is also a marker of the ependymal cells lining the ventricular wall (Carlen *et al.,* 2009; Chaker *et al.,* 2016; Shah *et al.,* 2018). More recently, the protein Troy (Tumor Necrosis Factor Receptor Superfamily, Member 19) has been used to isolate quiescent NSCs of the V-SVZ (Basak *et al.,* 2018). Quiescent NSCs, however, have been found to be negative for the widely used neural stem cell marker Nestin (Chaker *et al.,* 2016). As is evident here, there exist multiple markers of quiescent NSCs of the V-SVZ. What is not clear however, is whether the different markers isolate different subgroups of B-cells. This may be a possibility since the sequencing data generated by using different markers does indeed lead to significant differences in the results (Llorens-Bobadilla *et al.,* 2015; Dulken *et al.,* 2017; Basak *et al.,* 2018; Obernier and Alvarez-Buylla, 2019). Thus, it is of importance to the field and the continued study of B-cells to identify specific markers that can be used in a standardised manner.

As quiescent NSCs transition into activated NSCs, they maintain the expression of GFAP and GLAST whilst downregulating CD133 expression (Llorens-Bobadilla *et al*., 2015; Chaker *et al.,* 2016). Crucially, they also begin to express the Epidermal Growth Factor Receptor (EGFR; Llorens-Bobadilla *et al*., 2015; Chaker *et al.,* 2016). Activated NSCs also express Nestin and can express Ascl1 (Achaete-Scute family BHLH transcription factor 1; Lim and Alvarez-Buylla 2014). C-cells have been identified as expressing EGFR but downregulating the expression of the B-cell markers GFAP and CD133. They are also negative for the neuroblast marker CD24a (Dulken *et al.,* 2017). A-cells, or neuroblasts, are easily identified by their specific expression of the canonical markers CD24a, βIII-tubulin, Doublecortin (Dcx), and Polysialylated-Neural Cell Adhesion Molecule (PSA-NCAM; Mamber *et al.,* 2013).

Markers summarised in Table 1.1 below.

Table 1.1

Cell type	Abbreviation	Alternative nomenclature	Protein marker
Quiescent neural stem cell	qNSC	B-cell	GFAP, CD133, GLAST, Troy, SFRP1
Activated neural stem cell	aNSC	B-cell	GFAP, GLAST, EGFR, Nestin, Ascl1
Transit amplifying cell	TAC	C-cell	EGFR
Neuroblast	Nb	A-cell	CD24a, βIII-tubulin, Dcx, PSA-NCAM

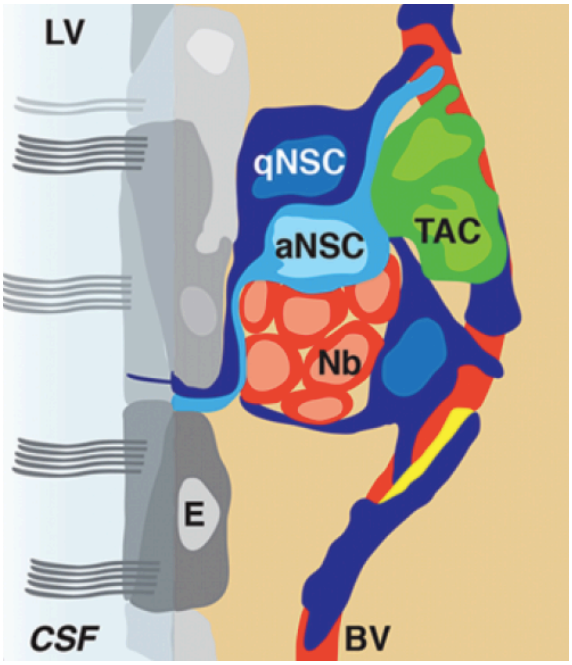


Figure 1.1. A diagram of relevant cell types of the V-SVZ stem cell niche. On the apical side of the niche multi-ciliated ependymal cells (**E**) enclose it from the lateral ventricle (**LV**) and the cerebrospinal fluid (**CSF**) that bathes it. The basal side of the V-SVZ is delimited by a vascular plexus formed in part by blood vessels (**BV**). B-cells are the neural stem cells of the niche. They are either quiescent (**qNSC**) or active (**aNSC**). The B-cells generate transit amplifying cells (**TAC**s or C-cells), which then generate neuroblasts (**Nb** or A-cells). Adapted from Chaker *et al.* (2016).

1.3 The V-SVZ germinal niche.

The epithelial niche is largely formed on the apical side by multi-ciliated ependymal cells which are arranged in pinwheel structures to form the ventricular wall (Mirzadeh *et al.,* 2008). On the basal side a vascular plexus separates the V-SVZ niche from the striatal parenchyma (Figure 1.2; Tavazoie *et al.,* 2008; Shen *et al.,* 2009). The quiescent NSCs send an apical projection towards the ventricular wall and extend a single primary cilium into the ventricular space—thus allowing them to sample the cerebrospinal fluid (CSF) flow and content (Mirzadeh *et al.,* 2008). The B-cells are located in specific locations along the ventricular wall, where the position correlates with the different types of olfactory bulb interneurons they will go on to produce (Figure 1.3; Merkle *et al.,* 2007). The position of the B-cells along the wall is determined by specific transcription factor-expression (Merkle *et al.,* 2014). Recent work has shown that these positional cues are established during embryonic development and maintained throughout the life of the animal (Fuentealba *et al.,* 2015). These findings provided the first evidence for heterogeneity within the pool of B-cells that reside in the V-SVZ—indicating that there may be many different subgroups of B-cells, which are somehow marked and sorted during embryonic development to later form particular cell types and occupy different areas of the V-SVZ niche.

The CSF on the apical side of the V-SVZ and the blood vessels on the basal side provide the niche with many signals that help to regulate neurogenesis (Chaker *et al.,* 2016). Sonic Hedgehog (SHH), Wnt, Notch, Bone Morphogenetic Proteins (BMPs), and Ephrins are some of the signalling pathways that are activated due to the received signals discussed below (Lim and Alvarez-Buylla, 2014). Furthermore, cell-cell contact between NSCs and endothelial and ependymal cells also help to promote the quiescence of B-cells (Chaker *et al.,* 2016). Neurotransmitters also work to regulate the cells of the V-SVZ niche. A dense plexus of serotonin (5-HT) projections originating in the Raphe nuclei contacts both ependymal and B-cells on the apical side, thereby regulating proliferation (Tong *et al.,* 2014). Dopaminergic deafferentiation has

been reported to have various effects on V-SVZ proliferation and neurogenesis. It is believed that some of this effect is exerted by dopamine release onto D3-type dopamine receptors (Kim *et al.,* 2010). Cells expressing ChAT have also been reported to reside within the V-SVZ niche; whose activity is also said to regulate proliferation and neurogenesis (Paez-Gonzalez *et al.,* 2014). Recent work by Fiona Doetsch's group has found that hypothalamic proopiomelanocortin neurons innervate the V-SVZ and promote the proliferation of a subgroup of B-cells (Paul *et al.,* 2017). Glutamate released from striatal neurons has been show to act on AMPA receptors on neuroblasts to increase their proliferation (Song *et al.,* 2017). Neuroblasts can themselves release GABA extrasynaptically that act on GABA-A receptors on B-cells to promote quiescence (Liu *et al.,* 2005). Activated NSCs and C-cells express the Notch ligand Delta-like-1 (Dll1) which acts on quiescent NSCs expressing Notch receptors to maintain their quiescence (Kawaguchi *et al.,* 2013). The lateral ventricle choroid plexus, which is a key regulator of brain homeostasis and CSF production, also releases factors that regulate quiescence, proliferation, and neurogenesis within the V-SVZ through the release of NT-3 which acts on TrkC-expressing quiescent NSCs (Delgado *et al.*, 2014; Silva-Vargas *et al.,* 2016).

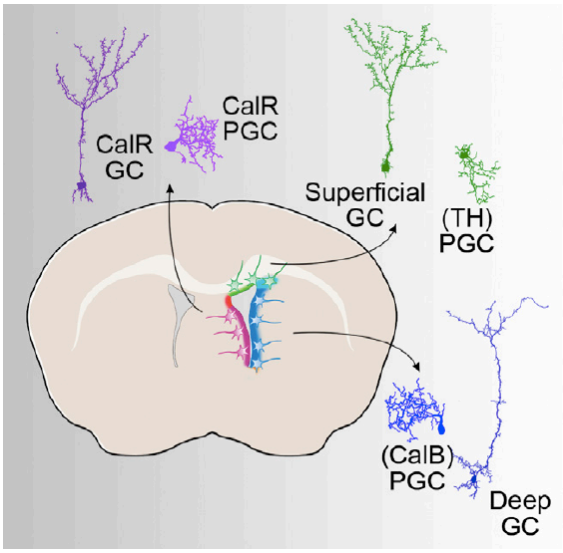


Figure 1.2. A diagram depicting the fate specification of adult B-cells. The schematic summarises the findings of the Alvarez-Buylla group on the fate specification of adult B-cells of the V-SVZ. They find the fate of B-cells are determined during embryonic development. They also find that this fate specification correlates to the expression of specific transcription factors and also the geographic location of the B-cells along the walls of the lateral ventricles. Such that B-cells from a certain area will express a certain transcription factor and will then generate a specific type of olfactory bulb interneuron. Adapted from Fuentealba *et al.* (2015).

1.4 The effects of ageing and sex on V-SVZ NSCs.

The number of quiescent and activated NSCs decrease dramatically as the animal ages, with concomitant changes to the very architecture of V-SVZ niche (Conover and Shook, 2011; Capilla-Gonzalez *et al.,* 2014). The exact reasons and mechanisms behind these changes are not yet fully understood, though released factors discussed above as well as genetic and epigenetic changes are hypothesised to play roles (Chaker *et al.,* 2016). A recent report by Anne Dulken's group found that quiescent NSCs have increasingly impaired lysosomal activity as the animal ages, which contributes to higher accumulation of protein aggregates shown using LAMP-2 and LysoTracker imaging. This failure to clear the protein aggregates is thought to trap the quiescent NSCs and prevent them from differentiating and proliferating—shown by preventing lysomomal acidification using bafilomycin A (Leeman *et al.,* 2018).

Sex is also thought to play a role in neurogenesis. In the V-SVZ, it has been shown that circulating prolactin in mice, increased in females during pregnancy, increases neurogenesis (Shingo *et al.,* 2003; Mak *et al.,* 2008). Similar findings have been reported with estrogen, which increases hippocampal neurogenesis in female rats (Pawluski *et al.,* 2009). A recent sequencing report from the lab of Fiona Doetsch has shed some light on the issue of sex differences (Mizrak *et al.,* 2019). Here, it was reported that males were significantly enriched in oligodendrocyte progenitors when compared to female mice. This study was not able to isolate and sequence the full spectrum of neurogenic niche cells (specifically qNSCs), thus future studies will be required to fully understand the exact interplay between sex and ageing in V-SVZ neurogenesis (Mizrak *et al.,* 2019).

1.5 The development of the B-cells of the V-SVZ.

B-cells are formed in the embryonic ventricular-subventricular zone as early as E11.5 when highly proliferative radial glia exit the cell cycle and become dormant until adulthood

(Fuentealba *et al.,* 2015). It is also at this stage that lineage-specification is thought to be imposed on the cells that will go on to become B-cells in the adult V-SVZ. The further specification of radial glia into B-cells is thought to be driven by the expression of Vcam1, GFAP, Cdkn1c (Kokovay *et al.,* 2012; Fuentealba *et al.,* 2015; Yuzwa *et al.,* 2017). The interplay of the expression of these proteins and how they exert such precise control over B-cell specification is not yet understood. But it is known that the precise fate of these B-cells, in terms of the cell types they will generate in the adult, is determined at this stage (Merkle *et al.,* 2014; Fuentealba *et al.,* 2015).

1.6 The dynamics of B-cell division.

As discussed above, the pool of B-cells is set during mid-gestation for the mouse and neurogenesis continues throughout the life of the animal (Fuentealba *et al.,* 2015). Thus, it is important to understand the processes that regulate quiescence and self-renewal. The exact dynamics of stem cell self-renewal at the V-SVZ are still not fully understood, though two recent reports shed some light on the process (Obernier *et al.,* 2018; Basak *et al.,* 2018). Stem cells can either divide symmetrically or asymmetrically. Symmetrical division can result in the generation of two B-cells or two C-cells, while asymmetrical division would result in one B-cell and one C-cell. Even before the emergence of direct evidence, symmetric division was thought to be the main mode of self-renewal since asymmetric division does not allow for the independent regulation of self-renewal and neurogenesis, while in the rodent brain there is overwhelming evidence for precise control of stem cell numbers and neurogenesis (Obernier *et al.,* 2019). Using the GFAP-GFP mouse line to label quiescent NSCs, the Alvarez-Buylla group showed that the stem cells of the V-SVZ divide symmetrically, but that there is also population asymmetry (Obernier *et al.,* 2018). It was found that approximately 80% of B-cells divide symmetrically to generate two C-cells, while 20% divide symmetrically to generate two B-cells (Figure 1.4). In this way, the niche is able to control both self-renewal and neurogenesis independently. An independent study using

Troy^{+} NSCs arrived at the same conclusions (Basak *et al.,* 2018). This finding goes some ways towards explaining why the numbers of quiescent NSCs decline with age. However, there remain many unanswered questions as to why there is such a propensity towards C-cell generation and not B-cell generation. For instance, are there cell-intrinsic differences in the B-cells that divide to produce C-cells and those that produce B-cells, or is this somehow controlled by the many different signals present within the niche? It is intriguing to consider the possibility that subgroups of B-cells divide differently—a hypothesis that will need to be tested in the future.

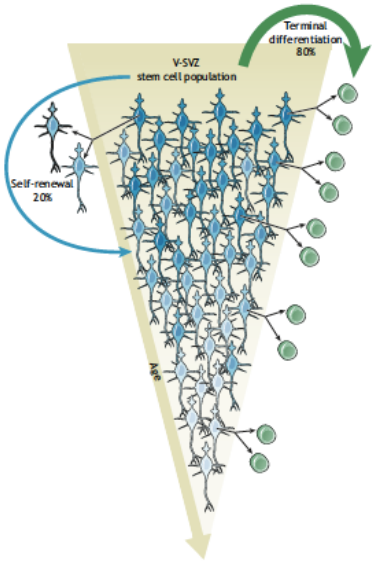


Figure 1.3. Stem cell self-renewal dynamics at the V-SVZ. This diagram summarises the current findings regarding the dynamics of self-renewal in the V-SVZ stem cell niche in an adult mouse brain. The overall numbers of quiescent NSCs decreases with age. This is thought to be a consequence of the way in which quiescent NSCs self-renew. The group of Alvarez-Buylla proposes that, while there is population asymmetry, the NSCs themselves self-renew symmetrically. They find that approximately 80% of NSCs divide symmetrically to produce two neuroblasts, while 20% self-renew to produce two more quiescent NSCs. Adapted from Obernier and Alvarez-Buylla (2019).

An overview of the striatum

1.7 The striatum and its function.

The striatum is one of the four nuclei which make up the basal ganglia, which is the largest subcortical structure in the mammalian forebrain. Itself comprised of two structures—the caudate nucleus and the putamen—the striatum plays key roles in movement planning, executive functions, and cognition (Graybiel, 2000). It is able to effect such varied actions because of the way its inputs and projections are organised. The striatum receives glutamatergic inputs from both the cortex and the thalamus and dopaminergic inputs from the substantia nigra, thus acting as a key convergence point for brain function. The striatum sends GABAergic outputs to the subthalamic nuclei and the pallidum, which then send further GABAergic projections to the thalamus, whose gluatamatergic efferents extensively innervate the overlying cortex (Bolam *et al.,* 2000). In this way, the striatum is able to control actions and movements while also regulating higher-order functions such as cognition.

Canonically, the striatum is understood to coordinate its functions via the two distinct circuits that emanate from it—the direct (striatonigral) and the indirect (striatopallidal) pathway (Figure 1.1). Regardless of the particularities of the pathways, the main signal is one of inhibition (Albin *et al.,* 1989; Chevalier and Deniau, 1990). This is due to the fact that the projection neurons of the striatum—the medium spiny neurons (MSNs)—are GABAergic. Due to the nature of the circuit make-up, however, this inhibition can have varied effects. Striatonigral, direct pathway, MSNs send GABAergic projections directly to the output nuclei of the basal ganglia, the substantia nigra pars reticulata (SNr) and the entopeduncular nucleus (EPN), which in turn project to the thalamus. This pathway leads to the inhibition of the inhibitory neurons of the output nuclei, thereby disinhibiting the thalamus (Bolam *et al.,* 2000). Thus, the direct pathway is said to promote movement-related behaviours. Striatopallidal, indirect pathway, MSNs project to and inhibit the

globus pallidus external segment (GPe) which inhibits the subthalamic nucleus (STN), which in turn sends excitatory glutamatergic projections to the SNr, thus leading to the inhibition of the thalamocortical projection neurons. Since the final action of the indirect pathway leads to the inhibition of thalamic neurons, it is said to inhibit movement-related behaviours (Bolam *et al.,* 2000). Though recent work, in particular the herculean efforts by the Costa group, suggests that the above functional model may be too simplistic to explain how motor planning is actually performed in an awake rodent, the anatomical parsing of MSNs is still of acute importance (Tecuapetla *et al.,* 2016).

1.8 Relevant cell types in the striatal parenchyma.

The striatum is mainly populated by MSNs, which are estimated to account for 90-95% of all striatal neurons (Kemp and Powell, 1971). The direct pathway neurons are molecularly identified and differentiated by their expression of Substance P and Dynorphin and the preferential expression of the κ-opioid receptor, as well as the D1 dopamine receptor (Kreitzer, 2009; Ophius *et al.*, 2014). The indirect pathway neurons conversely express Enkephalin (Penk), the D2 dopamine receptor and preferentially express the δ-opioid receptor (Kreitzer, 2009; Ophius *et al.*, 2014). It is the indirect pathway neurons with active Penk promoters that our IPNK mouse line labels (described below). Due to their variable expression of dopamine receptors, the MSN types respond differently to the dense dopaminergic innervations from the substantia nigra *pars compacta* (Surmeier *et al.,* 2007). The D1 receptors are thought to respond by increasing dendritic excitability in striatonigral MSNs, whereas D2 receptors are believed to decrease the dendritic excitability in striatopallidal MSNs (Surmeier *et al.,* 2007). Recently, a report that sequenced the MSNs of the striatum found that there was further heterogeneity within the cell types. The study also revealed that there was a great degree of overlap in the expression of some of the markers described above, which were previously thought to be expressed in a mutually exclusive manner

(Gokce *et al.,* 2016). Even though the two MSN types are, in general, molecularly distinct, their physiological properties are relatively similar. The only exception being that striatopallidal neurons are thought to be slightly more excitable (Kreitzer and Malenka, 2007).

As stated above, MSNs comprise 90-95% of striatal neurons, the remaining neurons are GABAergic and cholinergic interneurons that modulate the activity of the MSNs (Kawaguchi *et al.*, 1993). As with other interneurons, striatal interneurons modulate the activity of the projection neurons of the striatum. Of the GABAergic interneurons, those that express the calcium binding protein Parvalbumin (PV) make up the largest population (Kita *et al.,* 1990). The other GABAergic interneurons singly express the proteins Somatostatin (SST), Neuropeptide Y (NPY), Nitric Oxide Synthase (NOS), or Calretinin (Calb2; Kreitzer, 2009). Large cholinergic interneurons are also present within the striatal parenchyma, and can be immunochemically identified by their expression of the enzyme Choline O-acetyltransferase (ChAT; Calabresi *et al.*, 2000).

1.9 The IPNK mouse line.

The BAC-Penk-Cre$^{tg/+}$;Rosa26-tdTomato$^{tg/+}$ (IPNK) line was developed in the Minichiello laboratory to permanently label any cells with an active pre-pro-Enkephalin (Penk) promoter (Besusso *et al.,* 2013). As described previously, within the striatum, Penk is specifically expressed by indirect D2 receptor-expressing MSNs. Thus, the line was initially generated to label indirect pathway MSNs with tdTomato for further functional analyses of these cells—including single cell RNA sequencing.

Figure 1.5 shows a coronal brain section of IPNK mouse—here the endogenous fluorescence of the tdTomato reporter is visualised. As expected, there is strong staining of the striatum—indicating that is expressed in the correct brain region. Interestingly however, we also see strong staining of the walls of the lateral ventricles, among other other brain regions (see also Figure 3.1). We wanted to characterise, at the single-cell level, the Penk-tdTomato^{+} cells in the striatal

parenchyma as well those in the walls of V-SVZ. Following characterisation and validation, we also wanted to understand whether these cells could be recruited in disease states, such as Parkinson's disease (described below).

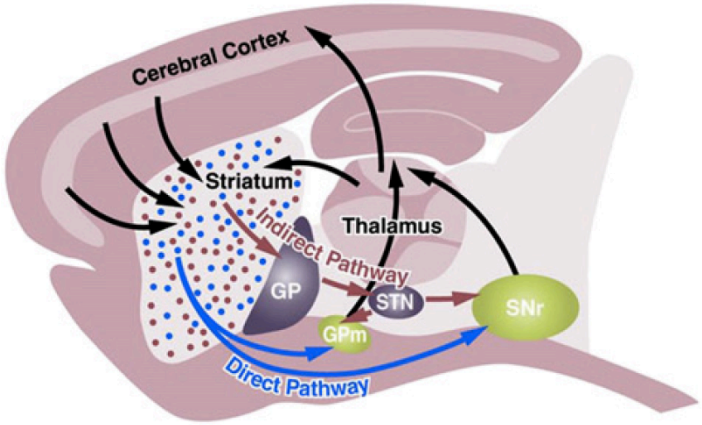


Figure 1.4. A simplified schematic of the direct and indirect pathways that originate in the striatum. The diagram indicates the glutamatergic innervation of the striatum by the cerebral cortex. The striatum then sends GABAergic projections to the output nuclei of the basal ganglia—entopeduncular nucleus (EPN) (globus pallidus internal segment [**GPm**] in primates and humans) and the substantia nigra pars reticulata (**SNr**)—via two distinct pathways. The output nuclei send inhibitory GABAergic projections to the thalamus. The first is labeled in ***blue*** and is called the direct or striatonigral pathway. It is called direct, because it directly sends projections to both output nuclei without requiring any intervening structures. The direct pathway leads to the inhibition of the inhibitory neurons, thus leading to disinhibition of the thalamus, and thus "promotes" movement-related actions. The indirect pathway is labeled in ***maroon*** and is so-called because it first sends projections to the globus pallidus external segment (**GP**), which then sends projections to the subthamalic nuclei (**STN**), which then sends excitatory glutamatergic projections to the two output nuclei. The GP usually inhibits the STN, but when the GP is inhibited, the STN is released from inhibition. When this occurs it further excites the EPN and SNr, leading to the inhibition of the thalamus. Thus the indirect pathway is said to "inhibit" movement-related actions. Adapted from Kreitzer and Malenka (2008).

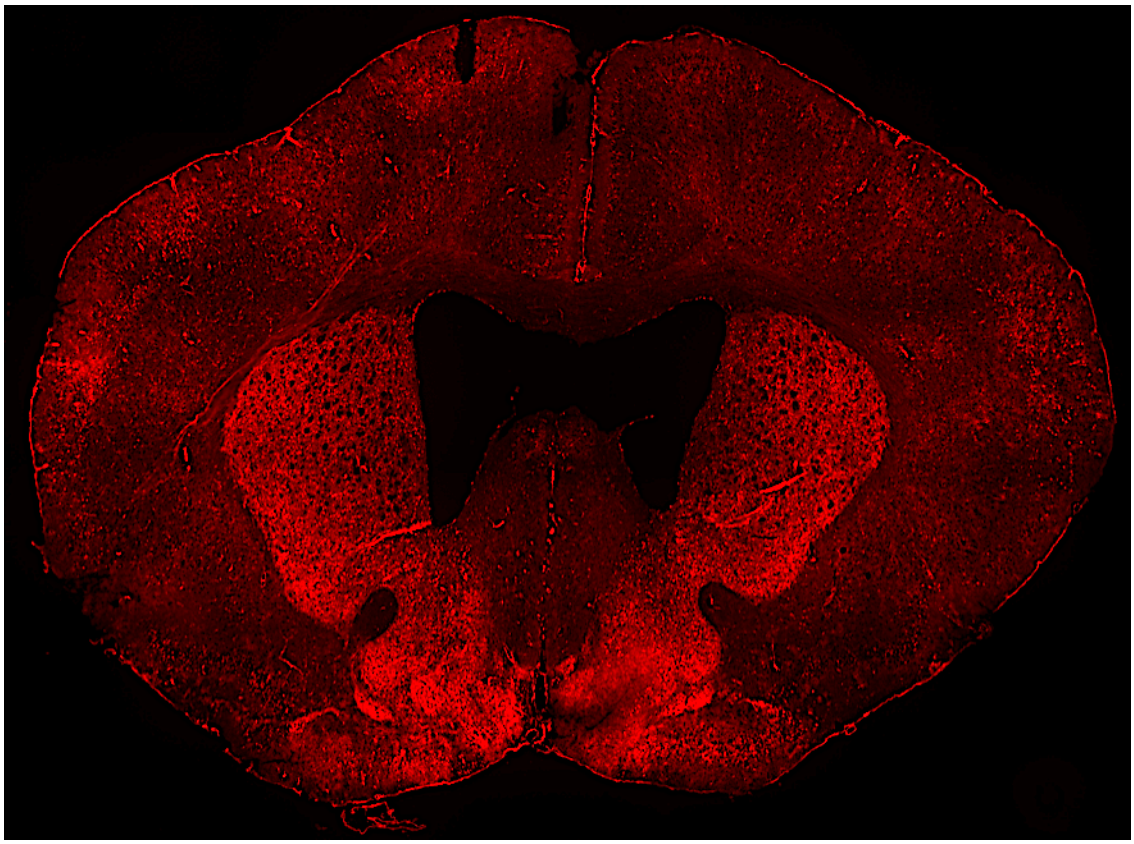


Figure 1.5. tdTomato Reporter expression in the IPNK line. A coronal brain section from an 8 week old IPNK mouse was mounted and imaged for endogenous tdTomato signal (no antibodies used). Strong staining can be detected in the striatal parenchyma, as predicted. Strong expression is also seen in the walls of the lateral ventricle, the third ventricle, and hypothalamus. Weak staining can be seen in the cortex as well other brain structures.

Neurogenesis and neurodegenerative diseases

1.10 An overview of Parkinson's disease.

Parkinson's disease (PD) is the most common movement disorder and the second common neurodegenerative disorder (de Lau and Breteler, 2006; Winner and Winkler, 2013). PD presents in many different ways, but the main motor symptoms are bradykinesia, rigidity, and a resting tremor (de Lau and Breteler, 2006). Non-motor symptoms include depression, anxiety, and cognitive defects, with 80% of patients in advanced stages developing dementia (Winner and Winkler, 2008). Intriguingly, deficits in olfaction are another key non-motor symptom—possibly implicating V-SVZ neurogenesis. In a manner similar to HD, the dysfunctional movements are thought to arise from altered signalling between the basal ganglia nuclei. The signaling is altered because the dopaminergic neurons of the substantia nigra *pars compacta*—which send dense projections to the striatum—degenerate, with approximately 60% of them being lost at the time of symptom onset (Dauer and Przedborski, 2003). This loss of dopaminergic input into the striatum, which normally regulates the dopamine receptor-expressing neurons of the direct and indirect pathway, accounts for the initial appearance of motor symptoms. Although some short-term drug therapies such as L-DOPA exist, there are currently no long-term therapies or cures (Winner and Winkler, 2008; Stoker *et al.,* 2018).

Though the etiology of PD is still yet to be completely understood, it seems that it arises due to a combination of genetic and environmental factors. Genetic factors are thought to account for 50% of early-onset cases, compared to only 2-3% of late-onset cases (Zeng *et al.,* 2018). The key neuropathological manifestation of the disease is the accumulation of alpha-Synuclein intracellular deposits in Lewy bodies and Lewy neurites (Spillantini *et al.,* 1998). When genetic factors are found to lead to PD, the mutations are mainly found in the alpha-Synuclein gene. The most common mutations being the A53T and A30P substitutions (Dauer and Przedborski, 2003).

Mutations in other genes such as *LRRK2*, *PINK1*, *Parkin*, *DJ-1*, *VPS35*, and *GBA1* are also associated with PD (Zeng *et al.,* 2018). Interestingly, it has also been found that lysomomal storage disorder genes are over-represented in PD (Robak *et al.,* 2017). This is intriguing because of the findings of Leeman *et al.* (2018) suggesting that lysosomal dysfunction traps qNSCs in a quiescent state—suggesting a possible link between the V-SVZ NSC niche and PD.

1.11 The role of V-SVZ neurogenesis in Parkinson's disease.

Unlike in HD, there exist contradictory reports regarding the roles of V-SVZ proliferation and neurogenesis in PD. It was first reported that proliferation in the V-SVZ was decreased in the brains of human PD patients (Hoglinger *et al.,* 2004). This finding was confirmed by a later study that saw a decrease in the number of $EGFR^+$, likely C-type, cells in the V-SVZ (O'Keeffe *et al.,* 2009). However, a more recent study has stated that proliferation in not compromised in human PD patients (van den Berge *et al.,* 2011). In this particular study, the patients were roughly ten years older than in the Hoglinger *et al.* study. Many factors, including, but not limited to, patient ages, method of collection, and staining could have led to this discrepancy. Nonetheless, this highlights the need to investigate V-SVZ proliferation and neurogenesis in models of PD.

Though there are genetic models of PD, the most popular means of modelling this condition rely on the use of neurotoxins such as 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), 6-hydroxydopamine (6-OHDA), paraquat, and retenone (He and Nakayama, 2009). The validity of the rodent models of PD are a point of concern as PD does not naturally present in rodents and must be simulated using the compounds listed above. None alone can create the exact temporal profile of the human disease nor can they recreate exactly cells and tissues are affected, thus models must be chosen with specific experimental questions in mind to investigate facets of the disease (Antony *et al.,* 2011).

MPTP was discovered when it was found that recreational drug preparations containing MPTP caused irreversible Parkinsonism in human users (Blume, 1983). In rodents, MPTP effectively reduces striatal dopamine levels and causes the degeneration of the substantia nigra *pars compacta* (Burns *et al.,* 1983). MPTP becomes cytotoxic when it is converted to MPP+ (1-methyl-4-phenylpyridinium) and taken up by nerve terminals.

It was first reported that MPTP administration led to a two-fold increase cell proliferation in nigral dopaminergic neurons in mice (Zhao *et al.,* 2003). The same study also found that there was neurogenesis in the substantia nigra, and that these nascent nigral neurons were derived from V-SVZ precursors (Zhao *et al.,* 2003). This finding has since been confirmed in mice, rats, and macaques by other groups (Shan *et al.,* 2006; He and Nakayama, 2009). However, reports that MPTP administration leads to a decrease in neurogenesis also exist. Hoglinger *et al.,* report that MPTP administration in mice mirrors what they report in human PD patients—in that proliferation is decreased. A similar decrease was also reported when MPTP was administered in macaques (Freundlieb *et al.,* 2006).

The recent finding that MPTP can also target migrating neuroblasts and thus that it does not selectively target nigral dopaminergic terminals, can explain some of these discrepancies (He *et al.,* 2006). Therefore, MPTP may directly decrease neurogenesis through a mechanism that does not actually mimic PD. Furthermore, the reported increase in neurogenesis may be explained as a response to the direct toxic effect of MPTP on the V-SVZ niche, and not as a response to PD-associated neurodegeneration (He and Nakayama, 2009).

The stereotactic administration of 6-OHDA is a popular method of inducing a Parkinsonian phenotype in mice, and is even considered the “gold-standard” model to analyse PD in animal models (Smith and Heuer, 2012). 6-OHDA cannot permeate the blood-brain barrier and thus needs to be injected directly into the desired target areas. There are some major advantages to using 6-OHDA in place of MPTP. The first being that MPTP can directly affect neurogenesis, by

targeting neuroblasts for cell death, thus confounding the analysis of the effects of Parkinson's on neurogenesis (He and Nakayama, 2009). Second, MPTP can have unstable effects depending on species, gender, strain, and age, whereas 6-OHDA is said to have a more stable and reproducible effect (Smith and Heuer, 2012). Third, MPTP is highly toxic and thus requires stringent protocols, handling methods, facilities, and permission, making it unfeasible for many laboratories. Finally, 6-OHDA creates a unilateral lesion, meaning that the un-injected contralateral side can be used as an internal control (Smith and Heuer, 2012).

6-OHDA is a structural analogue of the catecholamines, dopamine and noradrenaline. Thus, it only exerts its cytotoxic effect when taken up by catecholaminergic neurons (Perese *et al.,* 1989). The 6-OHDA neurotoxin can be injected directly into the medial forebrain bundle to cause an acute loss of the dopaminergic projection neurons of the substantia nigra *pars compacta*. The total loss of neurons can be seen as early as 24 hours and leads to a severe phenotype, and can even result in death in mice (Smith and Heuer, 2012). A safer and more progressive model of PD is generated by the direct injection of 6-OHDA into the dorsal striatum, where it is taken up by the pre-synaptic terminals of nigral dopaminergic neurons, subsequently causing their degeneration. In this paradigm, the lesion requires one to four weeks to develop.

As with MPTP models, the findings regarding the effect of 6-OHDA-mediated striatal dopaminergic deafferentiation on proliferation and neurogenesis are conflicting (see ***Discussion*** for an in-depth analysis). Several groups report an increase in V-SVZ cell proliferation in rodents, in response to 6-OHDA injection (Liu *et al.*, 2006; Parish *et al.*, 2007; Aponso *et al.*, 2008). While other groups report the opposite (Baker *et al.,* 2004; Winner *et al.,* 2006). Baker *et al.* (2004) report a decrease in neurogenesis, but Winner *et al.* (2006) reported that even though there was significant decrease in proliferation, this was accompanied by an increase in neurogenesis. While Liu *et al.* (2006) and Parish *et al.* (2007) report a significant increase in neurogenesis concomitant to the increase in proliferation. Finally, Aponso *et al.* (2008)—who did report an increase in

proliferation—do not find evidence of increased neurogenesis. Thus, the V-SVZ proliferation and neurogenic response upon 6-OHDA-mediated striatal lesioning remains unclear and in need of further analysis.

An overview of methods used to perform single cell analysis

1.12 Fluorescence-activated cell sorting.

Fluorescence-activated cell sorting (FACS) is a now-commonly used method to quantify and isolate cells based on the expression of particular cellular markers. At its simplest, the method relies on illuminating cells as they flow singly past a light source, and where this signal is then detected and used to decide whether to collect or discard the cell (Givan, 2004). Running FACS experiments first requires obtaining a suspension of separated cells. Thus, it can require the enzymatic digestion and trituration of tissues (Nath *et al.,* 2017; Vatanashevanopakorn *et al.,* 2017). This suspension is then run through the machine such that cells flow past the illuminator one at a time. Modern FACS instruments use lasers to illuminate samples at specific wavelengths—which is essential when using fluorophore-conjugated antibodies against desired cellular markers. Two lenses act as detectors and record both light scattering and fluorescence data. One lens is in the “forward” direction, while the other is in the orthogonal or “side” direction. The “forward” scatter is used as a proxy of the size or volume of the cell passing though the detection point. The “side” scatter primarily indicates the existence of irregularities on the surface of the cell or in the cytoplasm, thus it is used as a proxy for cell complexity (Givan, 2004). As mentioned above, the detectors can also pick up fluorescence emitted from samples in response to laser excitation. The experimenter uses the detected data to decide whether to collect the cell or discard it. This is done by passing the single cells between two high-voltage charge deflection plates and providing a charge to the cell droplet to either guide it to a collection vessel or a waste

container (Givan, 2004). Using this method, one can separate and collect cells on the basis of their expression of markers of interest.

1.13 High-throughput, microfluidic gene-expression analysis.

The gold-standard, and the now-ubiquitous, method for analysing gene expression is real-time quantitative reverse transcription polymerase chain reaction (RT-PCR). This is of interest because it helps experimenters to analyse how cells, or samples of cells change the expression of selected genes in response to changing experimental conditions. It can also be used as a means of validating RNA-sequencing data. It is now possible to perform these reactions in a high-throughput manner, using microfluidic platforms, to analyse the expression of up to 96 genes in up to 192 cells or samples, rapidly and simultaneously. This is now commonly done through combining the TaqMan Assay system with the Fluidigm microfluidic platform (Drissen *et al.,* 2016; Nath *et al.,* 2017).

The TaqMan system relies on the use of fluorogenic probes paired with conventional forward and reverse primers used in routine PCRs. On the 5' end, the probes contain a 6-carboxyfluorescein (FAM) fluorescent probe and a non-fluorescent quencher at the 3' end along with a minor groove binding moiety (Held *et al.,* 1996; Osman *et al.,* 2007). Thus, the probes are quenched when they are intact. *Taq* DNA polymerase has intrinsic 5'→3' exonuclease activity, which is key to TaqMan chemistry. The exonuclease activity removes the non-fluorescent quencher on the 3' end, thus allowing the FAM probe on the 5' end to fluoresce (Holland *et al.,* 1991). Thus, every time a primer is utilised by *Taq* polymerase, fluorescence will increase. The proportional increase in fluorescence can, then, be used as a proxy for the number of amplicons that are generated in the PCR reaction. The Fluidigm microfluidic platform allows one to perform this reaction on a miniaturised scale and across many cells and genes. The actual reaction is performed inside the BioMark instrument, where the fluorescence from each microfluidic well is

recorded and converted to the classic *ΔCt*-values seen in quantitative RT-PCR analyses (Drissen *et al.,* 2016; Nath *et al.,* 2017). The *ΔCt*-values can then be normalised to reference genes to obtain meaningful data regarding the expression of genes of interest across different samples or treatments. Specifics regarding the calculations can be found in ***Materials and Methods*.**

1.14 Single cell RNA-sequencing.

Single cell RNA-sequencing (RNAseq) has revolutionised the life sciences. It has allowed researchers to determine, with a relatively high degree of accuracy, the whole transcriptome of a cell at the time of sequencing. The technology has advanced to such a degree that the Human Cell Atlas project is seeking to sequence 30-100 million cells (Hwang *et al.,* 2018). Here, I provide a brief overview of the general concepts underlying modern RNA-sequencing protocols and also compare the two most popular sequencing methods—Smart-seq2 and DropSeq (Figure 1.6).

Both variants rely on the same method and platform to perform the actual sequencing, they only differ in the way the cells or samples are prepared prior to this sequencing. Thus, I will first describe the method by which the sequencing, proper, is done. Both methods rely on the Illumina platform to perform "genome sequencing using reversible terminator chemistry", or "sequencing by synthesis" (Bentley *et al.,* 2008). During sequencing, cDNA—synthesised from the full complement of captured and viable mRNA within the cell at the time of FAC sorting and lysis—is denatured into single-stranded DNA which is then immobilised on the glass surface of a flow-cell. The flow-cell is a proprietary reaction chamber on which the sequencing of the immobilised cDNA will be performed. To ensure accuracy, each cDNA fragment is sequenced multiple times, which involves iterative cycles of single-nucleotide extension by DNA polymerases. However, to perform this method in an orderly manner and to be able to gather actual sequencing data, one must perform the single-nucleotide extension in one direction and move progressively. This is accomplished by using special nucleotides that contain reversible

terminators (2'-O-azidomethyl 2'-deoxynucleoside triphosphates [A, C, G and T]) and different removable fluorophores attached to the terminators. The fluorescence signal is used a proxy to perform the base-calling. The nucleotides with the terminator can be ligated as normal, but the terminator will prevent the unwanted ligation of the next nucleotide (complementary to the template cDNA) such that base-calling can be performed in an orderly and directed manner. After the base-calling is complete and the data has been recorded, the fluorophore and terminator are removed by the addition of tris(2-carboxyethyl)phosphine (TCEP). The TCEP exposes the hydroxyl group necessary for the next incoming base to create a nucleotide bond. The TCEP is then washed out and the nucleotide+terminator solution is added once again so that the polymerase can add the next nucleotide. This is done iteratively until sequencing of the whole cDNA fragment is complete. The Illumina instrument automatically assigns a quality-score to each base-call, which is important for quality control assessments during later bioinformatics analysis (Bentley *et al.,* 2008).

Where Smart-seq2 and DropSeq differ is in the way the cDNA, generated from all the mRNA inside of a cell, are indexed or tagged. In other words, they differ in the way in which they keep track of which reads are coming from which cell. Smart-seq2 is a form of deep sequencing. This means that one can obtain up to 4-5 million reads per cell, depending on the make of the flow-cell and Illumina sequencer used in the experiment. The high read depth means that the experimenter has a higher chance of sequencing all of the mRNA within a cell—even the ones that are present in only a few copies. Smart-seq2 was developed at the Karolinska Institute to allow experimenters to use standard reagents to perform accurate deep sequencing (Picelli *et al.,* 2014). The method first involves reverse transcription by a reverse transcriptase, and subsequent amplification of the cDNA. The cDNA is then fragmented and tagged by Tn5 transposases in a process called tagmentation. The key process is the manner in which the tagmented cDNA is indexed. Special indices (i5 and i7) are attached to the tag that is added by the Tn5 transposase.

There are 8 i5 primers and 12 i7 primers, meaning that each well of a 96-well plate can be uniquely tagged, when the primers are used in combination. The tagged cells/samples are then sequenced on a flow-cell in the manner described above (Picelli *et al.,* 2014). Due to the limited number of indices, the Smart-seq2 method can only be used to sequence a relatively small number of cells. However, it does ensure a considerably higher read depth than the DropSeq method described below.

DropSeq was developed at the Broad Institute and allows for sequencing of thousands of cells, but at a much lower read depth (Macosko *et al.,* 2015). DropSeq relies on microfluidics and molecular barcodes to tag the cells prior to sequencing. A single-cell suspension is placed in one channel of a microfluidic device, with beads coated with unique primers/barcodes in another channel. One cell and one bead (coated in the same primer/molecular barcode) are then injected simultaneously towards another channel containing an oil solution. When the cell and the bead reach the oil solution they are encapsulated into a single oil droplet. The cells within the droplets are lysed, whereupon the unique primers bind to the free mRNA, thus barcoding the cell. The cDNA is then prepared and amplified in preparation for sequencing. Since an infinite number of barcodes can be generated, there is no upper limit to the number of cells that can be sequenced at one time. However, since the Illumina platform can only perform a finite number of reads this means that there is a negative correlation between sample size and read depth. Usually, the read depth for a DropSeq experiment will vary between 25,000-50,000, which is, at best, a 10-fold decrease in depth compared to Smart-seq2 (Macosko *et al.,* 2015; Hwang *et al.,* 2018). Also, due to the high-throughput nature of the method, there is a high attrition rate for cells because they may get damaged in the microfluidic device, they may not be incorporated into droplets, or the droplets may contain doublets or triplets.

Thus, the sequencing method that is required will depend on the nature of the experiment and the question being asked by the experimenter. In our case we were interested in sequencing

single cells from the striatal parenchyma (which are easy to isolate) and also cells from the lateral wall of the lateral ventricle (which are few in number and thus probably quite difficult to isolate). As mentioned above, DropSeq allows for the sequencing of large number of cells at low depth but it also has a high attrition rate of isolated cells due to the harsh nature of the microfluidic devices. Therefore, we reasoned that Smart-seq2 would be a more appropriate method for our experiment because we were in search of rare cells as well those more commonly found. Thus, the deep sequencing inherent to Smart-seq2 and relatively lower attrition rate of the method would more likely result in more successful sequencing of the cells from the V-SVZ.

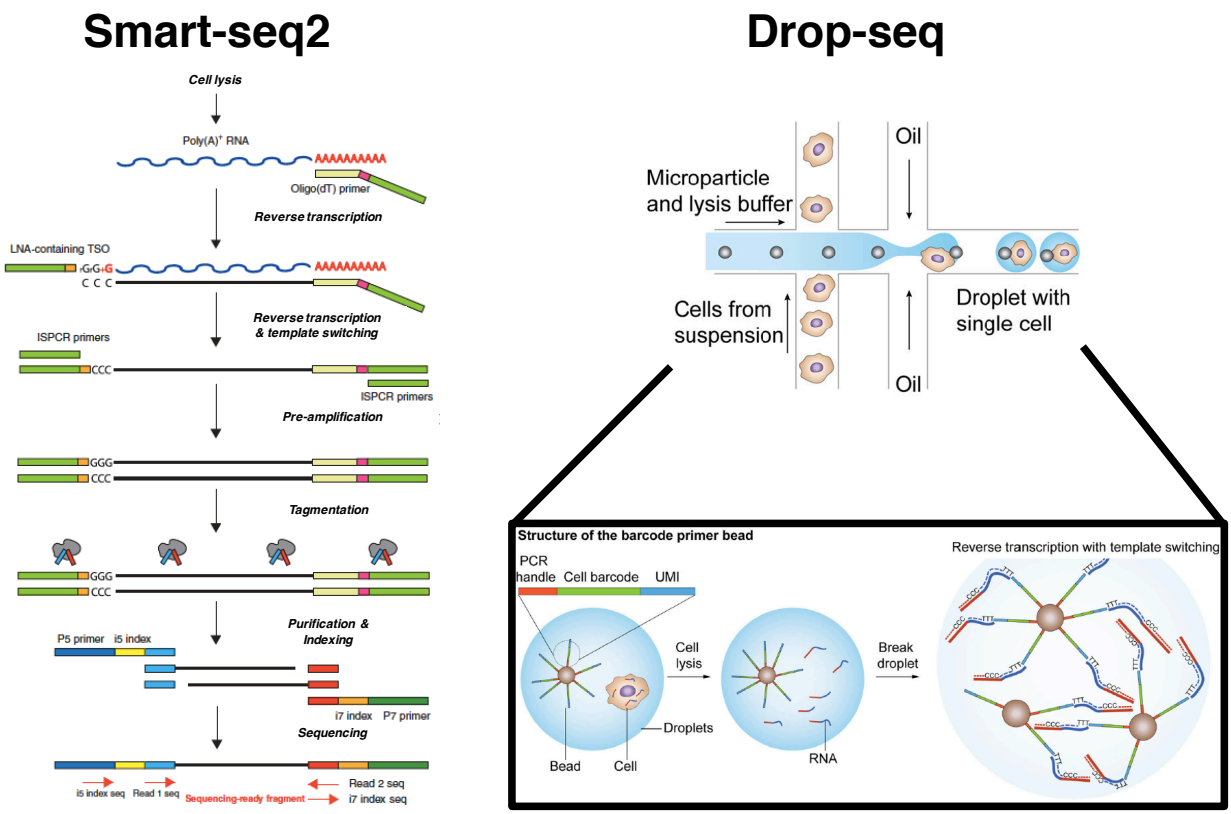


Figure 1.6. Summary of sequencing methods. An illustrated summary of the two major sequencing methods described in the ***Introduction***. The Smart-seq2 method is a method of deep sequencing developed at the Karolinska Institute (Picelli *et al.*, 2013). The major steps are first stand synthesis by reverse transcription, followed by pre-amplification and tagmentation. Cells are indexed by using the i5 and i7 indices, which is followed by sequencing on an Illumina platform such as the NextSeq. The Drop-seq method was developed by Macosko *et al.* (2015) and allows for the sequencing of thousands of cells at a lower depth than Smart-seq2. The illustration summarises microfluidic-based method of barcoding cells of interest. Sequencing proper is performed in the same manner as Smart-seq2.

Chapter 2 - Materials and Methods

2.1 Mouse lines.

Mouse Line	Description	Reference	Source
BAC-*Penk-Cre*	A BAC transgenic line expressing Cre-recombinase under the control of pre-pro-enkephalin promoter.	Besusso *et al.* (2013)	Minichiello Lab
Rosa26-tdTomato	A reporter line expressing tdTomato from the Rosa26 locus upon Cre-mediated removal of LoxP-flanked stop cassette.	Madisen *et al.* (2010)	The Jackson Laboratory
"***IPNK***": BAC-Penk-Cre$^{tg/+}$;Rosa26-tdTomato$^{tg/+}$	A crossed BAC transgenic line irreversibly expressing tdTomato in cells with an active pre-pro-enkephalin promoter. (Resulting from crossed lines BAC-*Penk-Cre and* Rosa26-tdTomato mentioned above).	Vatanashevanopakorn *et al.* (2017)	Minichiello Lab
Z/EG-eGFP	A reporter line expressing eGFP upon Cre-mediated removal of LoxP-flanked stop cassette.	Novak *et al.* (2000)	The Jackson Laboratory
Rosa26-EYFP	A reporter line expressing EYFP from the Rosa26 locus upon Cre-mediated removal of LoxP-flanked stop cassette.	Srinivas *et al.* (2001)	The Jackson Laboratory

All animals were kept on a mixed genetic background (C57BL/6J:129). Colonies were maintained by the Animal Technologists from Biomedical Services, Oxford University; and managed by my supervisor, Dr. Liliana Minichiello. Animals were housed and maintained according to UK legislation Animals (Scientific Procedures) Act 1986, as well as the University of Oxford Ethical Review Committee policy, and to the National and International laws and policies (EEC Council Directive 86/609, OJ L 358, 1, December 12, 1987; NIH Guide for the Care and Use of Laboratory Animals, NIH Publication No. 85-23, 1985 revised in 1995).

2.2 Solutions

0.1M Phosphate Buffer (PB; pH = 7.4): Na_2HPO_4 (1 M) 77.4 ml of, NaH_2PO_4 (1M) 22.6 ml for 1 litre)

Phosphate-buffered Saline (PBS): Phosphate buffer (0.01 M), KCl (0.0027 M) and NaCl (0.137 M), pH = 7.4 (prepared from Thermo Fisher Scientific PBS tablets, Cat. No. 18912014)

4% Paraformaldehyde (PFA) solution (1 litre): Paraformaldehyde (40 g) (Sigma, Cat. No. P6148) dissolved by continuous stirring in 800 ml 0.1M PB buffer (pH 7.4) by heating, keeping the temperature below 60 °C. The solution was cooled down to room temperature, adjusted to pH 7.4 with concentrated NaOH, and filtered through 0.2 micron filter. Aliquots were stored at -20°C. The aliquots were thawed slowly in a 4°C fridge when required for experiments.

Blocking buffer (*for fluorescence*): Fish gelatine (5%; Sigma, Cat. No. G7765) and 0.5% Triton-X (Sigma, Cat. No. T8787) in PBS

Blocking buffer (*for 3,3′-Diaminobenzidine [DAB]*): Normal goat serum (10%; Sigma, Cat. No. G9023), 0.3% carrageen (Sigma, Cat. No, 22048), 0.5% TritonX-100, 0.2% sodium azide in TBS.

Lysis buffer (Smart-seq2; per well): Triton (0.4%) + RNase inhibitor (2 μl), dNTP's (10mM; 1 μl; Thermo Fisher, Cat. No. 10319879), Oligo dT_{30VN} (10 uM; 1 μl; Nextera Kit).

Lysis buffer (microfluidic gene expression analysis; per well): CellsDirect 2x Reaction Mix (5 μl), Ambion SUPERase 12-In (0.1 μl), Tris/EDTA (TE) buffer (0.9 μl).

Borate buffer (1M): Boric Acid (H_3BO_3) (61.83g; 61.83 g/mol) + NaOH (10g; mw: 39.997 g/mol) in 1 litre dH_20.

Sucrose artificial cerebrospinal fluid (aCSF): Sucrose (75 mM), NaCl (87 mM), $NaHCO_3$ (25 mM), KCl (2.5 mM), $NaH_2PO_4{\cdot}H_2O$ (1.25 mM), $CaCl_2{\cdot}2H_2O$ (0.5 mM), $MgCl_2{\cdot}6H_2O$ (7 mM), Glucose (25 mM).

2.3 Antibodies

Primary Antibody	Host Species	Supplier	Cat. No.	Concentration
BrdU	Rat	Novus	NB500-169	1:200
CD11b-APC	Mouse	BioLegend	101211	1:400
CD133	Rat	Ebioscience	14-1331-80	1:500
CD24	Mouse	Ebiosceince	14-0247-80	1:100
Dcx	Rabbit	abcam	ab18723	1:200
Dlx5	Mouse	Santa Cruz	sc-398150	1:100
GFAP	Rabbit	Dako	GA524	1:500
Gja1	Mouse	Santa Cruz	sc-271837	1:100
Gjb6	Mouse	Santa Cruz	sc-514847	1:100
GLAST	Guinea Pig	Merck Millipore	AB1783	1:500

Kcnj13/KiR7.1	Mouse	Santa Cruz	sc-398810	1:100
Ki67	Rabbit	abcam	ab16667	1:5000
KiR7.1-ATTO488	Rabbit	Alomone Labs	APC-125-AG	1:50
NeuN	Rabbit	abcam	ab104225	1:500
Penk	Rabbit	Neuromics	RA14124	1:1000
Prlr	Mouse	Novus	NB300-561AF488	1:100
Sox11	Guinea Pig	Dr. Elisabeth Sock	--	1:2000
Sox9	Rabbit	abcam	ab185230	1:500
TH	Mouse	Millipore	Mab318	1:1000
Tuba1a	Mouse	Santa Cruz	sc-58666	1:100
Tubb3	Mouse	R&D Systems	MAB1195	1:1000

Secondary Antibody	**Supplier**	**Cat. No.**
Alexa Fluor 488 Donkey Anti-Goat IgG (H+L)	Invitrogen (Thermo Fisher)	A11055
Alexa Fluor 488 Goat anti-Guinea Pig IgG (H+L)	Invitrogen (Thermo Fisher)	A-11073
Alexa Fluor 488 Goat Anti-Rat IgG (H+L)	Invitrogen (Thermo Fisher)	A-11006
Alexa Fluor 488, Goat anti-Mouse IgG (H+L)	Invitrogen (Thermo Fisher)	A-11029
Alexa Fluor 555 Donkey anti-Mouse IgG (H+L)	Invitrogen (Thermo Fisher)	A31570
Alexa Fluor 555, Donkey anti-Rabbit IgG (H+L)	Invitrogen (Thermo Fisher)	A-31572
Alexa Fluor 647 AffiniPure Donkey Anti-Guinea Pig IgG (H+L)	Jackson Immunoresearch	706-605-148
Alexa Fluor 647 AffiniPure Donkey Anti-Rat IgG (H+L)	Jackson Immunoresearch	712-605-153
Alexa Fluor 647 Donkey anti-Goat IgG (H+L)	Invitrogen (Thermo Fisher)	A21447
Alexa Fluor 647 Donkey Anti-Mouse IgG (H+L)	Invitrogen (Thermo Fisher)	A31571
Alexa Fluor 647 Donkey Anti-Rabbit IgG (H+L)	Invitrogen (Thermo Fisher)	A31573

2.4 Preparation and perfusion of animals for immunochemistry.

Mice were anaesthetised with 50 µl of sodium pentobarbital at concentration of 200 mg/ml (Euthatal; Sigma, Cat. No. P3761-5G) by intraperitoneal injection and depth of anaesthesia was checked by pinching the paw and the base of the tail. The mice were considered anaesthetised when they were completely unresponsive but still breathing. All brains used were prefixed by transcardiac perfusion with 20 ml ice-cold 4% PFA (paraformaldehyde) preceded by clearing of blood by circulating 20 ml 0.1M PB. Cardiac perfusion was performed by opening the chest wall,

and inserting a needle in the left ventricle and then piercing the right atrium with small scissors. Then the 0.1M PB was perfused followed by ice-cold 4% PFA. This was performed using perfusion tubing attached to a 30G needle and a perfusion pump at a rate of 3-4 ml/min (Masterflex L/S easyload II, Cole Parmer) to slowly perfuse the solutions through the left ventricle of the animals. The animals were then decapitated and their brains dissected out. The brains were then post-fixed for 3 hours in 4% PFA at room temperature. They were then washed with 0.1M PB and cryoprotected. This was done by using successive sucrose solutions: 10% overnight, 20% overnight, 30% overnight (all made up in 0.1M PB; 4˚C). The brains were then embedded in a plastic mold containing OCT (Fisher Scientific, Cat. No. 23-730-571) and flash frozen in isopentane cooled in dry ice until the OCT hardened. Brains were stored in the -80°C freezer.

2.5 Cryosectioning.

All sections were made using a Leica Cryostat (CM3050S). The temperature of the chamber was set at -23°C and the objective was at -18°C. In preparation for sectioning, the brains were taken from the freezer and equilibrated within the Cryostat chamber to bring the temperature up to make the brains amenable to sectioning. Brains were oriented appropriately to either obtain coronal or sagittal sections. The cryostat was then set to section at the appropriate thickness. All images in this thesis were from sections of 30 μm thickness. The sections were quickly incubated in cold 0.1M PB to rinse off excess OCT and then transferred for long-term storage into 48-well plates containing 0.1M PB and 0.01% sodium azide at 4˚C.

2.6 Fluorescent immunochemistry.

Appropriate sections were selected and placed in 500 μl PBS in a 24-well plate and washed 3 times at 10 minutes per wash on a rocker (20 rpm) at room temperature. Sections were

then blocked in blocking solution (5% fish gelatine and 0.5% Triton-X in PBS) for 1 hour at room temperature. Antibodies were diluted in blocking solution at the appropriate concentrations. Sections were incubated in primary antibodies overnight at 4°C on a rocker. Primaries were washed off with PBS in 0.1% Triton-X in 3 washes of 10 minutes each. They were then incubated with secondary antibodies (in blocking solution, always at 1:1000) for 2 hours at room temperature. Sections were then washed 3 times with PBS, 10 minutes per wash. The sections were stained with DAPI (1:10,000 in PBS) for 30 minutes at room temperature. They were then washed 3 times with PBS, 10 minutes per wash. The sections were mounted onto SuperFrost slides (Fisher Scientific, Cat. No. 12372098), they were then dried and mounted with VectaShield (Vector Laboratories, Cat. No. H-1000) with a rectangular coverslip (Thermo Fisher, Cat. No. 22X30-1). The slides were then sealed with nail polish and allowed to dry before imaging.

2.7 Immunohistochemistry.

Sections were treated with 3% hydrogen peroxide (in TBS) at room temperature for 20 minutes, followed by 3 washes (10 minutes each) with TBS. Tissues were incubated in tissue blocking buffer 1 (10% normal goat serum, 0.3% carrageen, 0.5% TritonX-100, 0.2% sodium azide in TBS) at room temperature for 1 hour. Tissues were then incubated with the primary antibody (chicken anti-GFP, Abcam, 1:1,000) diluted in 1% normal goat serum, 0.3% carrageen, 0.5% TritonX-100 in TBS at room temperature overnight. After 3 washes with TBS, sections were incubated with the secondary antibody, biotinylated goat anti-chicken (1:200, Novex) diluted in 1% normal goat serum, 0.3% carrageen, 0.5% TritonX-100 in TBS at room temperature for 2 hours. After 3 washes, sections were incubated in Vectastain® ABC solution (Vector Laboratories, Cat. No. PK-6100) at room temperature for 30 minutes. Tissues were washed 3 times (15 minutes each) with TBS followed by 2 washes (15 minutes each) with TB. Tissues were then incubated in DAB solution for 1-2 minutes. When the colour developed, the reaction was stopped by adding

cold TB, and quickly rinsing several times. Sections were mounted on SuperFrost slides and air-dried in a fume hood for 2 hours. Mounted sections were dehydrated in xylene for 10 minutes. Vectamount permanent mounting medium (Vector Laboratories, Cat. No. H-5000) was then applied on the slide before cover slip was placed. The slide was then left to plasticise overnight.

2.8 Fluorescence microscopy.

The sections were fluorescently imaged with an upright widefield epifluorescence microscope from Leica (DM6000B) and images were acquired using monochrome camera (DFC365FX, Leica). For confocal imaging a Zeiss 780 upright LSM confocal microscope with MP laser was used. The following objectives were used for image acquisition: HCX PL APO 10x/0.40 CS, HC PL 20x/0.75 CS2 and HCX PL APO 40x/0.85 CORR CS. The images were then exported, edited, and quantified using the open-source FIJI software (NIH).

2.9 Penk-tdTomato$^+$ cell counting and estimation.

Three female IPNK brains 9 weeks of age were used in this study. Coronal sections of 30 μm thickness, spanning the entire striatum (3775 μm) were obtained. One in every 21 sections was used for antibody staining and cell counting, for a total of six coronal sections per antibody (anti-GFAP; anti-Sox11; anti-Ki67; anti-CD133). The sections began at Bregma Level: 1270 μm and terminated at Bregma Level: -1730 μm. The sections were treated as described above for fluorescent immunochemistry.

For the estimation of the number of Penk-tdTomato$^+$ cells at the lateral wall of the lateral ventricle, we derived the following formula:

$$Total\ number\ of\ cells\ across\ the\ striatum = \frac{l}{C \times n} \times N$$

where ***N*** is the total number of cells counted in all sections of each mouse, ***l*** is the length of the striatum (3775 μm), ***C*** is the cross-sectional width of each section used (30 μm) and ***n*** is the number of sections used (6) in the analysis.

2.10 Tissue processing for FACS.

Mice were anaesthetised as above, however instead of 0.1M PB and 4% PFA, 20 ml of artificial cerebrospinal fluid made up in sucrose was perfused. The brains were then dissected out in ice-cold platform. Sub-dissection of both striata was performed by opening the skull of the mouse and taking out the brain, this was then followed by separating both hemispheres and removing overlying cortex and hippocampus. The striatum is then dissected out and cleaned of corpus callosum and any remaining cortical tissue, this was followed by manual chopping of the tissue. The chopped tissue was placed in an Accutase (Stem Cell Technologies, Cat. No. 07920) enzyme solution to dissociate cells, for 30 minutes at 4˚C on a rotator. The digested tissue was then centrifuged for 2 min at 450 *g* and the pellet was resuspended in 1 ml of ice-cold Hibernate A (Brain Bits, Cat. No. HA). This was then triturated 10 times with a large-diameter (~1.3 mm) Pasteur pipette. 600 μl of the cloudy suspension containing dissociated cells was transferred to a 15 ml Falcon tube on ice. Another 600 μl of ice-cold fresh Hibernate A was added to the original Eppendorf tubes. The trituration steps were repeated two more times as above with medium- (~0.8 mm) and small-diameter (~0.4 mm) Pasteur pipettes. The last of the cells were collected from the original tube by adding 750 μl of ice-cold Hibernate A, inverting it several times and letting settle on ice. The cell suspension was then crudely filtered using 100 μm and 40 μm cell strainers.

A Percoll (Sigma, Cat. No. P1644) density gradient was set up as shown below:

Layer	Percoll (µl)	Hibernate A (µl)	1 M NaCl (µl)	Total (µl)
1 (bottom)	206.13	856.5	24.45	1087.08
2	162.63	900	19.13	1081.75
3 (top)	120.08	942.5	14.88	1077.45

The cell suspension was gently placed on the top layer. The gradient was centrifuged 430 *g* (2400 rpm) for 5 min at 4°C (Beckman-Coulter ultracentrifuge; SW41 rotor, acceleration at max and deceleration = 0). The top layer contains cellular debris and was aspirated out. The remaining layers were gently mixed by pipetting. The mixed gradient was placed back into the centrifuge to obtain a cell pellet: 5 min at 4°C at 550 *g* (2700 rpm). All the supernatant was removed and the pellet was resuspended in 1 ml Hibernate A. Cell counting was performed and LIVE/DEAD viability dye (Thermo Fisher, Cat. No. L34963) was added at a concentration of 0.25 *µ*l/1,000,000 cells and incubated on ice for 30 minutes. Dye was removed by pelleting the cells by centrifuging for 5 minutes at 550 *g* (at 4°C). Cells were resuspended in PBS and pelleted again to remove as much of the LIVE/DEAD dye as possible. Finally, the cells were resuspended in 1 ml Hibernate A.

2.11 Antibody incubation for FACS.

Cell suspensions were then treated with appropriate antibodies for sorting and isolation. CD11b-APC (BioLegend, Cat. No. 101211) was added at a concentration of 1:400 and KiR7.1-ATTO488 (Alomone Labs, Cat. No. APC-125-AG) was added at 1:50. Antibodies were incubated for 15 minutes in the dark, on ice. Following incubation the cells were pelleted by centrifuging for 5 minutes at 550 *g* (at 4°C). The supernatant was discarded and the pellet was resuspended in Hibernate A in preparation for FACS.

2.12 FACS protocol.

All sorts were performed on the FACSAria II Cell Sorter (BD Biosciences). A 100 μm nozzle was used for cell sorting. The "Cell" gate was placed in the position that it is in, through trial-and-error work done by Dr. Vatanashevanopakorn (Vatanashevanopakorn *et al.,* 2017). Forward angle light scatter (FSC) and orthogonal light scatter (SSC) by cells were measured using a 488 nm laser and 488/10 nm band pass filters. This was used to remove cell doublets by using width measurements from the light scattering pulses. Viable cells were differentiated from non-viable cells by exclusion of LIVE/DEAD violet intensity, since only dead cells should have perforated membranes and thus be able to take up the dye. LIVE/DEAD fluorescence is detected by excitation at 405 nm and measuring emission using a 450/50 nm. Expression of tdTomato by viable cells is detected by fluorescence using excitation at 561 nm and measuring emission using a 582/15 nm band pass filter. Expression of APC by viable cells is detected by fluorescence using excitation at 633 nm and measuring emission using a 675/20 nm band pass filter. Expression of ATTO488 by viable cells is detected by fluorescence using excitation at 488 nm and measuring emission using a 515/30 nm band pass filter. Purity tests were performed for each sample by re-running isolated cells through the FAC sorter, along with proper "fluorescence-minus-one" (FMO) controls. Purity was 100% but recovery of the cells was extremely low—between 1-10%. This is probably due to very sensitive nature of the cells in question.

Gating protocols and logic are described in detail in the appropriate sections and figure legends.

2.13 Cell processing for Smart-seq2.

Lysis Buffer

For single cell RNA sequencing, using the Smart-seq2 method, single cells were collected (by FACS) into single wells of a 96-well plate (ThermoScientific, Cat. No. AB0800). The wells contained lysis buffer [0.4% Triton + RNase inhibitor (2 μl); dNTP's (10 mM; 1 μl) Oligo dT_{30VN} (10

uM; 1 μl)]. Upon completion of the sort, the plates were sealed quickly spun down (700 *g*) and then placed in a container containing dry ice. This was then transferred for short-term storage in a -80°C freezer.

Reverse Transcription

The first step in the process is reverse transcription. This is done by using the following reagents at the stated volumes and concentration, per well:

Reagent	**Volume (μl)**
Superscript II first strand buffer (5x)	2
DTT (100 mM)	0.5
Betaine (5 M)	2
MgCl2 (1 M)	0.1
RNase Inhibitor (40 U/μl)	0.25
TSO (100 μM)	0.1
water	0.8
Superscript II RT (200 U/μl)	0.25
total	*6*

6 μl of this solution was added to each well. The plates were then vortexed and spun down quickly. The reaction was performed using a PCR machine using the program below:

42 °C	**50 °C**	**42 °C**	**70 °C**	**4 °C**
90 min	2 min	2 min	15 min	-
	x10			

Preamplification

Following this, preamplification of the cDNA is performed using the following reagents, per well:

KAPA HiFi HS Ready Mix (2x) (KAPA Biosystems, Cat. No. KK2601)	12.5
ISPCR primers (10 μM)	0.125
water	2.375
total	15

The following protocol is used:

Step	Temperature (°C)	Time
1	98	3 min
2	98	20 sec
3	67	15 sec
4	72	6 min
5	Repeat Step 2-4, 24 times	
6	72	5 min
7	4	Hold

cDNA Purification

After preamplification, the cDNA product is purified using magnetic beads (Agincourt Ampure XP beads [Beckman Coulter, Cat. No. A 63881]). The beads were first equilibrated to room temperature. 8 μl of water and 25 μl of beads were then added to each well. This was allowed to incubate to ensure that cDNA would bind to the beads when placed on a magnetic stand. The cDNA bound beads were then washed twice with 80% ethanol and the pellet was then allowed to dry. The pellet was passively resuspended with elution buffer (Qiagen, Cat. No. 19086), by removing the plates from the magnetic stand. This was then gently mixed by pipetting and then placed back on the magnetic stand to separate out the magnetic beads from the eluted cDNA. The eluted cDNA was then pipetted out and placed into fresh 96-well plates.

Bioanalyser Analysis

Following cDNA purification, the quality of the cDNA needs to be checked. This was done by performing Bioanalyzer analysis (all materials from Agilent Technologies). In short, this method is a miniaturised DNA electrophoresis gel. The High Sensitivity DNA Dye Concentrate and the High Sensitivity DNA Gel Matrix were equilibrated to room temperature for 30 minutes. 15 μl of the High Sensitivity DNA Dye Concentrate was pipetted into the vial containing the High Sensitivity DNA Gel Matrix and mixed by vortexing. The mixed solution was then filtered using a spin filter and microcentrifuging for 10 minutes at room temperature at 2240 *g*. This mixed and filtered solution is referred to as the Gel-Dye Mix. The Bioanalyzer analysis is performed using proprietary microfluidic chip called a High Sensitivity DNA Chip. 9 μl of the Gel-Dye Mix was pipetted into each well of the chip and loaded into the well using a proprietary pressure-injection platform called a Priming Station. The ladder comes as part of the kit and was pipetted into the chip. 1 μl of cDNA was then pipetted into the wells. The chip contains 11 wells for samples, thus 11 cDNA samples were chosen to sample the whole plate. Following sample loading, the chip was gently vortexed and placed into the Bioanalyzer machine to run the chip. Once the chip run is completed the program generates a plot that relates fluorescence intensity to the sizes of the cDNA fragments. This plot can be used to determine the general distribution of fragment sizes and thus the quality of the cDNA.

Tagmentation and Indexing

Library preparation (tagmentation and indexing) is performed using the Nextera XT kit (Illumina, Cat. No. FC-131-1096). A new 96-well plate is used to perform the tagmentation.
2.5 μl of Tagmentation Buffer is added to each empty well, followed by 1.25 μl of cDNA and 1.25 μl of Amplicon Tagment Mix. The plates were then sealed, vortexed and spun down prior to performing the tagmentation on a PCR machine using the following program:

Cycle	Temperature (°C)	Time
1	55	5 min
2	4	Hold

The tagmentation needs to be neutralised, otherwise the Tn5 transposase will continue to generate fragments. This done by adding 1.25 μl of Neutralisation Buffer to each well and incubating for 5 minutes.

Following tagmentation, indexing of each sample can be performed. This is done by adding 8 X 12 primers (such that there are 96 unique combinations of primers). Into each well was added:

Reagent	Volume (μl)
Nextera PCR master mix	3.75
Index 1 primers (N7xx)	1.25
Index 2 primers (N5xx)	1.25

The primers were attached to the tagmented fragments using the following program:

Cycle	Temperature (°C)	Time
1	72	3 min
2	95	30 sec
3	95	10 sec
4	55	30 sec
5	72	30 sec
Repeat 3-5 12X	--	--
6	72	5 min
7	10	Hold

Following library preparation, purification of the indexed cDNA fragments by magnetic beads as outlined above. Bioanalyzer analysis was also performed to check the quality of the cDNA at this point as well.

Sample Pooling

Since all the samples from a 96-well plate are loaded onto a single flow cell, the concentration of DNA in each well needs to determined so that the final DNA concentration of the pooled samples will be 4 nM. This concentration measurement is done using the Quant-It PicoGreen system (Thermo-Fisher, Cat. No. P7589). The system uses a standard curve of known DNA concentrations and the fluorescence those concentrations of DNA generate, to estimate the concentration of the DNA in the analysed wells. A standard curve is first made by serially diluting the 2 μg/ml stock of DNA by 50% in 1X TE buffer, each well having a final volume of 30 μl. The standard curve has the following concentrations:

Standard	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
DNA concentration (μg/ml)	2	1	0.5	0.25	0.125	0.0625	0.03125	0

The PicoGreen reagent needs to be diluted to 1:200 (in TE) before it can be used. For each sample well, 29 μl of TE was added with 1 μl the cDNA. Then 30 μl of the PicoGreen reagent (at 1:200) was added to each well, including the DNA standard wells. These plates were then read on a fluorescent plate reader (ClarioStar) using an excitation of 483/14 nm and an emission filter at 530-30 nm with automated gain and focus adjustment. The readings from the standard curve are used to generate an equation of the slope of the line of best fit. This equation can then be used to estimate the concentration of the wells.

Once the concentrations are determined, an Excel spreadsheet called “Sample Pooling” is used to calculate the volume of the cDNA sample that needs to be taken from each well to arrive at the final pooled cDNA solution. In this set of experiments the final concentration was 4 nM at a final volume of 400 μl.

Preparation of the flow cell.

The pooled samples need to be denatured and further diluted before they are loaded onto the NextSeq machine using the NextSeq® 500/550 High Output Kit v2 (75 cycles; Illumina Cat. No. TG-160-2005). This was done by adding 5 μl of 0.2N NaOH and incubating at room temperature for 5 minutes. The reaction was stopped by adding Tris-HCl (200 mM). The library is diluted down to 20 pM by adding 985 μl of proprietary HT1 buffer. It is diluted once again to the final concentration of 1.8 pM by adding 117 μl of the pre-diluted library to 1183 μl of the HT1 buffer. This total 1.3 ml solution is then pipetted into the chip. The chip and the requisite buffers (that come as part of the kit) were loaded into the NextSeq machine along with a new flow cell. The NextSeq then performs the sequencing and generates a FASTQ file with the sequencing data that can then be used to perform the bioinformatics analysis.

2.14 Bioinformatics Analysis.

All bioinformatics analysis was performed with help from the Computational Biology Research Group at the Weatherall Institute for Molecular Medicine, using different packages in R (R Core Team, 2013, URL: http://www.R-project.org/). The first analysis that was performed was the requisite, and rough, quality control analysis of the read data itself using the package FastQC (Babraham Institute, URL: http://www.bioinformatics.babraham.ac.uk/projects/fastqc). This quick analysis is performed to ensure that the base callings are of high enough quality to perform bioinformatics analysis on, which is computationally intensive. The FASTQ file generated by the NextSeq encodes quality control data about every single nucleotide read, corresponding to the amount of signal that was incorporated when the base calling was performed by the TIRF microscope sensor and NextSeq software. The FastQC package allows for quickly gauging this information before proceeding to more rigorous quality control analysis following the alignment of the reads. The base calling for both experiments passed this analysis.

The second analysis that was performed was the alignment of reads. The reads were aligned to a widely-used mouse reference genome (NCBI GRCm38 - mm10; C57BL/6J strain). This was done through the package STAR (**S**pliced **T**ranscripts **A**lignment to a **R**eference; Dobin *et al.,* 2013). cDNA is generated from RNA, thus the introns that are present in genomic DNA are absent. STAR aligns cDNA to genomic DNA by iteratively aligning the largest part of a cDNA fragment (called *seed1*) to the reference genome. The remaining part of the cDNA (called *seed2*) is then aligned to a reference genome. This is done iteratively (both in the forward and reverse directions) so that each cDNA fragment is aligned genomic DNA, in a process called “stitching”. However, not all parts of the cDNA will be properly mapped (due to genomic variability and errors during first strand synthesis, preamplification, tagmentation, indexing, and sequencing). In these situations, the number of unaligned, multimapping, or even featureless reads are recorded and can later be used as an assessment of the quality of the sequencing and the samples; this is called “scoring” (Figure 3.7). Cells with less than 60% uniquely mapped reads were excluded from both analyses. To ensure that each cell was deep-sequenced, cells with less than 500,000 uniquely mapped reads were also excluded.

While the STAR package is able to align the reads to the genome, it does not provide any data regarding which genes those reads are aligned to. STAR does, however, provide a .BAM file with the chromosomal location of every mapped read. Using the featureCounts tool in the subRead package one can simultaneously determine which genes the reads are mapped to and, also, the number of times the gene was sequenced, or counted (Liao *et al.,* 2014). However, the counts for genes will vary depending on the reads per sample and the length of the gene. To account for this, the counts need to be normalised. For both analyses, this was done using a median of ratios method, also known as normalisation by size-factor. This is different from the crude counts-per-million (CPM) or reads-per-kilobase of transcript-per-million (RPKM) methods. This size-factor method works by first finding the geometric mean of the expression of the gene

across all samples. All the samples are then divided by this geometric mean to give a ratio of the sample to the geometric mean. Then the median of the ratios is determined, across all the samples. This median is called the size-factor. Now, all the samples are divided by this size-factor to arrive at the normalised expression value for the gene in all the samples. This method is superior to the other methods because it can account for large variations in gene expression and also sample dropouts. This normalisation was done using the Scater package (McCarthy *et al.,* 2017).

Following the normalisation, we clustered the cells based on the normalised expression of genes using the PAGODA (Pathway and Gene Set Overdispersion Analysis) methodology, as part of the SCDE package (Kharchenko *et al.,* 2014; Fan *et al.,* 2016). This method clusters cells based on statistically evaluating the coordinated expression variability of previously-annotated gene pathways (as in Gene Ontology analysis) on top of the regular method of evaluating the coordinated expression of *de novo* (non-annotated) gene sets. It also works by generating error models for each sample, which like in Scater, take into account sequencing depth and dropouts. The differential expression of genes between the clusters generated by PAGODA was determined by one-to-one comparison between the groups. The adjusted *p*-value that is listed in the results is the output of the SCDE package and is called “adjusted” because it accounts for multiple hypothesis testing.

Pseudotime analysis was also performed on the clusters to estimate the temporal relationship between cells and clusters. This is analysis was done using the TSCAN package (Ji and Ji, 2016). This is a PCA-based method, in that it uses the clusters in the PCA plot to build a minimum-spanning tree to get an estimate of the pseudotemporal ordering. The minimum-spanning tree approach seeks to find the path with the shortest possible distance between the centres of clusters.

Gene Ontology (GO) analysis was performed using the GOsummaries package (Kolde and Vilo, 2015). Gene Ontology is a standardised repository of the function of genes. Since the SCDE analysis is able to generate differentially expressed genes, GO analysis can be used to find differences in pathways that are encoded by the differentially expressed genes. The GOsummaries package simply allows for the visualisation of the enrichment of pathways. The output is described in the figure legend for Figure 3.30.

2.15 High-throughput gene expression analysis.

This method is outlined in detail in our published manuscript (Nath *et al.,* 2017). Prior to high-throughput gene expression analysis, cells were processed and FAC sorted as described above. Where there was a difference was in the lysis buffer and the number of cells being sorted. For the validation of the elimination of microglial signature, 50 cells were sorted into the lysis buffer below. For the validation of the capture of Group1 cells using the anti-KiR7.1 antibody, 10 cells were sorted into the lysis buffer below:

Reagent	**Volume (µl)**
CellsDirect 2x Reaction Mix	5
Ambion's SUPERase 12-In	0.1
TE Buffer	0.9
(Total)	*(6)*

The above reagents come as part of a kit: CellsDirect™ One-Step qRT-PCR Kit (Thermo Fisher, Cat. No. 11753100). The lysis buffer was placed into 0.2 ml thin-walled PCR tubes and the appropriate number of cells, from the appropriate gates were sorted into the tubes. The tubes were quickly spun down and frozen in dry ice and placed in the -80°C freezer for storage.

The reverse transcription and pre-amplification were done in the following way. The gene expression analysis performed here is a high-throughput quantitative RT-PCR. Thus, specific

primers for genes of interest need to be used. To this end, the following “TaqMan Assays” were obtained from Thermo Fisher Scientific:

Assay Name	Assay ID
Enpp2	Mm00516572_m1
Ttr	Mm00443267_m1
Kcnj13	Mm01287179_m1
Folr1	Mm00433355_m1
Prlr	Mm04336676_m1
Igfbp2	Mm00492632_m1
CD59	Mm01276239_m1
Darpp-32	Mm00454892_m1
GLAST	Mm00600697_m1
CD133	Mm01211402_m1
GFAP	Mm01253033_m1
Gja1	Mm01179639_s1
Itih3	Mm00434548_m1
Sox9	Mm00448840_m1
Sox11	Mm01281943_s1
Dcx	Mm00438400_m1
Gad1	Mm04207432_g1
Snap25	Mm01276449_m1
Tubb3	Mm00727586_s1
Penk	Mm01212875_m1
Drd1	Mm02620146_s1
Drd2	Mm00438545_m1
Tmem119	Mm00525305_m1
Sfpi1	Mm00488142_m1
Csf1r	Mm01266652_m1
Canx	Mm00500330_m1
Gusb	Mm01197698_m1
Rn45s	Mm04277571_s1

These FAM-MGB dye conjugated probes come at a concentration of 20X. For the experiment they need to be diluted down to 0.2X and combined for the appropriate experiments, into a mixture

called the “0.2X Assay Mix” in TE buffer. Once this step is completed, the following solution is made and added to each bulk-cell sample:

Reagent	**Volume (μl)**
0.2x TaqMan Assay Mix	2.5
Ambion’s SUPERase 12-In	0.1
RT/Taq (SS III Platinum) Mix	1.2
(Total)	*(3.8)*

After the reagents were added and the tubes vortexed and spun down, they were placed in a PCR machine to perform the reaction. The following program was used:

Step	**Temperature (°C)**	**Time**
Reverse Transcription	50	15 min
Activation of RTase/Activation of Taq	95	2 min
Preamplification:	95	15 sec
	60	4 min
	Repeat for 18 cycles	--
Elongation	72	10 min

Once the process is completed the tubes are stored at 20°C in preparation for Fluidigm gene expression analysis on the Fluidigm microfluidic chip. When ready to perform the analysis on the chip, each sample was thawed and 40 μl of TE buffer was added to each tube. Preparation of the cDNA samples and TaqMan assays (the latter were aliquoted in the clean room when the lysis buffer was prepared) were done exactly as stated in the standard Fluidigm protocol. The protocols can be found on the Fluidigm website: https://www.fluidigm.com/documents. The reagents in the protocol are: TaqMan Universal PCR Master Mix (2X) (Life Technologies, Cat. No. 4304437) and 20X GE Sample Loading Reagent (Fluidigm, Cat. No. 100-7610). The protocols differ slightly depending on the type of chip being used. In these experiments the 48.48 chip was used (Fluidigm). The chip was filled and loaded into the BioMark device in preparation for the actual qRT-PCR.

The following analytical method was adapted from a manuscript published by the Nerlov group, based at the MRC Weatherall Institute of Molecular Medicine in Oxford (Drissen *et al.,* 2016). Following the run, the data was visualised using the Fluidigm software "Real-Time PCR Analysis". From here the dataset was exported (*ΔCt* values) as a .CSV file. This .CSV file was then converted to an .XLSX file, such that it could be easily analysed in Microsoft Excel. The data can be handled in any number of ways depending on what the experiment demands. The most common way, and the way it was analysed for the data in this thesis, is to represent and evaluate the data by a heat map. However, the processed *ΔCt* values can be used to perform statistical analysis between bulk cell groups. To perform the subsequent analysis the *ΔCt* values of each sample needed to be clearly separated. We used the "filter" function to filter the dataset for one of the genes—this now showed the *ΔCt* values of all the samples for that particular gene. This was copied and pasted into a new sheet and repeated for all the genes. Thus, we had each sample's expression level (*ΔCt* values) of each gene organized in adjacent columns. We replaced each "999" value (which indicates that no expression was detected) with an "x" for analysis purposes. This had to be changed to an "x" or the subsequent analysis would have been rendered meaningless. The first quality control step was to remove all samples that did not express the reference gene(s). The expression level of each gene had to be normalised. This was done by normalising the expression of each gene (in each sample) to that sample's expression of a reference gene (either Canx or Rn45s). As such, each gene's expression level was subtracted from the expression level of the reference gene:

$$= \Delta Ct \text{ (reference gene)} - \Delta Ct \text{ (gene being tested)}$$

Following this, we replaced all the error "#VALUE!" cells with "x". This results from samples where the expression of a gene is not detected. For each gene, the average expression level across all samples was calculated. Then we centered the expression of each gene around 0. This allowed us to determine gene expression relative to other samples of interest. This was done by

subtracting the mean expression of each gene (across all samples) from the expression level of the gene in each sample:

$$= \Delta Ct\,(\text{gene}) - \text{mean}(\Delta Ct\,(\text{gene}))$$

Then, as before, we replaced the error "VALUE!" cells with an empty cell. The empty cells represent "undetected expression." This file was imported into a program that allows for the graphical visualization of the gene expression data. The online open-source software Morpheus (Broad Institute) was used to generate heat maps.

2.16 Tissue processing for wholemount lateral ventricle staining.

Wholemount ventricular staining allows for the *en face* analysis of the lateral wall of the lateral ventricle. The protocol that was employed to obtain the images in this thesis was nearly identical to the widely-used protocol described in Mirzadeh *et al.* (2010). Briefly, the animal is terminally anaesthetised and decapitated (without cardiac perfusion) and the brain is dissected out. The brain is then cut along the interhemispheric fissure and a coronal cut is made so that the caudal hippocampus can be seen. The hippocampus and cortex are then pulled and cut away to reveal the lateral ventricle. The tissue was then trimmed to obtain the cleanest possible cut of the lateral ventricle (and the still-attached striatum). It was then fixed overnight at 4°C in 4% PFA. The wholemounts are then washed extensively using PBS in preparation for immunostaining. The immunostaining method was identical to the fluorescent immunochemistry method described above, except that primary and secondary incubations were performed for 48 hours. Following immunostaining, the wholemounts need to be sub-dissected to ensure that they can be mounted onto glass slides to enable confocal and epifluorescence imaging. Thus, the remaining striatum and overlying cortex were dissected away using a 22.5° microsurgical stab knife under a dissecting microscope. This resulted in a relatively flat section of the lateral wall of the lateral ventricle of 200-300 μm thickness. The section was mounted as described previously, using

Vectashield. The slides were left to settle and cure for at least 24 hours at 4°C before any imaging was performed.

2.17 CLARITY tissue processing and imaging.

The method is slightly modified from the standard protocol described in Tomer *et al.* (2014) and was done in collaboration with Dr. Christoffer Lagerholm at the Wolfson Imaging Centre at the Weatherall Institute for Molecular Medicine. The brains underwent cardiac perfusion in the manner described before using both 0.1M PB and 4% PFA. They were then post-fixed for 24 hours at 4°C with 4% PFA. The brains were then rinsed with PBS several times. They were then left to be washed further in fresh PBS for 24 hours at 4°C. Following the washing the brains were incubated overnight in the hydrogel monomer solution:

	Add for 10ml	**Final**
Acrylamide (40%)	1 ml	4%
VA-044 (10%)	0.25 ml	0.25%
10X PBS	1 ml	1x
H_2O	7.75 ml	-
(Total)	*(10 ml)*	

The brains were then polymerised for 3 hour in the absence of oxygen at 37°C. This was followed by a 24-hour incubation at room temperature for in X-CLARITY ETC (electrophoretic tissue clearing) buffer (Logos Biosystems, Cat. No. C13001). The actual clearing was performed using the X-CLARITY instrument for 15 hours (Logos Biosystems, Cat. No. C30001). The brains were then rinsed carefully with PBS and left in PBS overnight to remove all ETC buffer. The clarified brains were then ready to be imaged using a confocal microscope (Zeiss 780 LSM upright confocal microscope with MP laser) using a special CLARITY objective (LD Plan-APOCHROMAT 20X/1.0) at the Weatherall Institute for Molecular Medicine. The brains were superglued on a platform and completely submerged in a refractive index matching solution called the X-CLARITY Mounting Solution (Logos Biosystems, Cat. No. C13101). Following imaging of the tissue volume,

the images were stitched and analysed using the Arivis InViewR software. The video was produced using tools in the InViewR program. The stills presented in the thesis are screenshots of the film.

2.18 BrdU analysis.

5-bromo-2'-deoxyuridine (BrdU) was obtained from Sigma (Cat. No. B5002) and made up in 0.9% NaCl. It was made up to 20 mg/kg and injected at a final volume of 200 mg/kg (this equaled 10 μl/g). The mice were intraperitoneally injected as per Home Office guidelines.

To visualise the BrdU in cryosectioned slices, sections were treated in 2N HCl for 45 minutes at room temperature. The sections were then thoroughly rinsed with borate buffer (pH = 8.5) and left to incubate in said buffer for 15 minutes at room temperature. All other steps were identical to the immunofluorescence staining protocol described previously.

2.19 Stereotactic injection of 6-OHDA

All experimental procedures on animals were conducted in accordance with the Animals (Scientific Procedures) Act, 1986 (United Kingdom). Male and female IPNK mice aged from 2 to 4 months were used for all experiments. Animals were housed with their littermates with free access to food and water in a dedicated housing room with a light/dark cycle (7 a.m. to 7 p.m.).

The injections were performed in collaboration with Prof. Peter Magill's group with help from Dr. Natalie Doig. Mice were anaesthetized with isoflurane (0.5–3.0% oxygen gas mixture 2 l/min) and given buprenorphine (0.1 mg/kg of body weight) for peri-operative analgesia during stereotaxic injections.

Initially, the mice were injected at a single site: AP: -0.1 mm from bregma; ML: 2.0 mm from bregma; DV: -2.25 mm from bregma. The injections for the main experiment were targeted unilaterally to the striatum at 2 sites (rostral site: +0.6 mm anterior, +2.3 mm lateral from Bregma

at a depth of -2.5 mm from the brain surface and caudal site: -0.1 mm posterior and +2.4 mm lateral to Bregma at a depth of -2.4 mm from the brain surface; Paxinos and Franklin, 2013). Either 6-hydroxydopamine (6-OHDA [6 mg/ml] dissolved in saline and 0.02% ascorbic acid) or the vehicle alone (0.02% ascorbic acid in saline) were injected (750 nl per site) at a maximum rate of 100 nl/min using a glass micropipette lowered to the target site; the micropipette was left in place 5 minutes before and after the injection. Recovery of mice was monitored daily.

2.20 Statistical analysis.

All statistical analysis was performed on GraphPad Prism 7. The test used to compare treatments in Chapter 5 was an unpaired t-test. Differences were considered significant when $p < 0.05$. The statistical information listed in Chapter 3, specifically the $p_{adjusted}$ values, were obtained from the analysis performed on the SCDE package. Differences were considered significant when $p_{adjusted} < 0.05$.

Chapter 3 – Results I: Using the Penk-tdTomato (IPNK) mouse line to isolate and sequence adult neural stem cells.

3.1 Penk-tdTomato expressing cells at the ventricular/sub-ventricular zone of the lateral ventricle.

The BAC-Penk-Cre line has been extensively characterised and published first in Besusso *et al.* (2013) and then in Vatanashevanopakorn *et al.* (2017). In the latter, it has been crossed to a reporter line described in Madisen *et al.* (2010), expressing the tdTomato fluorescent protein upon Cre-recombination. This cross resulted in the Penk-tdTomato strain (IPNK) that is used throughout this thesis. Briefly, the IPNK line leads to the intense and irreversible expression of the tdTomato fluorescent protein in Enkephalin-expressing cells. It was designed to label the medium spiny neurons (MSNs) of the indirect pathway of the striatum, which are known to express Enkephalin (as opposed to the direct pathway neurons). However, when we closely analysed cryosectioned IPNK brains under an epifluorescence microscope we found intense Penk-tdTomato expression in the ventricular/sub-ventricular zone (V-SVZ) in addition to its canonical expression in the striatal parenchyma (Figure 3.1). We noted that not all the cells of the V-SVZ are identified using the Penk-tdTomato reporter, but rather a subset of them. The tdTomato identified cells of variable morphology, which indicated that it might be identifying several different cell types within the stem cell niche. Interestingly, we found that the morphological characteristics of some of these Penk-tdTomato^{+} cells mirrored those seen in electron micrographs of quiescent NSCs, or B-cells, of the V-SVZ niche—one key trait being the characteristic bipolar projections that run parallel to the lateral wall of the lateral ventricle (Figure 3.2; Doetsch *et al.,* 1997; Obernier *et al.,* 2018). In the figure, we can also see that the Penk-tdTomato^{+} cell is also intensely stained

for GFAP—a marker of quiescent NSCs of the V-SVZ (discussed at length in ***Chapter 4***; Doetsch *et al.,* 1997; Doetsch *et al.,* 1999).

We wanted to eliminate the possibility that the identified cells at the V-SVZ could be expressing tdTomato due to leaky expression of the tdTomato reporter and not the specific activation of CRE in Penk-expressing cells. Thus, we sought to demonstrate similar expression patterns in the V-SVZ in other reporter lines. To this end, we were able to demonstrate similar cell types in the Penk-Z/EG reporter line and the Penk-Rosa26-EYFP line used in Besusso *et al.,* 2013 (Figure 3.3a, and Figure 3.3b, respectively) using immunostaining with 3,3'-diaminobenzidine (DAB). Taken together, these results indicated that the IPNK line identifies a sub-group of cells within the stem cell niche that is the V-SVZ. Thus, we began a series of experiments to characterise and analyse these cells at the single-cell level.

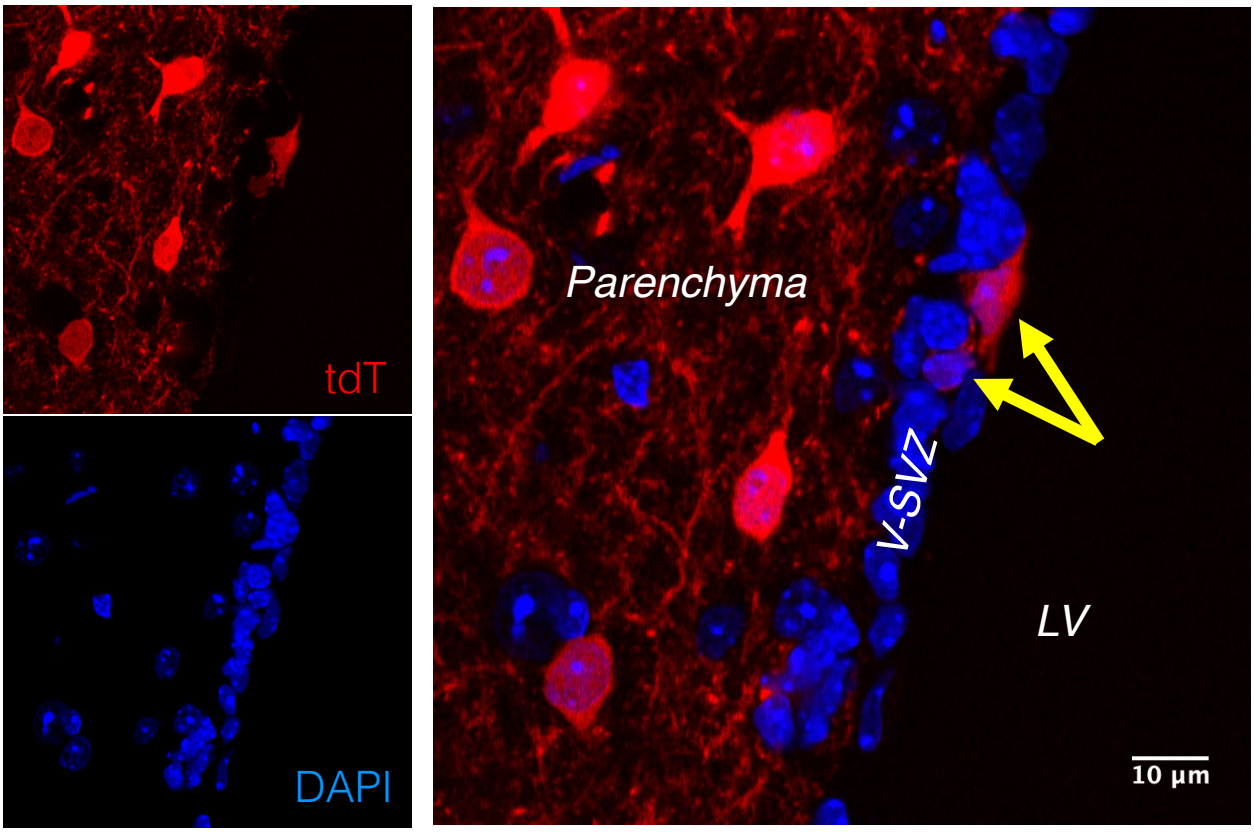


Figure 3.1 Penk-tdTomato is expressed by some cells residing in the V-SVZ stem cell niche. An IPNK mouse brain (female; 2 months old) was cryosectioned and stained for DAPI and visualised using an upright confocal microscope. The V-SVZ is demarcated by the thin, dense, and continuous line of DAPI-stained nuclei. This separates the ventricular space, to the right, and the striatal parenchyma, to the left where Penk-tdTomato$^+$ cells are canonically seen. These cells display the characteristic morphology of MSNs. The two ***yellow arrows*** indicate two Penk-tdTomato$^+$ cells residing within V-SVZ germinal niche.

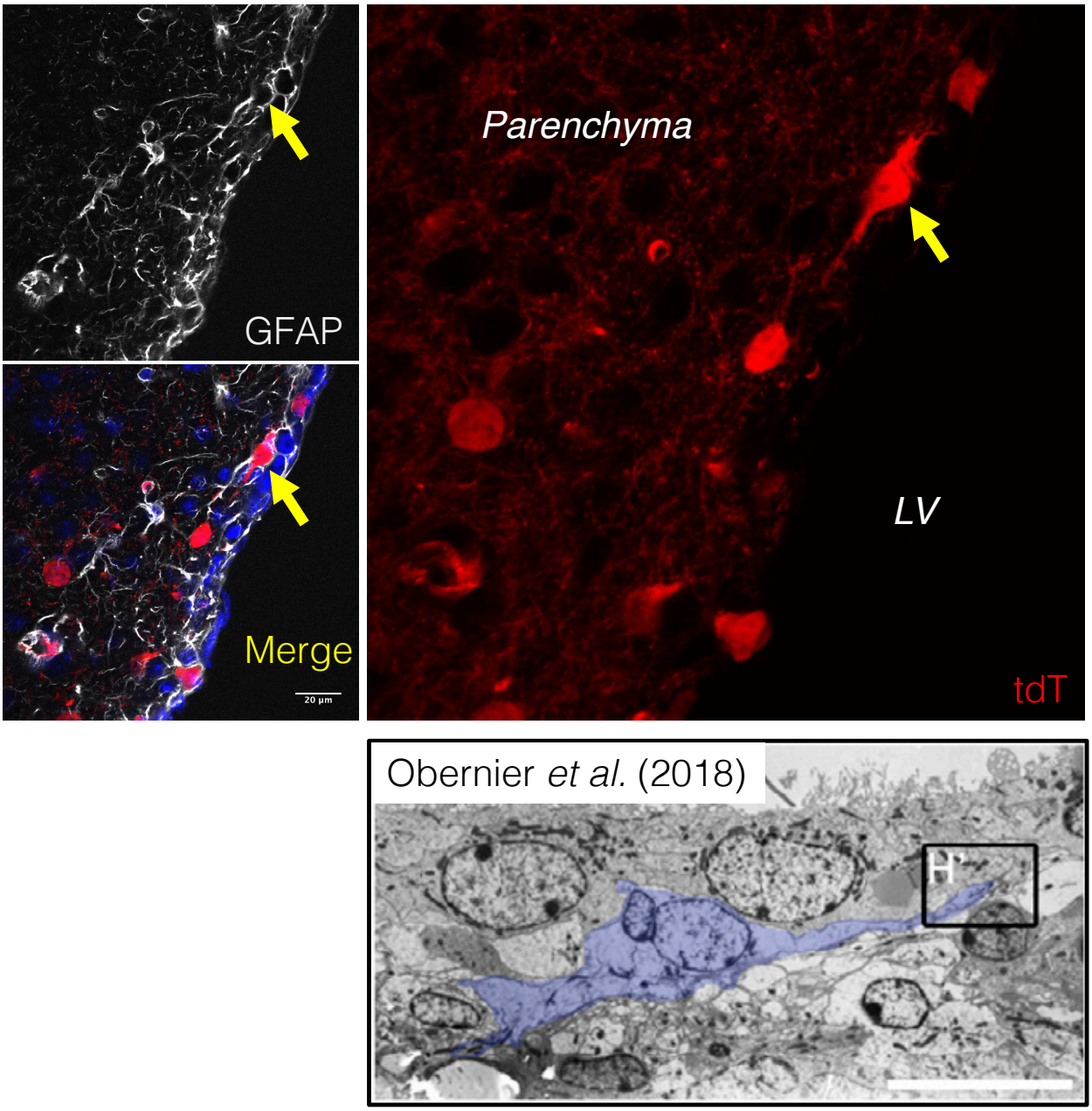


Figure 3.2 Penk-tdTomato⁺ cells of the V-SVZ possess the morphology of B-cells. An IPNK mouse brain (female; 2 months old) was cryosectioned and stained for DAPI and GFAP and visualised using an upright confocal microscope. The ***yellow arrow*** indicates a Penk-tdTomato⁺ cells residing within V-SVZ germinal niche with projections running parallel to the V-SVZ. This cell also stains positively for the quiescent NSC marker GFAP. At the bottom is an electron micrograph taken from Obernier *et al.* (2018; *with permission*) of a B-cell. Notice the almost identical morphology and placement of the Penk-tdTomato⁺ cell and the B-cell within the V-SVZ niche.

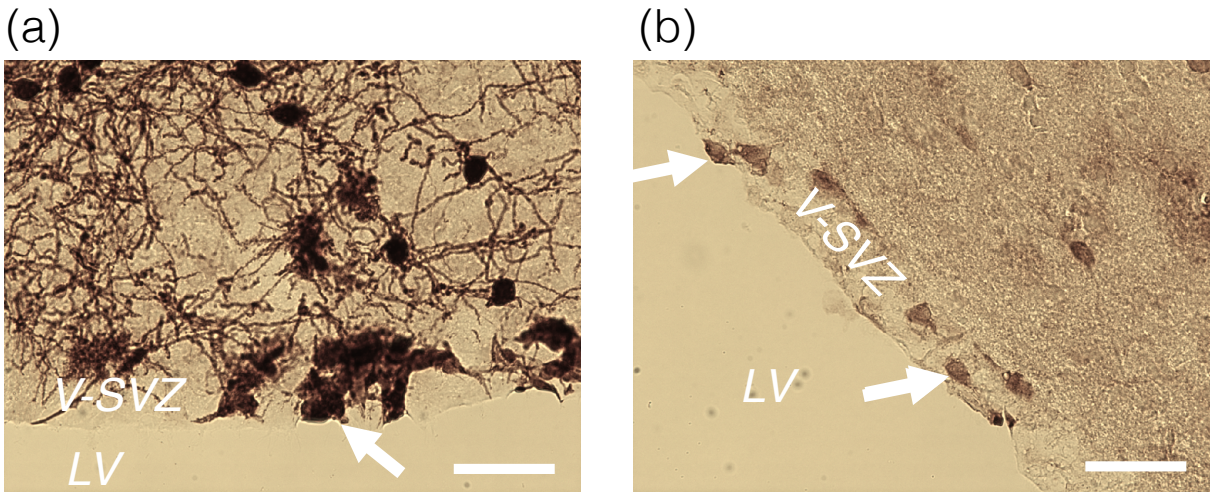


Figure 3.3 Two other fluorescent reporter lines crossed to the BAC-Penk-CRE strain replicate a similar expression pattern in the V-SVZ. (a) A Penk-Z/EG-GFP mouse brain section immunohistochemically stained with 3,3'-diaminobenzidine (DAB) for the GFP reporter. Clusters of GFP^+ cells from BAC-Penk-CRE activation can be seen at the V-SVZ. **(b)** A Penk-Rosa-YFP mouse brain section DAB stained for the YFP reporter. GFP staining activated by the BAC-Penk-CRE expressing cells can be found within the V-SVZ of this reporter line as well. Both images visualised using an upright epifluorescence microscope. Arrows point to reporter-expressing cells at the V-SVZ. Scale bars = 30 µm.

3.2 Optimising the isolation protocol for single cells from acutely dissected striatal tissue.

Single cell analysis of the Penk-tdTomato$^+$ of the V-SVZ first required an isolation protocol that was replicable and would allow us to keep sufficient numbers of acutely dissected neurons alive while also ensuring that they were properly dissociated. A previous D.Phil. student in the laboratory, Chinnavuth Vatanashevanopakorn, developed an isolation protocol for striatal MSNs for bulk cell analysis. I adapted and optimised that protocol so that we could capture live single cells reliably (Vatanashevanopakorn *et al.,* 2017). This required extensive work over several months because of how sensitive mature neurons are to changes in their environment. The mechanical stress caused by the isolation procedure further damages the neurons. Taken together these two factors can lead to widespread cell loss (Saunders *et al.,* 2019). As described in detail previously by Vatanashevanopakorn *et al.* (2017), the protocol is centered around the use of differential centrifugation of a discontinuous gradient to separate out neurons from debris generated during the dissection and enzymatic dissociation steps. Figure 3.4a shows a representative discontinuous gradient following centrifugation. As a preliminary indication of whether we had been able to enrich tdTomato$^+$ cells, we plated the neuronal fraction on a coverslip and imaged using an epifluorescence microscope. Figure 3.4b indicates that we were able to collect what resemble tdTomato$^+$ cells using this isolation protocol.

We initially sought to perform a comparison between the V-SVZ cells and the MSNs. Thus, to analyse V-SVZ cells as well as the MSNs, we dissected out the lateral wall of the lateral ventricle along with the whole striatum, under the assumption that we would be able to separate out the V-SVZ cells from MSNs during the bioinformatics analysis stage. Thus, I used this method to create cell suspensions of V-SVZ cells and MSNs.

3.3 Fluorescence-activated cell sorting-mediated isolation and capture of single Penk-tdTomato$^+$ cells.

The method has been previously described in detail in the ***Materials and Methods***. Briefly, the fluorescence-activated cell sorting (FACS) was performed to enable the isolation of single and live Penk-tdTomato^{+} cells. Extensive optimisation of this process was performed in the laboratory and published in (Vatanashevanopakorn *et al.,* 2017). A cell suspension of Penk-tdTomato cells (both V-SVZ cells and MSNs) was sorted using a combination of forward and side scatter, along with LIVE/DEAD staining (to select for live cells), while the detection of tdTomato fluorescence permitted the purification single Penk-tdTomato^{+} cells. To ensure that we were isolating our target cell population we performed purity tests for each sample by re-running isolated cells through the FAC sorter, along with proper "fluorescence-minus-one" (FMO) controls. Purity was 100% but recovery of the cells was extremely low—between 1-10%. This is probably due to very sensitive nature of the cells in question. A further and more detailed analysis of the purity of the sorted populations is the sequencing itself—we can determine from the sequenced cells whether or not we have captured the target cell population or whether there was contamination of any kind. Figure 3.5 shows the gating procedure for this isolation method.

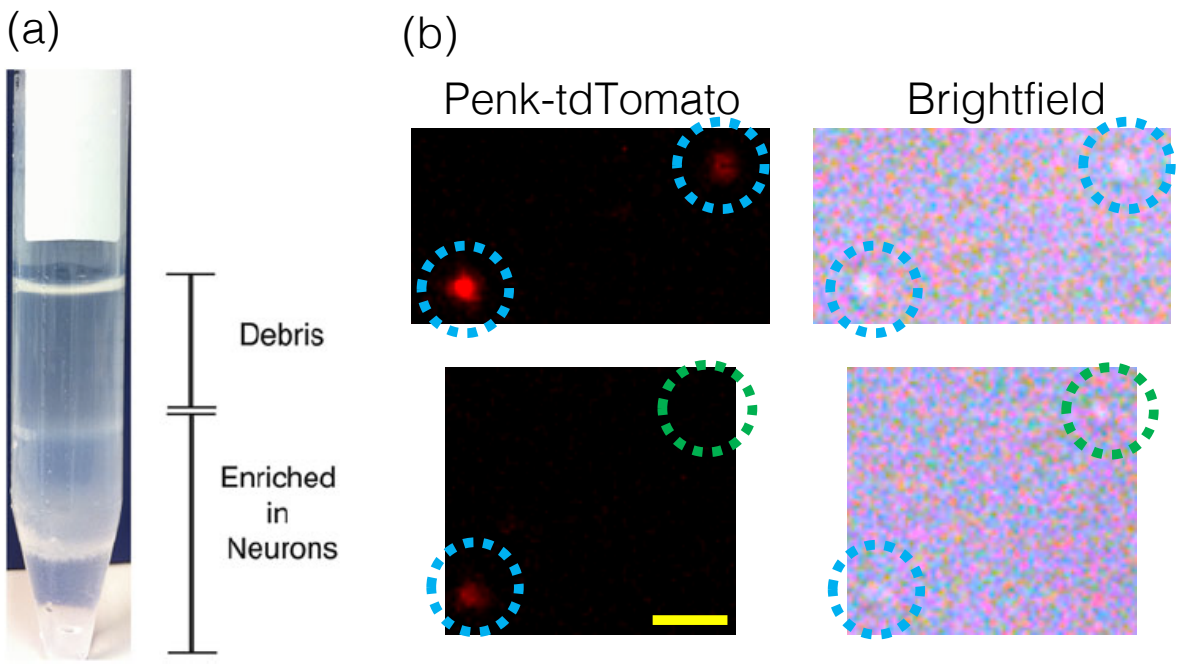


Figure 3.4. Optimised differential centrifugation method to separate live neurons from debris. (a) The image shows the result of differential centrifugation in a Percoll gradient. The debris is aspirated and discarded. *Adapted from Vatanashevanopakorn et al. (2017).* **(b)** Single neurons dissociated from the striata of a 5-month-old IPNK mouse. Tissue was enzymatically and mechanically dissociated and isolated by ultracentrifugation in a Percoll based density gradient column. Material from the purified neuronal layers was plated on a glass coverslip and imaged using an upright epifluorescence microscope. Penk-tdTomato^{+} cells can be identified by their endogenous tdTomato signal. The corresponding brightfield image reveals the general morphology of the cells. Structures that stain postively for tdTomato and have the general shape of a dissociated neuron are highlighted by ***blue*** squares. The ***green*** circle identifies a cell that is tdTomato negative. Scale bar = 5 µm.

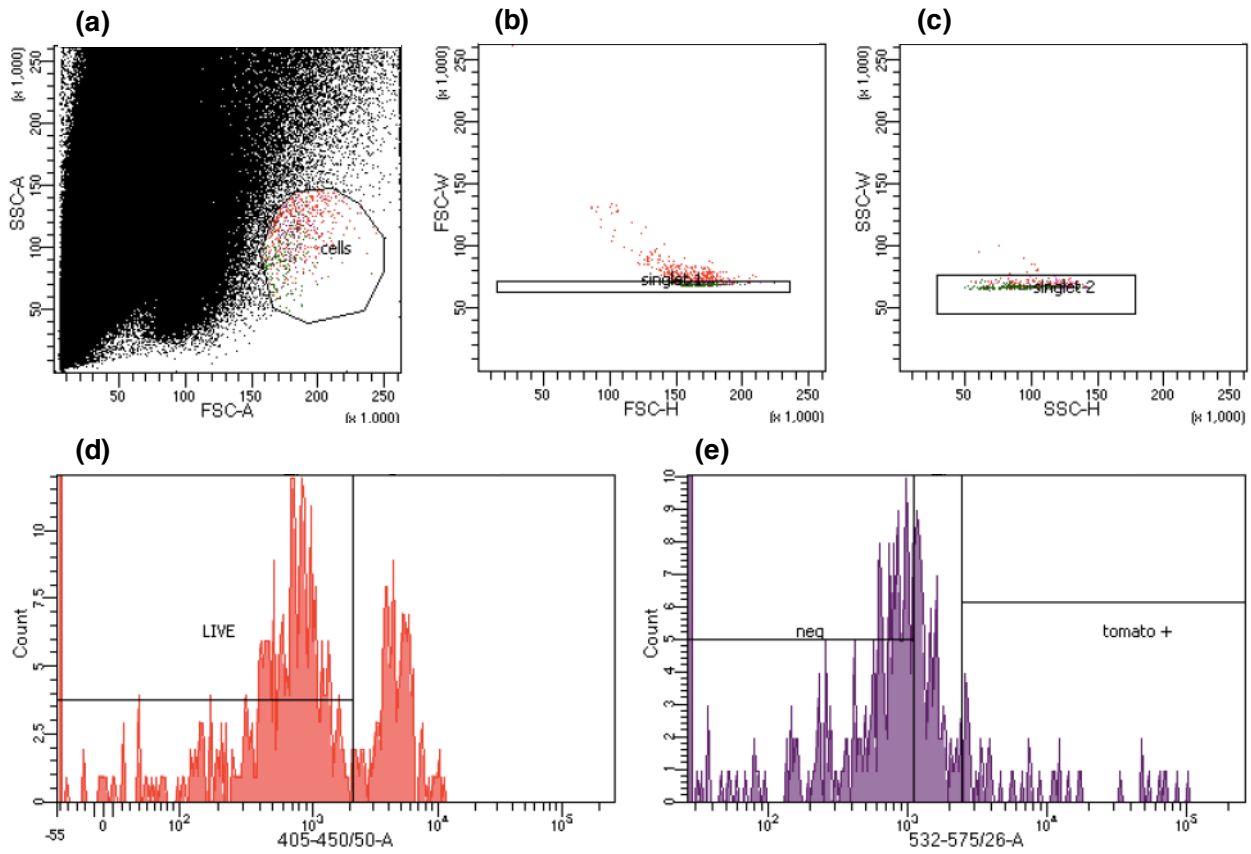


Figure 3.5. A representative gating protocol for fluorescence-activated cell sorting-mediated isolation and capture of single Penk-tdTomato⁺ cells . (a) Single neurons dissociated from the striata of a 7-month-old IPNK mouse and stained with LIVE/DEAD stain to identify dead neurons. Cells were sorted using a FACSAria Fusion Cell Sorter. All sort events are plotted according to forward and side scatter. The indicated "cell" gate was chosen after extensive optimisation by previous work in the lab, in conjunction with other independent, published work (Vatanashevanopakorn *et al.,* 2017). Black dots identify all events. Red dots identify the subset of cells that possess the morphological characteristics we identify as those of cells. **(b)** and **(c)** Forward and side scatter plotted to select strictly for singlets. **(d)** The singlet cells from (b) and (c) are then plotted on a frequency histogram to determine their intensity of LIVE/DEAD staining. Using extensive empirical methods, the lab has determined that live cells are those with a fluorescence intensity to the left of the line indicated. **(e)** The "LIVE" cells from (d) are plotted on a frequency histogram to determine the intensity of the tdTomato fluorescence signal. Cells to the right of the "tomato+" line are taken as Penk-tdTomato⁺ cells.

3.4 StriatumSeq1: Deep RNA sequencing of single Penk-tdTomato$^+$ cells.

As stated above, the goal of this experiment was to sequence single V-SVZ cells in parallel to the MSNs of the striatal parenchyma, so that we could perform a comparison between the cell types while also dissecting the molecular nature of the Penk-tdTomato$^+$ cells in the V-VSZ. To this end, we isolated the lateral ventricles and striata from four IPNK male mice, aged approximately four months. Each preparation was kept separate so that we could determine which cell originated from which mouse. We named this experiment StriatumSeq1.

The cell suspensions were treated as described above and the optimized FACS protocol was used to isolate and capture single Penk-tdTomato$^+$ cells. We followed the "Smart-seq2" deep RNA sequencing method developed at the Karolinska Institute (Picelli *et al.,* 2014). This method is a refinement of the original shotgun sequencing approach employed by Venter *et al.* (1995). It relies on sequencing small fragments of DNA and then aligning the sequenced fragments *post hoc*. The single cells were sorted into a lysis buffer along with appropriate reagents to perform reverse transcription and cDNA amplification. Approximately 96 cells, from each mouse brain, were collected, for a total of four 96 well plates or approximately 384 cells. We then performed cDNA fragment analysis to determine whether a cell was indeed alive, and if the RNA was intact and thus appropriate for cDNA amplification. Figure 3.6 a-c shows the representative cDNA profiles (obtained using the Agilent Bioanalyzer platform) of two single cells (a-b) from StriatumSeq1 along with an idealised profile from a live cell with intact RNA from the original Picelli *et al.* (2014) manuscript (c). This analysis was performed for each plate of single cells, and the analysis indicated that each plate generally contained cells with excellent cDNA fragment profiles. Thus, we proceeded with the next phase of the RNA sequencing.

We used the Tn5 transposase, in a step called tagmentation, to obtain cDNA fragments of manageable lengths and to simultaneously add primers to perform the sequencing. Fragment analysis was also performed following this step to ensure that tagmentation occurred properly.

Figure 3.6 d-f shows two representative profiles from two single cells from StriatumSeq1 (d-e) along with an idealised profile from Picelli *et al.* (2014; f). While the fragments from StriatumSeq1 run slightly larger, on average, from the idealised profile they still indicate that the tagmentation process was successful and thus that the fragments could be sequenced appropriately.

The sequencing was performed on the Illumina NextSeq platform using NextSeq 550 System High-Output Kit flow cells. Each flow cell generates 400 million reads and since we used one flow cell per 96-well plate, it meant that we would be sequencing at a depth of approximately 4 million reads per cell (later bioinformatics analysis revealed that the average sequences per sample was 3,663,069)—ensuring that, indeed, we were performing deep sequencing.

3.5 Initial bioinformatics analysis of the RNA sequencing data.

A total of 375 cells were initially sequenced from four male IPNK mice, approximately 16 weeks of age. We then carried out the bioinformatics analysis in collaboration with the Computational Biology Research Group at the MRC Weatherall Institute of Molecular Medicine, Oxford. Read alignment and quality control checks were performed to determine whether the sequencing data was of high enough quality to continue with further bioinformatics analysis. The NextSeq-generated reads were aligned using a software called Spliced Transcripts Alignment to a Reference (STAR; Dobin *et al.,* 2013) against the mouse genome assembly (NCBI GRCm38 - mm10). Aligning the reads gave us a snapshot of the genes being expressed but it did not provide any data regarding the amplitude of gene expression. To realise this, gene expression levels were quantified as read counts using the featureCounts function (Liao *et al.,* 2014) from the Subread package (Liao *et al.,* 2013) with default parameters.

3.6 Quality control analysis of the RNA sequencing data.

We then began the quality control analysis to remove low-quality cells (criteria are outlined below). This is crucial because the inclusion of low-quality cells in the dataset can significantly alter the conclusions drawn from bioinformatics analysis. The quality control thresholds that we employed were quite stringent, this was because of the issue discussed earlier regarding the sensitivity and poor viability of acutely isolated mature neurons. The first criterion was that for a cell to remain in the dataset it had to have >60% uniquely mapped reads. A non-uniquely mapped read is one that maps to multiple loci, meaning that one cannot unambiguously assign that read to one gene. Figure 3.7 shows a histogram for each of the samples and the percentage of reads that were uniquely mapped, unmapped, unassigned, multimapped, and featureless reads. Thus, using this criterion, 11 cells were removed from the initial 375 sequenced cells. Figure 3.8 is an extension of the analysis performed in the previous figure. Here, we analysed the distribution of uniquely mapped reads by plate (or flow cell). This was done to ensure that there were no significant differences in the read qualities generated in each of the flow cells—which would generate batch effects and thus confound future bioinformatics analysis. Although slight differences in the amplitude can be seen, the general distribution of the uniquely mapped reads are quite similar.

The second step in the quality control pipeline was to remove samples that had <500,000 unique reads mapped to genes. Recall that as per the experimental design we expected approximately 4,000,000 reads per cell. This analysis identified 8 cells that were below threshold. However, 5 cells overlapped with the 11 cells identified in the previous quality control step. Thus, we removed a total of 14 cells from the initial 375 sequenced cells leaving a total of 361 cells that would be further analysed. The mean number of detected genes for these cells was 1,382.

Another quick way to determine, at the population level, if the sequenced cells were of high quality and thus behaving as expected is to plot the expression level of all genes against the number of cells expressing those genes (Figure 3.9). It is expected that only some genes (such

as housekeeping genes) will be expressed at high levels by almost all of the cells, while most genes will be expressed at lower levels by fewer cells (these would be the genes that generate cellular heterogeneity in cell populations). Indeed, we found that our cell population displays the pattern expected of a population of high-quality cells.

Finally, as another check of the basic quality of the cells, we analysed their expression of known housekeeping genes. We hypothesised that if the cells were of high-quality they would express at least one of several housekeeping genes at high levels. In Figure 3.10 we plotted the expression of beta-Actin (Actb) as well as Calnexin (Canx) and beta-Glucuronidase (Gusb). Canx and Gusb were identified, in an analysis performed by the Surmeier laboratory, as the most stably expressed genes in striatal neurons and thus we adopted them as housekeeping/reference genes (Plotkin *et al.,* 2014). We found that nearly all the cells from all four plates expressed Actb at similarly high levels, indicating good quality cells. The expression of Canx and Gusb was more variable, but in general both were expressed at generally high levels by a large majority of the cells.

Taken together the quality control analysis indicated that we had 361 high-quality cells which could be further analysed bioinformatically.

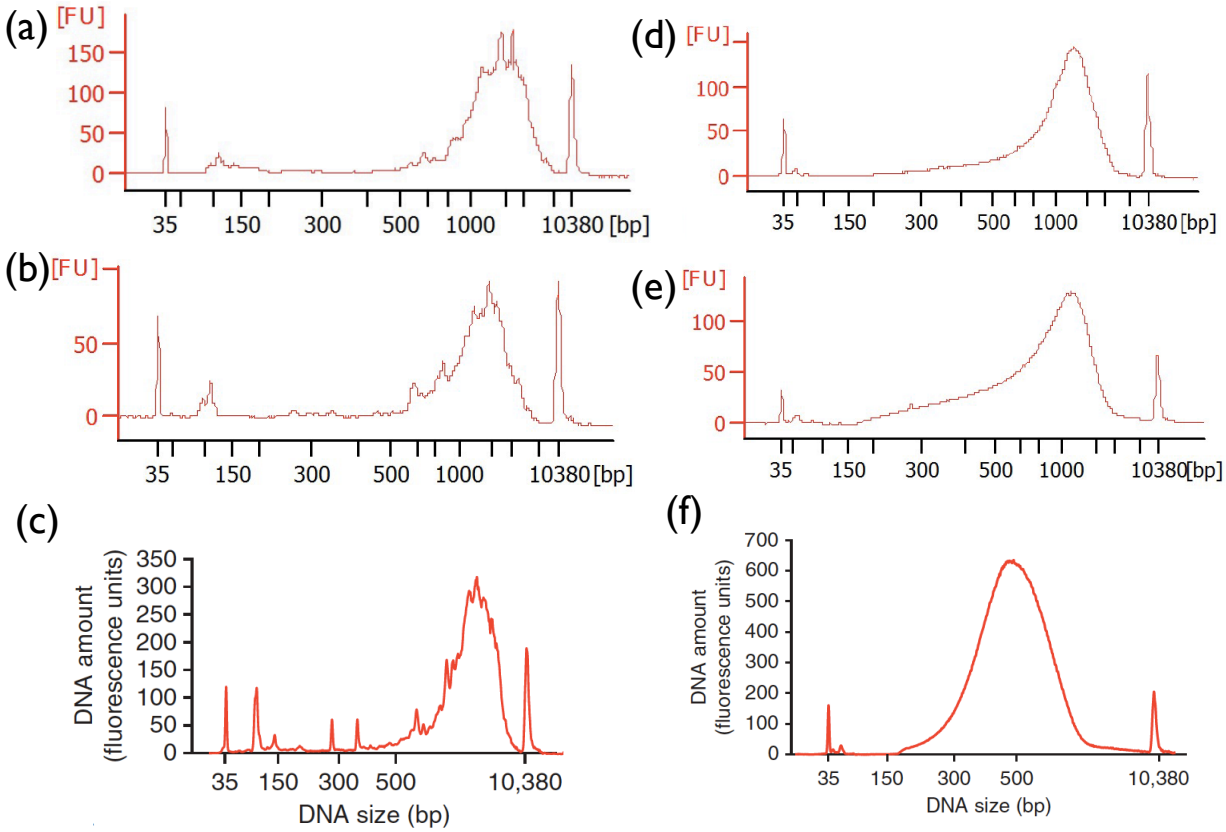


Figure 3.6. (a-c) Representative bioanalyzer profiles following cDNA amplification. Single neurons dissociated from the striata of 4-month-old IPNK mice. Following reverse transcription, amplification, and magnetic bead cleanup, the quality of the cDNA was assessed using the Agilent Bioanalyzer platform. Traces in **(a)** and **(b)** are cDNA profiles from single cells from two different plates. **(c)** An idealised profile of an undegraded cell from Picelli *et al.* (2014). **(d-f) Representative Bioanalyzer profiles following tagmentation.** Single neurons dissociated from the striata of 4-month-old IPNK mice. Following tagmentation and magnetic bead cleanup the quality of the cDNA was assessed using the Agilent Bioanalyzer platform. Traces in **(d)** and **(e)** are cDNA profiles from single cells from two different plates. **(f)** An idealised profile from Picelli *et al.* (2014).

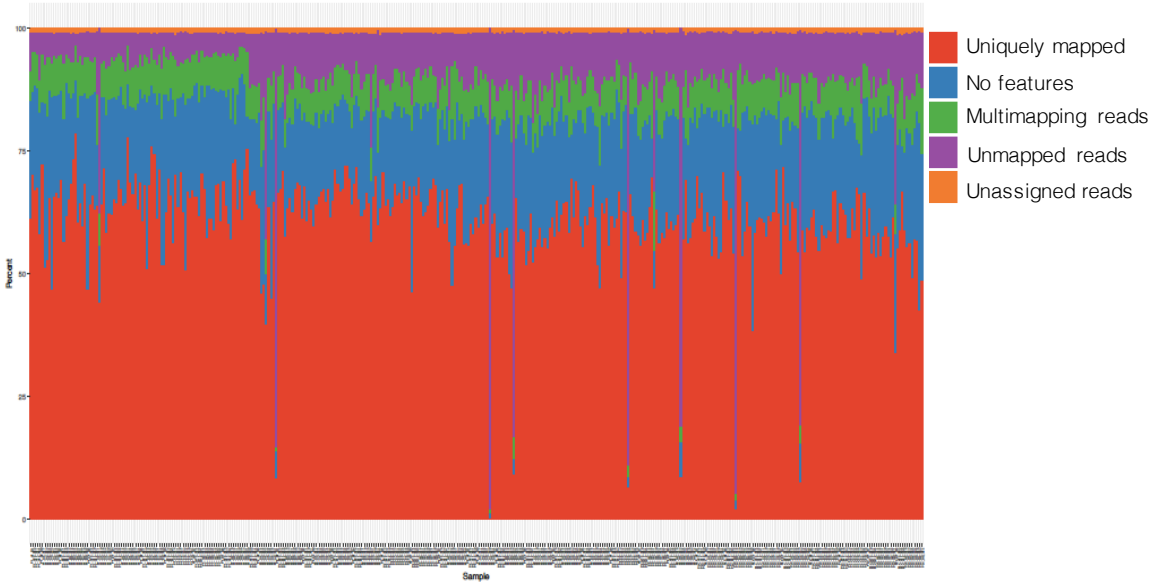


Figure 3.7. A summary of read types in all 375 sequenced cells. Barplot shows the percentage (y-axis) of uniquely mapped reads (RED) in every cell (x-axis). The percentage of reads without features (BLUE), unmapped reads (PURPLE), unassigned reads (ORANGE), and multimapping reads (GREEN) are also plotted. All except 14 cells have greater than 60% uniquely mapped reads, indicating high quality.

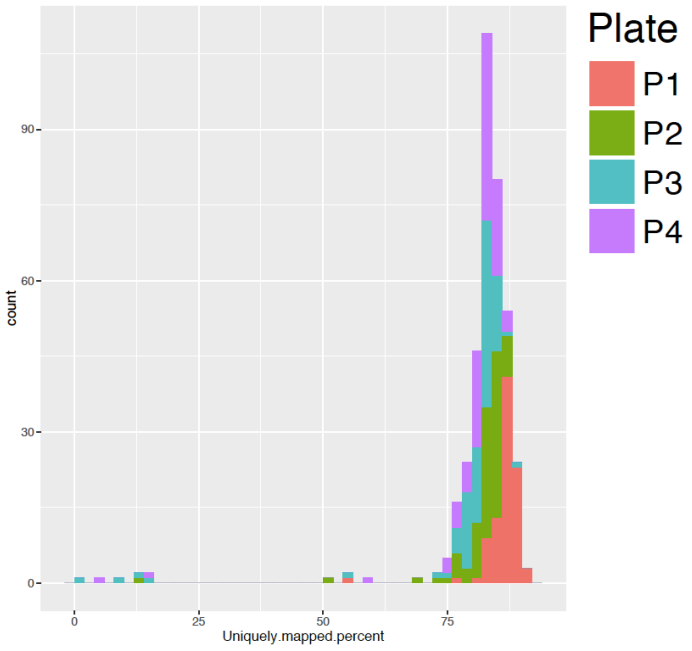


Figure 3.8. The distribution of uniquely mapped reads, by plate. A frequency histogram, generated using fastQC data, shows the distribution of the percentage of uniquely mapped reads by plate, indicated by different colours. This plot indicates the majority of the cells in each plate have >75% uniquely mapped reads. This is an indication of a high-quality dataset.

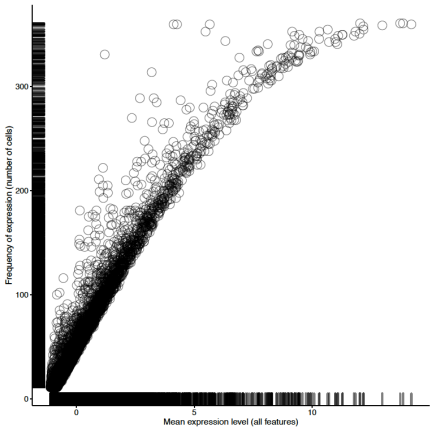


Figure 3.9. Plot of expression frequency versus expression level for all detected genes. Each circle represents a gene. The expression level of each gene is plotted against the number of cells that the gene is detected in. The plot shows that very few genes are expressed at high levels in nearly all the cells—these would be so-called "housekeeping" genes. The plot also shows that most genes are expressed at low levels and that this expression in not consistent among all the sequenced cells.

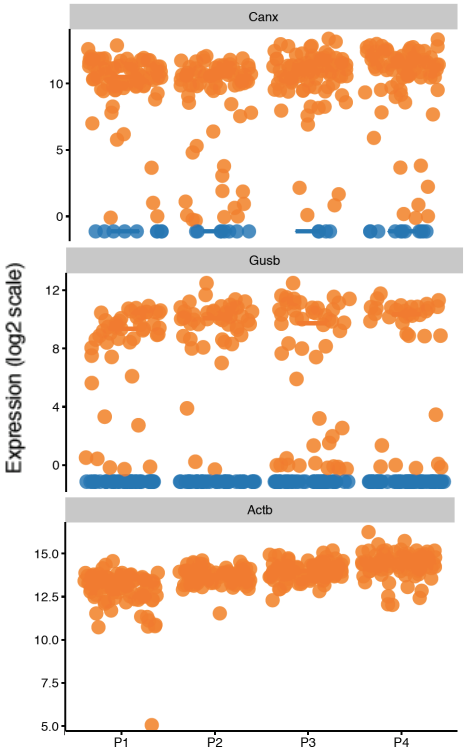


Figure 3.10. Plots showing the expression levels of three "housekeeping"/reference genes in single cells. Each cell was analysed for the expression of three housekeeping/reference genes: Actb, Canx, and Gusb (on the *y*-axis). Cells are grouped by plate (on the *x*-axis). ***Orange*** indicates cells in which expression was detected and ***blue*** indicates cells in which the expression of the gene was not detected.

3.7 Unbiased clustering of Penk-tdTomato^{+} cells.

Clustering of the 361 cells was performed by principal component analysis (PCA) as part of the PAGODA (Pathway and Gene Set Overdispersion Analysis) methodology (Fan *et al.*, 2016). This method was employed because it has been shown to account for possible batch effects that arise in single cell sequencing experiments. Figure 3.11 shows the result from this initial clustering. Three different clusters emerged from this analysis: one large cluster called Group320, and two smaller clusters called Group15 and Group26; each group was named after the number of cells that comprised it. Our initial hypothesis was that Group320 consisted of mature MSNs while Group15 and Group26 were two different cell types from the V-SVZ. Thus, we performed more in-depth gene expression analysis to test this hypothesis.

We began by performing differential gene expression analysis using the SCDE package (Kharchenko *et al.,* 2014) on the clusters that were generated by the PAGODA analysis. In essence, we used this analysis to determine which genes were upregulated and which genes were downregulated in each of the clusters. When we performed the analysis on the three clusters we found that the most upregulated genes in Group320 were not those that were canonically expressed by MSNs—they were instead microglial genes (Holtman *et al.,* 2015). Figure 3.12 compares the expression of four established microglial markers (P2ry12, Csf1r, Tmem119, Cx3cr1) in Group320, Group26, and Group15. We saw that the microglial genes were massively upregulated in Group320 while they were virtually undetected in the two smaller clusters. Thus, the Group320 cluster was contaminated with a microglial signature while the two smaller clusters, Group26 and Group15, were not. Why the two smaller clusters were not contaminated is most likely due to the identity of the different cell types that the clusters represent and thus their increased robustness and resistance to degradation and microglial attack (discussed further in ***Results*** and more extensively in the ***Discussion***).

In light of this, we decided to concentrate on Group26 and Group15 and redo the PAGODA and differential gene expression analysis on only the 41 cells from these two groups. Figure 3.13 shows the PCA plot that resulted from this analysis. Interestingly, we saw that the cells re-cluster into the same two groups—one with 26 cells and the other with 15 cells. Seeing that these cells were unaffected by microglial contamination and also unaffected by the removal of the cells displaying the microglial signature, we continued our analyses on these two clusters of cells.

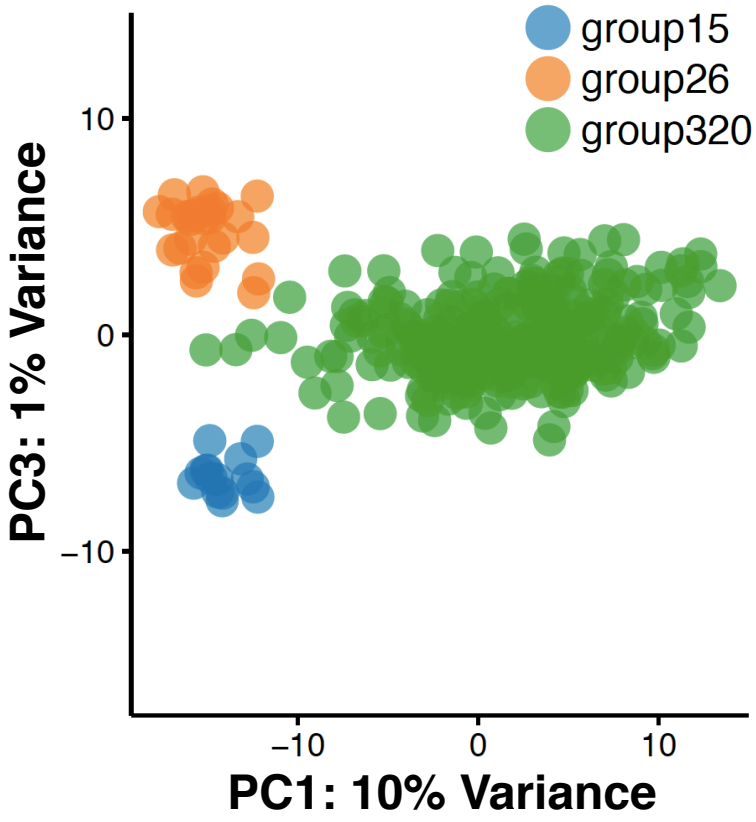


Figure 3.11. A PCA plot of the 361 cells reveals three distinct clusters. The normalised gene expression matrix from the PAGODA (Pathway and Gene Set Overdispersion Analysis) methodology was used to perform principal component analysis using the 500 most differentially expressed genes. The cells are plotted for their 1st and 3rd principal components. This analysis generates three clearly separated clusters: Group15, Group26, and Group320.

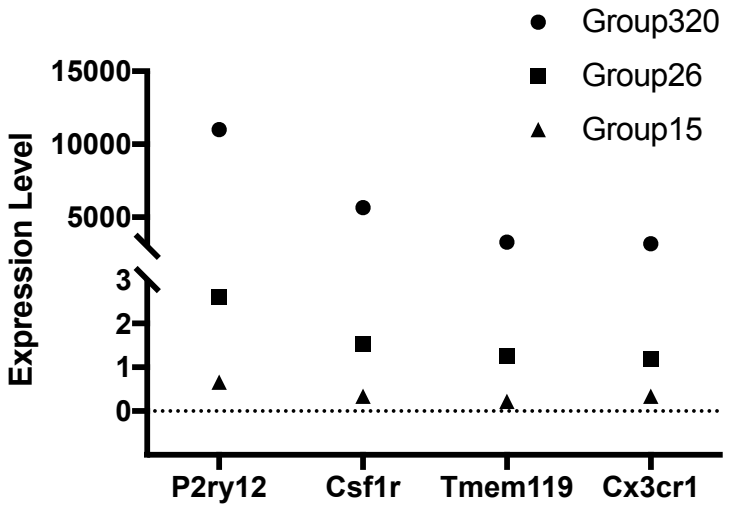


Figure 3.12. A graph of the differential expression levels of four selected microglial markers across the three clusters. The average expression values for each gene was taken from the SCDE analysis matrix and plotted on the *y*-axis (note the break in the axis). The microglial signature genes are listed on the *x*-axis.

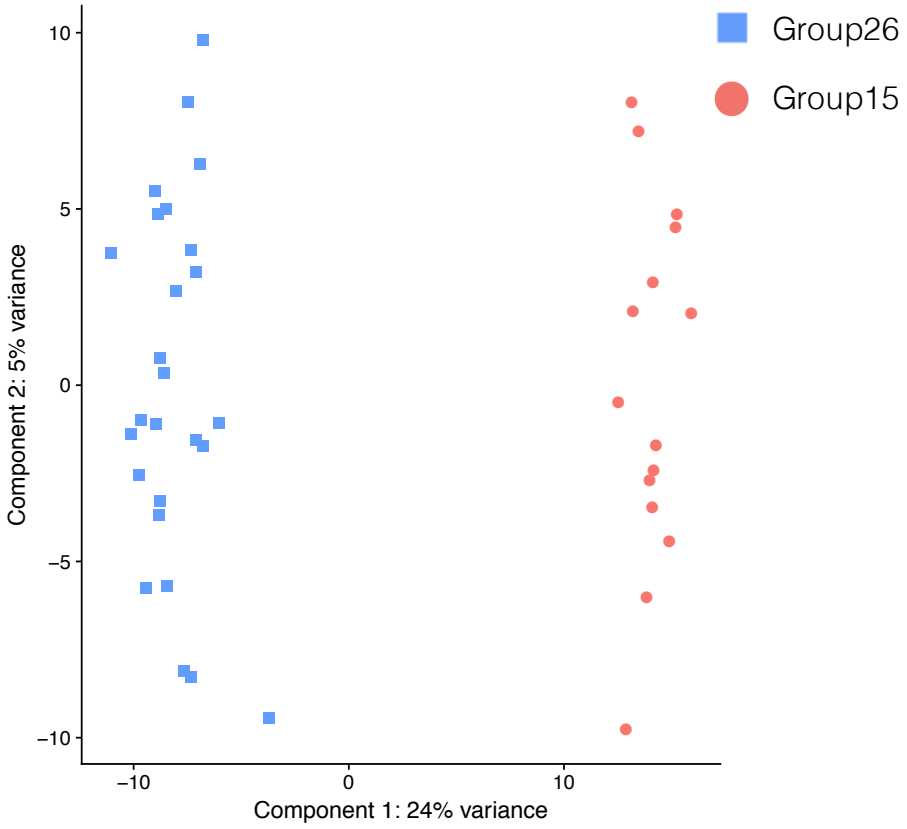


Figure 3.13. A PCA plot of 41 cells from Group26 and Group15 in absence of Group320. The cells from Group320 were omitted and PAGODA analysis was performed as before. The normalised expression matrix was used to generate the PCA plot using the 500 most differentially expressed genes. When the cells are plotted for their 1st and 2nd components, Group26 and Group15 separate as seen before in **Figure 3.11**.

3.8 Establishing the cellular identities of the clusters.

In line with our hypothesis that the two smaller clusters represent cell types originating from the V-SVZ and not the striatal parenchyma, we did not detect the expression of the mu-kappa-, and delta-opioid receptors, which are known to be expressed by MSNs and critical for their function. Ctip2 (also identified as Bcl11b) is a transcription factor known to be critical for the development of MSNs, and the cyto-architecture of the striatum as a whole (Arlotta *et al.,* 2008). We, therefore, analysed if these two clusters expressed this key transcription factor—shown alongside the expression of β-actin as a reference (Figure 3.14). In support of our hypothesis, we found that this key gene is virtually not expressed by the cells that form these clusters—Group26 and Group15—and therefore, they are likely not MSNs.

To eliminate the possibility that the clusters represent other striatal parenchymal cells, we next analysed whether the cells in the two clusters expressed any interneuronal genes. Striatal interneurons are generated in the medial ganglionic eminence of the embryonic brain and most express the transcription factors Nkx2.1 and Mash1 (Marin *et al.,* 2000). We did not detect the expression of either *Nkx2.1* or *Mash1* in our dataset. This provided the initial evidence that neither of the two clusters represent mature striatal interneurons. To analyse this further, we queried the dataset for the expression of mature interneuron markers (Tepper *et al.,* 2010). While we did detect some expression of the glutamate decarboxylase genes, *Gad1* and *Gad2* (genes required for GABA synthesis), we did not detect any expression of the vesicular inhibitory amino acid transporter (*Viaat*)—which is required for the packaging and release of GABAergic vesicles (Westmoreland *et al.,* 2001). Furthermore, we did not detect the expression of specific markers of interneurons known to populate the striatum: Parvalbumin (*PV*), Neuropeptide Y (*NPY*), Somatostatin (*SST*), Nitric Oxide Synthase (*NOS*), Calretinin (*Calb2*), or Choline O-Acetyltransferase (*ChAT*). The combined lack of general interneuronal transcription factors,

GABA synthesis machinery, and specific mature interneuronal markers led us to conclude that the two clusters did not represent populations of striatal interneurons.

Finally, we determined whether the clusters may present oligodendrocytes or astrocytes. We did not detect the expression of the oligodendrocyte lineage specific transcription factor *Sox10* (Kuhlbrodt *et al.,* 1998). We also did not detect the expression of immature oligodendrocyte marker *NG2* and *PDGF-α* (Baumann and Pham-Dinh, 2001). Finally, we did not detect the expression of the mature oligodendrocyte marker Olig2. We determined that the clusters also did not represent astrocytes due to their lack of expression of two key astrocyte genes: *Aldh1L1* and *S100β* (Cahoy *et al.,* 2008).

Taken together, this analysis led us to conclude that the clusters of cells in Group26 and Group15 were neither MSNs, interneurons, oligodendrocytes or astrocytes, as we had hypothesised initially. Having eliminated this possibility, we began to analyse whether the two clusters did indeed present cell types inhabiting the V-SVZ germinal zone.

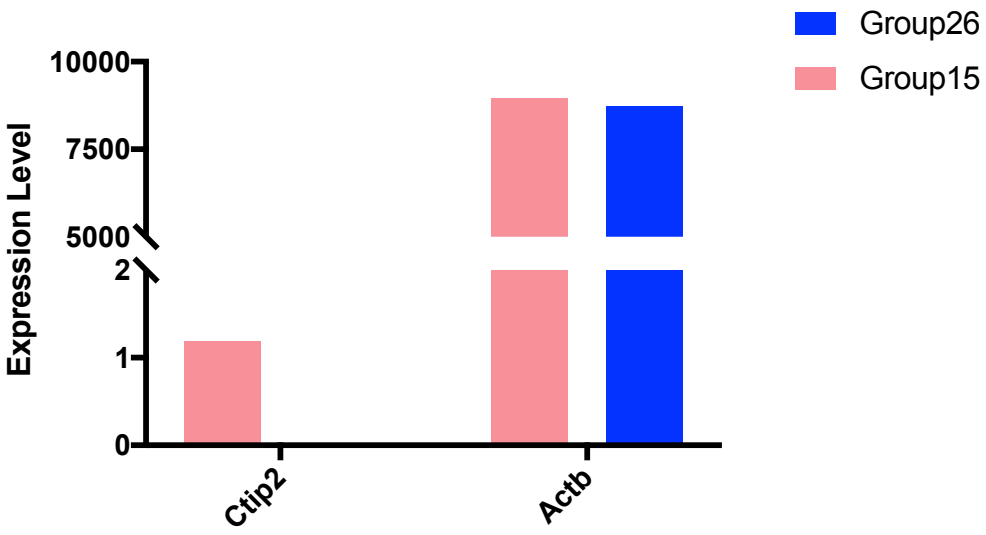


Figure 3.14. The mature MSN marker Ctip2 is virtually absent in Group26 and Group15 cells. The SCDE-derived average expression levels for the *Ctip2* and *Actb* genes plotted for Group26 and Group15. *Ctip2* is present at barely detectable levels in Group15 and completely absent in Group26. The "housekeeping" gene β-actin (*Actb*) is expressed at similarly high levels in both clusters. Note the break in the *y*-axis.

3.9 The Penk-tdTomato^{+} clusters display molecular signatures of V-SVZ neural stem cells at two stages of differentiation.

The first evidence that the clusters may represent V-SVZ NSCs at two different stages of differentiation came from their mutually exclusive expression of several established markers (Figure 3.15). Firstly, *Dcx* (a widely-accepted marker of terminally differentiating NSCs or neuroblasts) was found to be expressed in Group26 and not Group15—though this difference was not statistically significant, as determined by the SCDE software. We also queried the expression of *βIII-tubulin*, which is another canonical marker of neuroblasts, and found it to be significantly upregulated ($p_{adjusted}$ = 0.002) in Group26 and undetected in Group15. Sox9 has been identified as a key protein expressed in quiescent NSCs, where its expression is required to maintain stem-ness and prevent terminal differentiation (Scott *et al.,* 2010). We found that it was upregulated, though this was not statistically significant, in Group15 and very minimally expressed in Group26. We then analysed the groups for the expression of Vascular Cell Adhesion Molecule-1 (*Vcam1*), a gene which is known to be expressed by the B-cells of the V-SVZ germinal niche (Kokovay *et al.,* 2012). We did not detect the expression of this NSC marker in Group26 but we did find it to be upregulated in Group15, though not significantly. Insulin-like Growth Factor Binding Protein 2 (*Igfbp2*) has been shown to be crucial for hippocampal neurogenesis, while it has also been postulated that quiescent B-cells should be responsive to circulating insulin-like growth factor (Zhang *et al.,* 2007; Mirzadeh *et al.,* 2008). We found that *Igfbp2* was significantly upregulated ($p_{adjusted}$ = 0.002) in Group15 and undetected in Group26. Taken together, this initial expression analysis suggested that we had isolated cells, on the basis of their expression of Penk-tdTomato, that are part of the V-SVZ stem cell niche. Group15 cells express markers of quiescent NSCs and therefore may represent a subset of B-cells, while Group26 cells express markers of neuroblasts or A-cells. Summarised in Table 3.1.

To analyse the expression of the genes of interest in a more granular manner, we then visualised them within and between the clusters of cells in a PCA plot. Sox11 is a transcription factor known to be expressed by terminally differentiating NSCs (Bergsland *et al.,* 2006). Figure 3.16a shows that the expression of *Sox11* is restricted to the cells forming the Group26 cluster. This upregulation is also statistically significant ($p_{adjusted}$ = 0.01). Neural Cell Adhesion Molecule 1 (*Ncam1*) is also known to be expressed by neuroblasts (Boutin *et al.,* 2009). Figure 3.16b shows that the expression of *Ncam1* is also restricted to Group26, and that this upregulation is statistically significant ($p_{adjusted}$ = 0.001). CD24a is a surface antigen that has been shown to be expressed by migrating neuroblasts (Calaora *et al.,* 1996). We found that its expression is virtually restricted to Group26, though the upregulation was not statistically significant (Figure 3.16c). This further supports the idea that Group26 cells are terminally differentiating NSCs and therefore neuroblasts. Summarised in Table 3.1.

Prolactin has been shown to increase neurogenesis in the V-SVZ (Shingo *et al.,* 2003; Mak *et al.,* 2007) and in a recent study published by the laboratory of Alvarez-Buylla it has been postulated that this is mediated by the expression of the prolactin receptor (Prlr) on the quiescent B-cells (Obernier *et al.,* 2018). In line with the hypothesis that the Group15 cells are quiescent NSCs, we found that the expression of *Prlr* is restricted to Group15 cells and that upregulation is also statistically significant ($p_{adjusted}$ = 0.002; Figure 3.17a). We also found that the KiR7.1 inwardly rectifying potassium channel (*Kcnj13*) is highly expressed and restricted to the Group15 cluster ($p_{adjusted}$ = 0.002; Figure 3.17b). This is a key finding, because Pruss *et al.* (2011) show that KiR7.1 is expressed in the "early" ($GFAP^+$) neurospheres derived from the V-SVZ but is absent in the "late" (Dcx^+) neurospheres. This supports the fact that the Group15 cluster is a molecularly distinct population of cells, and that they are most likely quiescent NSCs of the V-SVZ. Findings summarised in Table 3.1.

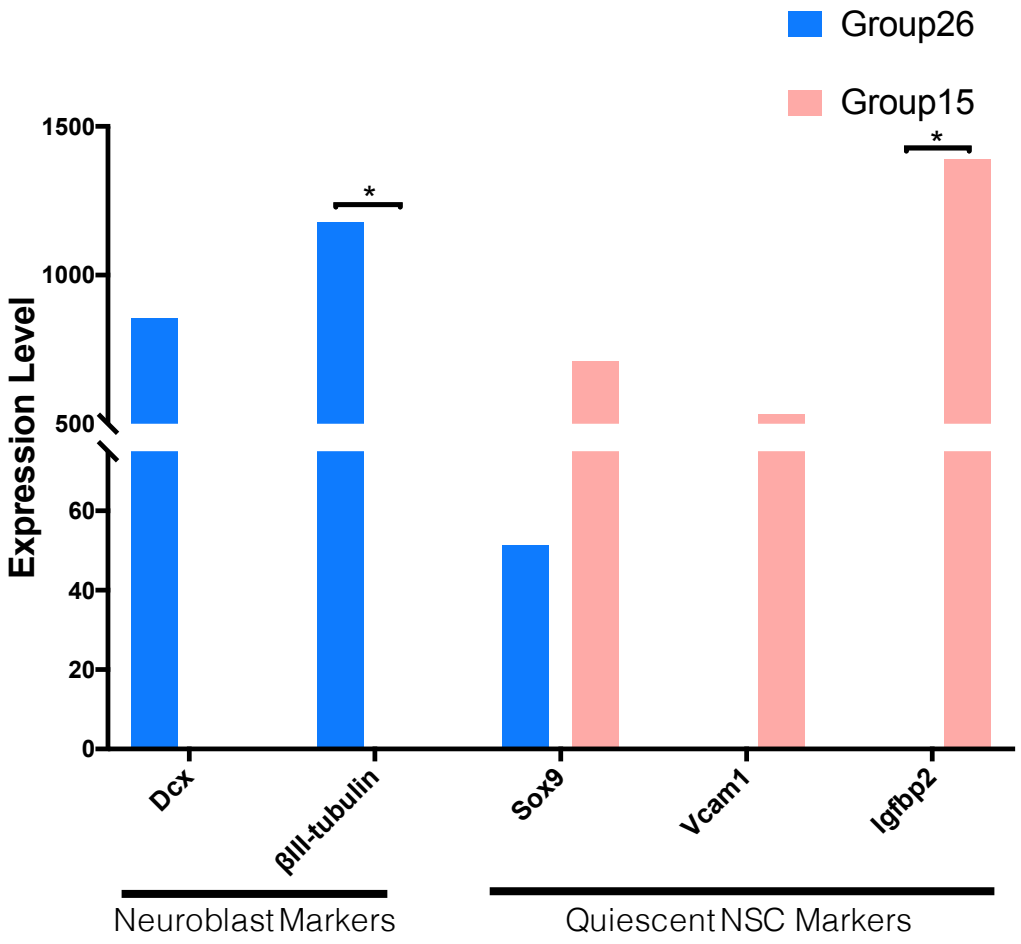


Figure 3.15. The differential expression of neuroblast markers in Group26 and quiescent stem cell markers in Group15. The average expression levels for each of the genes, as determined by SCDE, is plotted for Group26 and Group15. *Dcx* expression is not detected in Group15 and expressed relatively highly in Group26 (855.2), SCDE analysis states that this difference is not significant between the two groups. *βIII-tubulin* is expressed highly in Group26 (1176.8) but is not detected in Group15, this difference is significant ($p_{adjusted}$ = 0.002). *Sox9* is weakly expressed in Group26 (51.4) and relatively highly expressed in Group15 (710.4), this difference is not found to be significant. *Vcam1* expression is not detected in Group26 but is found to be expressed in higher, though not significantly, levels in Group15 (531.6). *Igfbp2* expression is not detected in Group26, but is found to be significantly upregulated in Group15 (1390.1; $p_{adjusted}$ = 0.002).

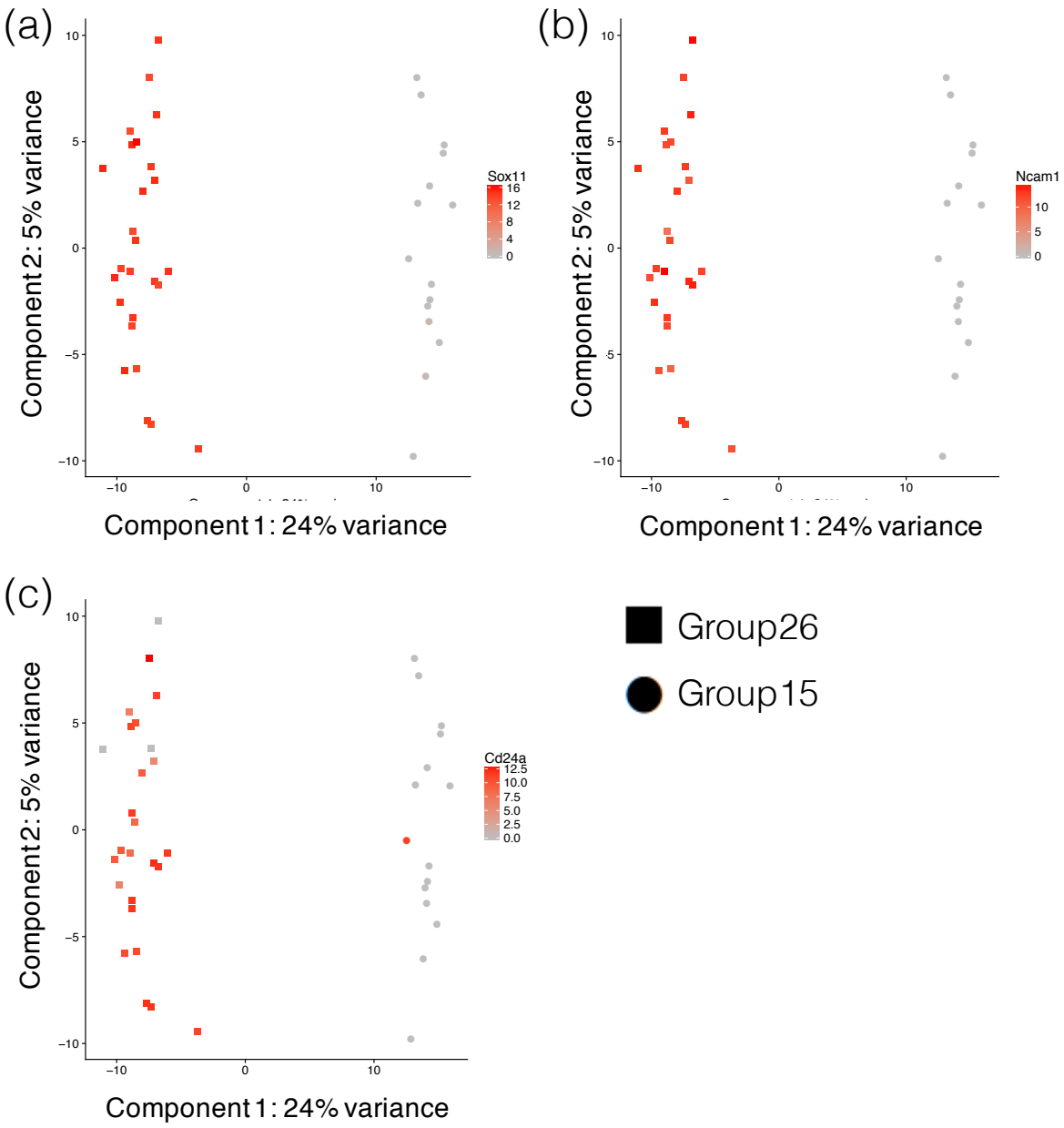


Figure 3.16. The differential expression of neuroblast markers in Group26. The PCA plot is identical to that in Figure 3.13 (plotted using the 500 most variable genes) except that the expression level for each gene of interest in each cell is also plotted. **(a)** Depicts the expression of *Sox11* in each of the sequenced cells. SCDE analysis finds it is significantly upregulated in Group26 ($p_{adjusted}$ = 0.01; average expression: Group26 = 10501.4, Group15 = 0.09). **(b)** Depicts the expression of *Ncam1* in each of the sequenced cells. It is significantly upregulated in Group26 ($p_{adjusted}$ = 0.001; average expression: Group26 = 2997.3, Group15 = 0). **(c)** Depicts the expression of *CD24a* in each of the sequenced cells. Its expression is virtually restricted to Group26, but this difference is not found to be significant (average expression: Group26 = 909.7, Group15 = 90.7).

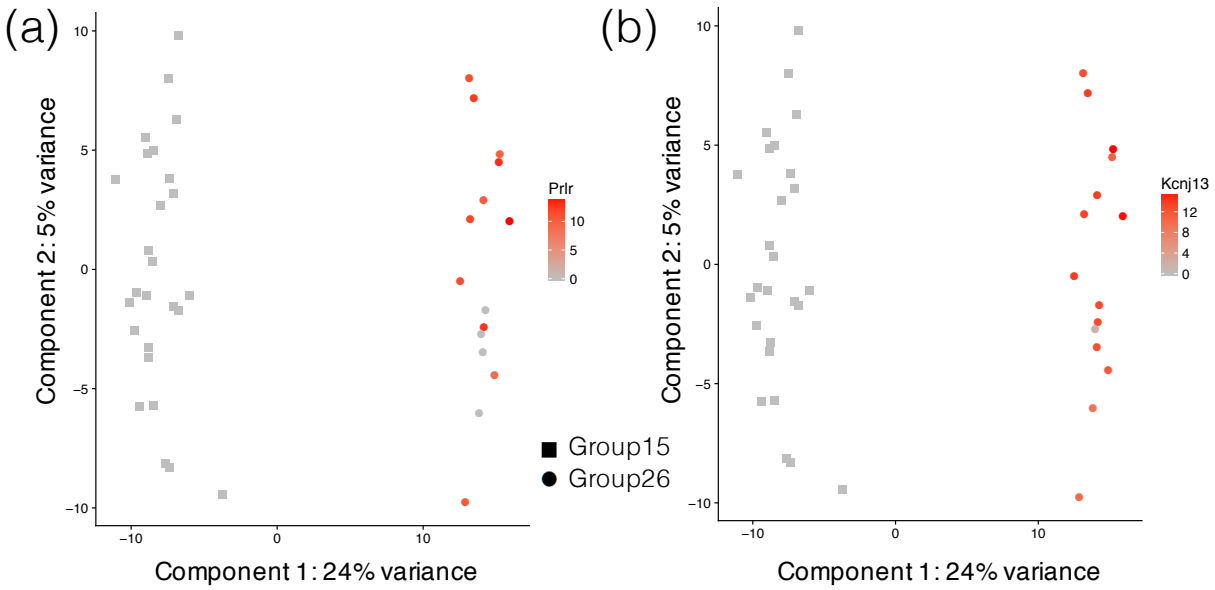


Figure 3.17. The differential expression of quiescent neural stem cell markers in Group15. The PCA plot is identical to that in Figure 3.16 (plotted using the 500 most variable genes) except that the expression level for each gene of interest in each cell is also plotted. **(a)** Depicts the expression of *Prlr* in each of the sequenced cells. SCDE analysis finds it is significantly upregulated in Group15 ($p_{adjusted}$ = 0.002; average expression: Group15 = 1081.4, Group26 = 0). **(b)** Depicts the expression of *Kcnj13* in each of the sequenced cells. It is significantly upregulated in Group15 ($p_{adjusted}$ = 0.002; average expression: Group15 = 4157.2, Group26 = 0).

Table 3.1

Gene Markers	
Group15	**Group26**
Sox9	*Dcx*
Vcam1	*βIII-tubulin*
Igfbp2	*CD24a*
Prlr	*Sox11*
Kcnj13	*Ncam1*

3.10 *In situ* validation of the StriatumSeq1 clusters on IPNK brain sections.

We then wanted to test whether we could indeed detect the markers of either cluster *in situ* in the V-SVZ of an IPNK brain section before proceeding with a more refined single cell sequencing experiment. Figure 3.18 shows a representative image from this preliminary analysis; here we immunostained a cryosectioned IPNK brain with antibodies against Sox11 (Group26 marker) and Sox9 (Group15 marker). It can be seen that Penk-tdTomato^{+} cells colocalise with either Sox11 or Sox9, but not with both. As expected, Sox11 and Sox9 are present in the striatal parenchyma where they are known to label differentiated neurons and astrocytes, respectively. This preliminary result supported the existence of the clusters *in situ* and provided the first evidence that stem cells within the V-SVZ germinal niche expressed Enkephalin at some point during their development, and therefore that they could be identified using our Penk-tdTomato reporter line. In order to confirm this, we sought to repeat the single cell sequencing under more refined conditions to ensure that we could initially exclude cells displaying a microglial signature.

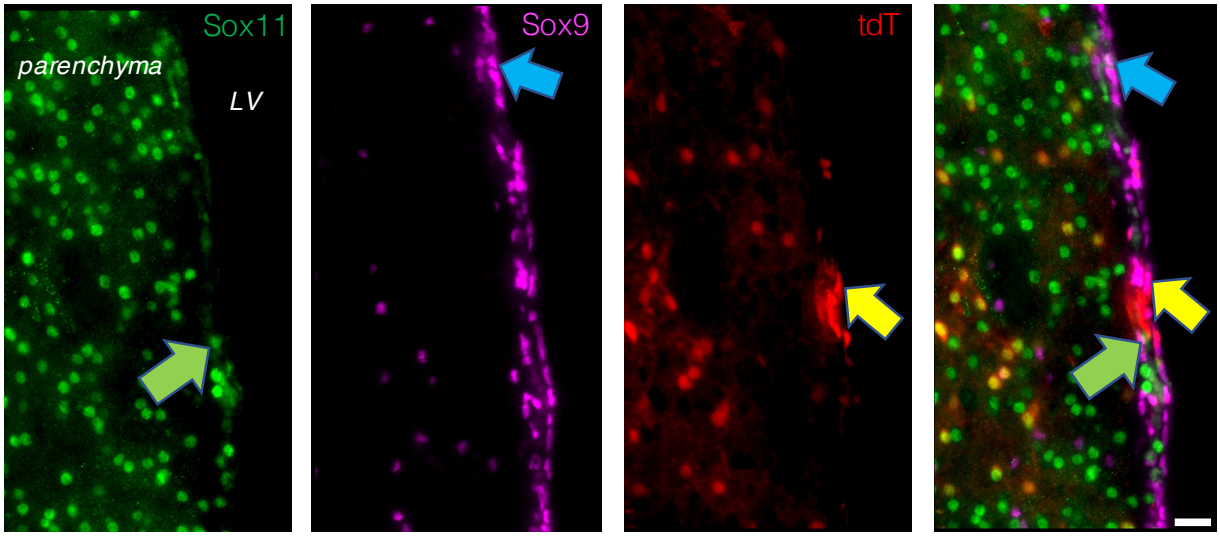


Figure 3.18. *In vivo* validation of the expression of neuroblast and quiescent NSC markers in Penk-tdTomato^{+} cells of the V-SVZ. Representative IPNK brain section (30 µm) immunostained with antibodies against Sox11 (in ***green***) and Sox9 (in ***violet***). Within the LV, Penk-tdTomato^{+} cells are either Sox9^{+} (yellow arrow) or Sox11^{+} (green arrow). We did not see triple colocalisation at the LV. There are cells that are Penk-tdTomato^{-}/Sox9^{+}/Sox11^{+} (blue arrow). Scale bar (in white) = 20µm.

3.11 Elimination of the microglial signature by FACS.

We hypothesised that one reason why the microglial signature arose in the Group320 cluster was because of the Penk-tdTomato$^+$ cells dying due to mechanical stress and environmental changes and were thus being engulfed by microglia. Therefore, we needed a method to eliminate any cells (regardless of whether they were Penk-tdTomato$^+$) displaying a microglial signature during the cell sorting step. The cell surface antigen CD11b (also called the Macrophage Antigen Alpha Polypeptide) is known to be an effective and widely used marker of macrophages (Zhang *et al.,* 2008). We reasoned that if we could eliminate any cells that were immunopositive for CD11b (CD11b$^+$) and thus only collect cells that were Penk-tdTomato$^+$/CD11b$^-$, we would be able to eliminate the microglial signature and capture pure V-SVZ cells. To perform this, we obtained a CD11b antibody conjugated with the far-red fluorophore allophycocyanin (APC; 650nm absorbance maximum), so that we could perform FACS-based purification. IPNK brains were prepared for FACS as described above but now with the additional step of a CD11b-APC antibody incubation. These preparations were then sorted (Figure 3.19a). While Penk-tdTomato$^+$ were quite rare, Penk-tdTomato$^+$/CD11b$^-$ cells were rarer still. Although the cells were rare it seemed that the CD11b antibody was effective in eliminating the microglial signature, thus allowing us to eliminate microglia cells from the cell suspension.

Since Penk-tdTomato$^+$/CD11b$^-$ were rare, we also analysed where the majority of Penk-tdTomato$^+$/CD11b$^-$ lie in the FACS plot and if we could come up with an alternative strategy to maximise our ability to collect these cells. Through trial-and-error, we found that a large proportion of the cells that were Penk-tdTomato$^+$/CD11b$^-$ were located in the middle peak of the LIVE/DEAD plot (Figure 3.19b; heretofore called “M-Peak”). This led us to believe that this peak may be enriched in live Penk-tdTomato cells. Thus, we collected samples from this peak to test whether this gate was in fact enriched in neurons and if we were successful in eliminating cells with a microglial signature.

3.12 Validation of microglial signature elimination and M-Peak cells.

Prior to proceeding with RNA sequencing (which is extremely labour intensive and costly) using the strategy outlined above, we wanted to know with certainty that it could eliminate the microglial signature. We have developed and published a method to perform microfluidic-based high-throughput gene expression analysis of mature neurons, both in bulk- and single-cell paradigms (Nath *et al.,* 2017). These methods allow to assess, with high sensitivity, whether cells are expressing certain genes of interest. Thus, I reasoned that I could use this method, to first, test the sorted cells for their expression of key microglial genes to determine whether we had in fact eliminated Penk-tdTomato$^+$ cells with a microglial signature. Secondly, we would also be able to determine whether the Penk-tdTomato$^+$/CD11b$^-$ cells collected from the M-Peak were in fact neurons. Finally, we also included three other gating paradigms as a means of quality control.

To perform this analysis, we created cell suspensions from the striata of three male IPNK mice, of three months of age. We then prepared the samples for sorting as outlined previously. Three replicates of 50 cells were collected from each of the brains for the following five gating conditions: Penk-tdTomato$^+$/CD11b$^-$ (classical gate); M-Peak (Penk-tdTomato$^+$/CD11b$^-$; new gate); Penk-tdTomato$^+$/CD11b$^+$ (classical gate); Penk-tdTomato$^-$/CD11b$^+$ (classical gate); Penk-tdTomato$^-$ (no selection for Cd11b; classical gate). The cells were collected and prepared for the high-throughput gene expression analysis. We tested each population for striatal MSN genes (*Penk, Drd1, Drd2*), a general neuronal marker (*Snap25*), and finally also for three microglial genes (*Tmem119, Sfpi1, Csf1r*). The experiment was performed and analysed as described previously in ***Materials and Methods*** as well as in Nath *et al.,* (2017). The *ΔCt* values for the three replicates were averaged. The means for the three biological replicates were also averaged to produce the heatmap in Figure 3.20, using the Morpheus software (Broad Institute). The heatmap produces a visual output of the relative expression levels for each of the genes across the five gating paradigms. From the analysis, we saw first that the Penk-tdTomato$^+$/CD11b$^-$

(classical) gate allowed us to isolate cells that are depleted in the microglial markers while being enriched for some neuronal markers. The M-Peak (Penk-tdTomato^{+}/CD11b^{-}) produced a population of cells that were similarly depleted in microglial markers, but expressed neuronal genes at higher levels than the classical gating paradigm. Since the classical gate led to the isolation of stem cells in the sequencing experiment and since it seemed to express mature neuronal markers at lower levels, we reasoned that this gate was enriched in V-SVZ cells, while the M-Peak (Penk-tdTomato^{+}/CD11b^{-}) was enriched in mature neurons of the striatum. The three other gating paradigms capture cells that highly express microglial markers and which are, in turn, reduced in the expression of neuronal markers.

Having completed this analysis, we felt confident that we would be able to specifically sort and sequence single cells of the V-SVZ using the Penk-tdTomato^{+}/CD11b^{-} (classical) gate.

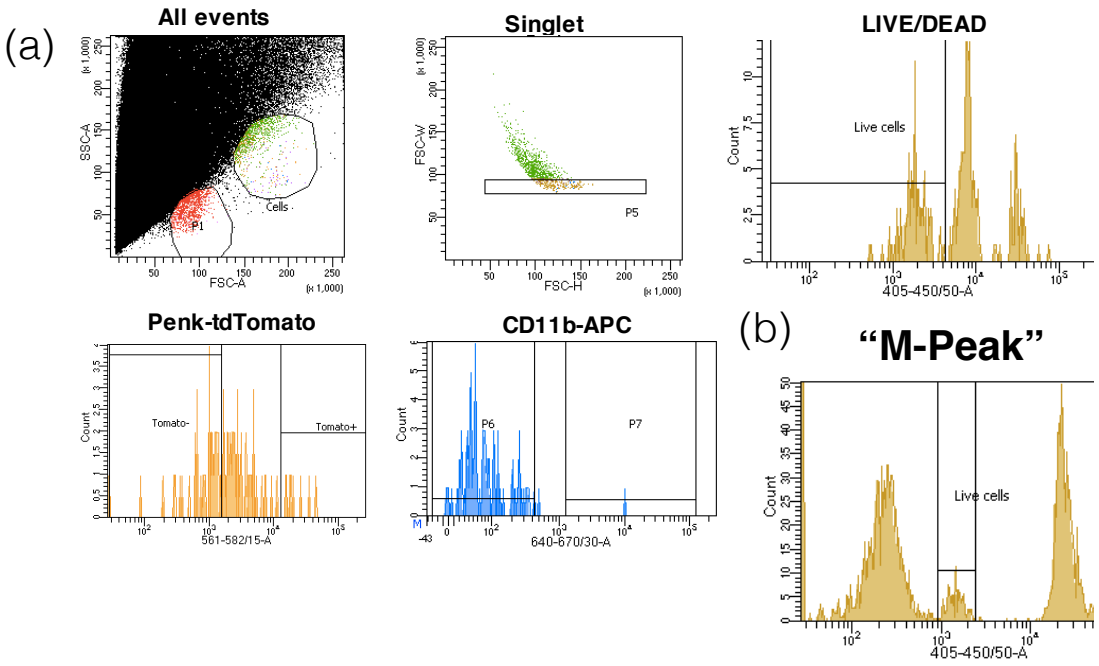


Figure 3.19. (a) Elimination of the microglial signature using an antibody against CD11b. The gating protocol is identical to that shown before in Figure 3.5 except for the addition of an antibody against CD11b. This is detected using the APC fluorophore, shown in the panel at the bottom right in ***blue***. The cells in the "P6" gate are collected and said to be CD11b^{-}. **(b) Using a new gating mechanism to capture a novel cell population.** Only the LIVE/DEAD plot is shown (as all other plots are identical) and within the plot, the cells that are in the "Live cells" gate are captured. Because the population lies in the middle peak, this population is termed the "M-Peak" population. The middle peak can also be seen in LIVE/DEAD plot in **(a)**.

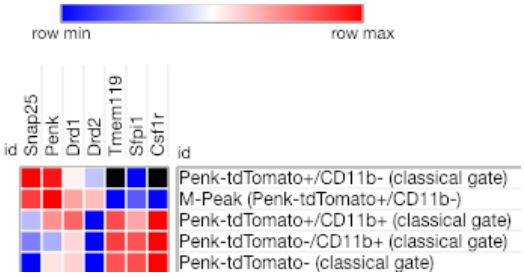


Figure 3.20. Elimination of the microglial signature and the isolation of Group1 cells and mature neurons. The averaged relative gene expression heatmap generated using Morpheus is shown. Three replicates of 50 cells were collected from each of the three brains for the following five gating conditions: Penk-tdTomato^{+}/CD11b^{-} (classical gate); M-Peak (Penk-tdTomato^{+}/CD11b^{-}; new gate); Penk-tdTomato^{+}/CD11b^{+} (classical gate); Penk-tdTomato^{-}/CD11b^{+} (classical gate); Penk-tdTomato^{-} (no selection for Cd11b; classical gate). These samples were then prepared for Fluidigm Biomark analysis using the TaqMan assay system. Each population was tested for expression of microglial genes (*Tmem119, Sfpi1, Csf1r*), mature MSN markers (*Penk, Drd1, Drd2*), and a pan-neuronal marker. The *ΔCt* values for the three replicates were averaged. The means for the three biological replicates were also averaged to produce the heatmap. Red indicates higher expression, blue indicates lower expression.

3.13 StriatumSeq2: Deep RNA sequencing of single Penk-tdTomato$^+$ cells of the V-SVZ.

Our first attempt, StriatumSeq1, provided us with key insights about the Penk-tdTomato$^+$ cells that reside in the V-SVZ. However, having discovered how to more specifically isolate these cells, we wanted to repeat the single cell RNA sequencing. This would allow us to see if the isolation and sequencing of the putative NSC populations could be replicated, albeit in a much more refined experimental set up.

To achieve this, we sought to collect a full 96-well plate of cells from the Penk-tdTomato$^+$/CD11b$^-$ (classical) gate. However, our earlier sorts indicated that these cells were exceedingly rare. Thus, we prepared cell suspensions from the striata of five different female IPNK mice (9 weeks of age)—keeping each sample separate. Of the initial 96 cells, 20 cells were excluded prior to sequencing due to either poor cDNA amplification or tagmentation. The 76 cells were processed, sequenced, and quality control analysis performed as described for StriatumSeq1. A further 37 cells were excluded during the quality control analysis as result of having <60% uniquely mapped reads or <500,000 uniquely mapped reads. The average number of reads that mapped to exonic regions was 1,873,440 (again, ensuring that we were performing deep sequencing). The average number of detected genes was 1,020.

3.14 Clustering of the Penk-tdTomato$^+$ cells.

As for StriatumSeq1, clustering of the 76 cells was performed by principal component analysis (PCA) as part of the PAGODA methodology (Fan *et al.,* 2016) using the 500 most variable genes. Figure 3.21 shows that analysis produces three distinct clusters, compared to the two that were generated in StriatumSeq1. This could be due to the fact that a more stringent gating paradigm was used for StriatumSeq2, but also to the fact that in the second experiment we did use younger animals (2 months old versus 4 months old) and it could be that this allowed us to isolate the extra group (see ***Discussion***). We then determined whether the clusters

generated in StriatumSeq2 overlapped with the clusters generated in StriatumSeq1. Indeed, in Figure 3.22 we saw that the two larger groups overlapped almost completely while the smaller, as expected, did not. Since we had successfully repeated the experiment and the clustering, it indicated that the clusters resulted most likely due to the biological features of the cell types and not due to just *in silico* parsing and analysis.

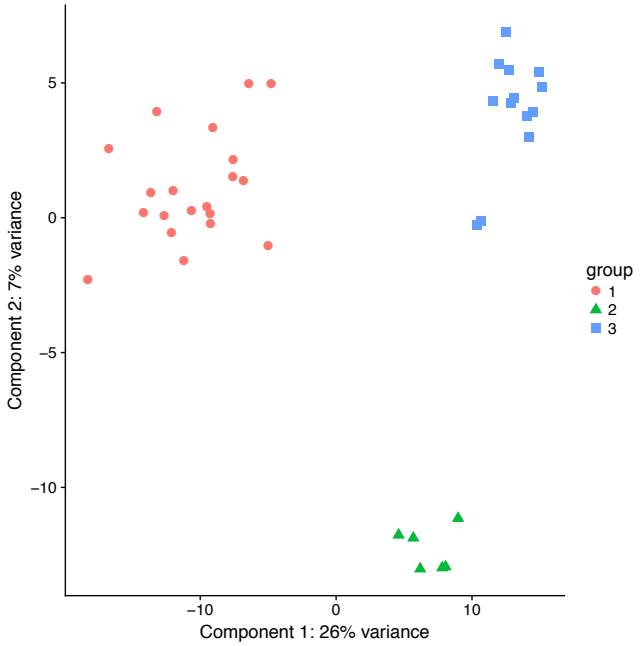


Figure 3.21. Sequenced cells in StriatumSeq2 produce three distinct clusters of cells. Following quality control, for the remaining 39 cells the 500 most variably expressed genes were used to generate the PCA plot. When the 1st and 2nd components are graphed, three distinct clusters emerge—heretofore called Group1, Group2, and Group3.

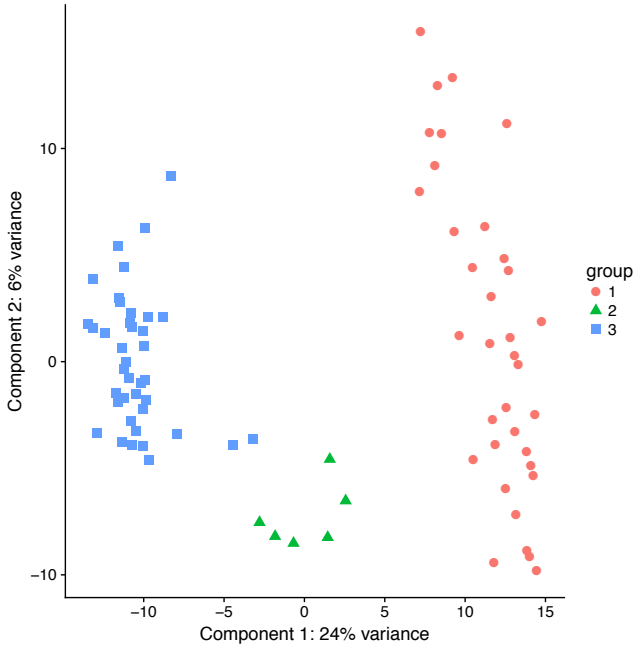


Figure 3.22. Sequenced cells in StriatumSeq2 and StriatumSeq1 reproduce the same clusters. The cells from StriatumSeq1 and StriatumSeq2 are plotted on the PCA plot. Group1 overlaps with the previous Group15. Group3 overlaps with the previous Group26. Group2 is unique to StriatumSeq2, as the six cells do not overlap with any cells in StriatumSeq1.

3.15 The Penk-tdTomato^{+} reporter line captures V-SVZ NSCs at three stages of the differentiation arc.

We then performed further gene expression analysis on the clusters to determine if they were indeed NSCs of the V-SVZ at different stages of differentiation, as the gene expression analysis for Group15 and Group26 in StriatumSeq1 indicated. To do this, we employed a bioinformatics tool called Pseudotime analysis performed with the TSCAN software package (Ji and Ji, 2016). Stated simply, the method allows for the ordering of cells (or clusters of cells) in a pseudo-temporal path by detecting gradual changes in the transcriptomes of said cells or clusters. We reasoned that, using our dataset, it should be able to order the cells and the three clusters in a differentiation arc—from the most quiescent NSCs to terminally differentiating neuroblasts. Figure 3.23 shows the result from that analysis. It places what we have termed Group1 at one end of the pseudo-temporal path, Group3 at the opposite end, and places Group2 in the middle. This analysis provides the relative distance between groups, but does not provide information regarding the absolute beginning and the absolute end of the arc—this has to be done manually by the expression of known markers. In our case, in StriatumSeq1, we had labeled Group15 as the quiescent NSC group and Group26 as the neuroblast group based on marker expression. Thus, we repeated the Pseudotime analysis with the StriatumSeq1 and the StriatumSeq2 data sets (Figure 3.24). Here we see that Group15 (StriatumSeq1) overlaps with Group1 (StriatumSeq2), indicating that Group1 is likely also quiescent NSCs. We also found that Group26 (StriatumSeq1) overlaps with Group3 (StriatumSeq1), indicating that Group3 is likely a cluster of neuroblasts. Interestingly, Group2 resides between the putative quiescent NSCs and neuroblasts within this differentiation arc.

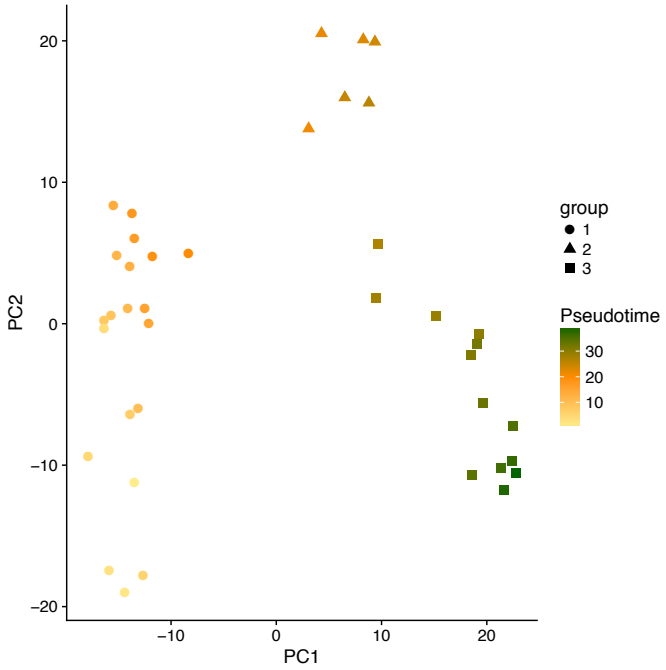


Figure 3.23. Pseudotime analysis of Striatumseq2 clusters places Group1 and Group3 at opposite ends of a differentiation arc. Pseudotime analysis of the 39 cells was performed using the TSCAN package. The cells are ordered algorithmically using a cluster-based minimum-spanning tree approach, this is projected on the plot using the first two principal components. The Pseudotime units indicate how close two cells are pseudotemporally. This analysis placed Group1 and Group3 at the opposite ends of the pseudotime path, with Group2 in the middle.

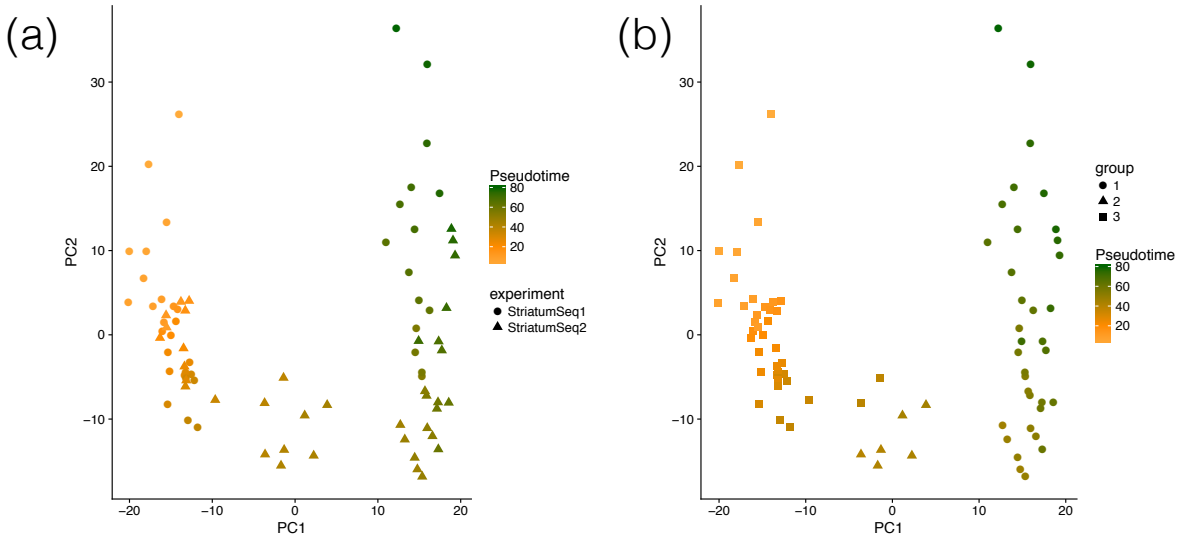


Figure 3.24. Pseudotime analysis of the combined cells from StriatumSeq1 and StriatumSeq2. The analysis was performed on both groups as stated in Figure 3.23. **(a)** The plot differentiates cells from StriatumSeq1 and StriatumSeq2. **(b)** The plot differentiates the three different clusters (Group1 = Group15; Group3 = Group26). The pseudotemporal ordering is maintained when both experiments are combined.

3.16 Cell cycle analysis.

We then analysed the cell cycle state of the cells to see if this was consistent with the hypothesis that Group1 and Group2 were subsets of quiescent NSCs and Group3 was a cluster of neuroblasts. Two groups have independently and directly analysed the cell cycle state of FAC sorted stem cells of the V-SVZ (Leeman *et al.,* 2018; Morizur *et al.,* 2018). Using either fluorescent dyes or DNA analogues, they found that quiescent NSCs are arrested in the G0/G1 phase and that activated NSCs and neuroblasts are in the S/G2/M phase. Thus, if Group1 and Group2 were indeed quiescent NSCs, one would expect them to be largely arrested in the G0/G1 phase of the cell cycle. Likewise, if Group3 was a cluster of neuroblasts they should largely be in the S/G2/M phase. We performed this type of analysis bioinformatically by assigning each cell a cell cycle phase based on their expression of well-annotated cell cycle genes (Figure 3.25a). While this is not a direct analysis of the cell cycle state of each of the groups, it does indicate that Group1 and Group2 were largely in the G1 phase, while Group3 was largely in the G2/M phase. We found similar results when we performed this analysis on the data from StriatumSeq1 (Figure 3.25b). Group15 cells were largely in the G1 phase, while Group26 is more variable—cells are in a mixture of G2/M and G1 phase.

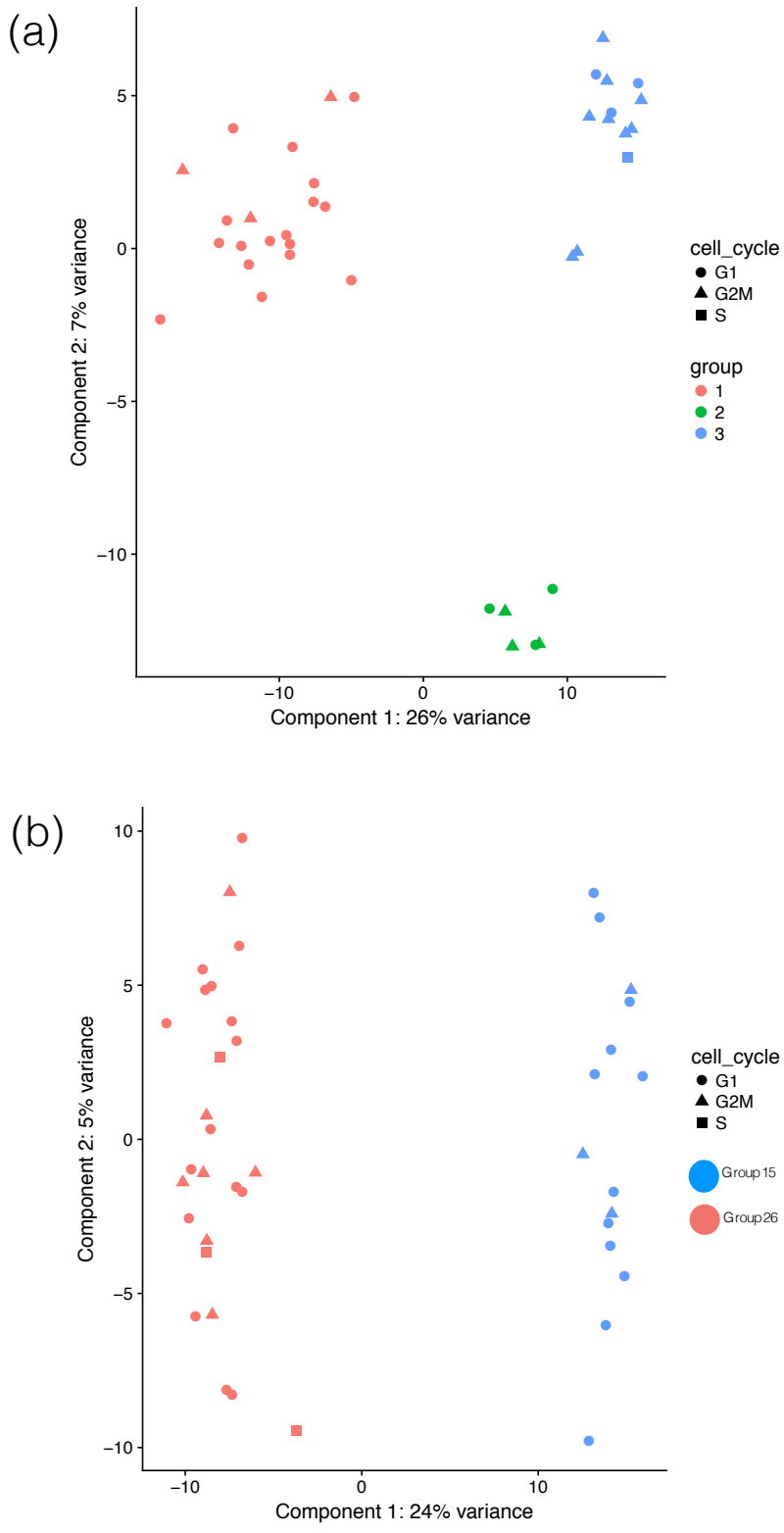


Figure 3.25. Cell cycle analysis of the sequenced cells. The cells from StriatumSeq2 **(a)** and StriatumSeq1 **(b)** were given an estimate of their cell cycle phase based on their expression of annotated cell cycle genes. The classifications are projected onto each cell in the respective PCA plots.

3.17 Analysing gene expression of the three clusters in StriatumSeq2.

The Pseudotime and cell cycle analyses indicated that Group1 was likely a cluster of quiescent NSCs, Group3 was a cluster of neuroblasts, and that Group2 was some intermediate group of cells—possibly activated NSCs. We began by producing a heatmap of the gene expression for the most differentially expressed genes in each group (Figure 3.26). This provided a global picture of gene expression and showed that the clusters did have unique patterns of gene expression. We proceeded then to perform specific gene expression analysis to confirm this by using the most differentially expressed, and therefore upregulated genes, for each of the groups—as determined by differential gene expression analysis using SCDE. We began by analysing gene markers for Group1 cells. Figure 3.27a shows the expression of *Kcnj13* (Kir7.1) in StriatumSeq2 cells. Like in Striatumseq1 (Group15), *Kcnj13* is intensely and specifically expressed in Group1 cells ($p_{adjusted1v2}$ = 5.08E-10, $p_{adjusted1v3}$ = 2.39E-10). This is line with the finding, in neurospheres, that KiR7.1 is expressed by "early" stem cells and not by "late" stem cells (Pruss *et al.,* 2011). As in StriatumSeq1, we also found that *Prlr* was highly and specifically expressed in Group1 (Figure 3.27b; $p_{adjusted1v2}$ = 6.56E-06, $p_{adjusted1v3}$ = 2.46E-10). This expression overlapped completely with the expression observed in Group15 when both experiments were combined (data not shown). In StriatumSeq2 we also found that Folate Receptor 1 (*Folr1*) gene was highly and differentially expressed in Group1 (Figure 3.27c; ($p_{adjusted1v2}$ = 4.07E-07, $p_{adjusted1v3}$ = 2.39E-10). Thus, these three markers (*Kcnj13, Prlr, Folr1*) are highly expressed in Group1 (putatively, quiescent NSCs) and could be used *in vivo* to identify these cell types for functional analyses. To complete the analysis for Group1 we also analysed the two markers we used to determine that Group15 cells were quiescent NSCs in StriatumSeq1. We found that *Vcam1* was specifically upregulated in Group1 ($p_{adjusted1v2}$ = 0.006, $p_{adjusted1v3}$ = 2.68E-07). *Igfbp2* was similarly upregulated in Group1 ($p_{adjusted1v2}$ = 0.03, $p_{adjusted1v3}$ = 0.0004). Secreted Frizzled Related Protein 1 (*Sfrp1*) is a recently discovered marker of neural progenitor cells both in the V-SVZ and the prefrontal cortex (Zhong

et al., 2018; Mizrak *et al.,* 2019). We found this progenitor marker to be highly upregulated in Group1 ($p_{adjusted1v2}$ = 0.03, $p_{adjusted1v3}$ = 0.002). Finally, functional and immunochemical analyses by the Alvarez-Buylla group found that the B-cells of the V-SVZ receive long-distance 5-HT_{2C} projections from the Raphe nuclei. With the further aid of patch-clamp recordings, they concluded that B-cells must express the 5-HT_{2C} receptor (*Htr2c*; Tong *et al.,* 2014). We found that this key gene is specifically upregulated in the Group1 cluster ($p_{adjusted1v2}$ = 0.05, $p_{adjusted1v3}$ = 0.0002). This is the first demonstration of 5-HT_{2C} receptor gene expression in cells of the V-SVZ in an RNA sequencing experiment and further evidence that Group1 cells (and Group15 cells from StriatumSeq1) are, indeed, the B-cells of the V-SVZ. Therefore, by extension, our analysis identifies novel and specific gene and protein markers for B-cells (*Kcnj13, Prlr, Folr*). This is of acute importance as, currently, no specific markers of B-cells exist—much less transporters and receptors that can be modulated extracellularly. Group1 markers are summarised in Table 3.2.

We also then looked for marker genes of Group2, to better understand the identity of this group as it was unique to StriatumSeq2. We found that Group2 specifically and highly expressed the gene *Slc1a3*, otherwise known as the Glutamate Aspartate Transporter (GLAST; $p_{adjusted2v1}$ = 1.96E-07, $p_{adjusted2v3}$ = 0.1; Figure 3.28 a). GLAST is a well-established marker for the quiescent neural stem cells of the V-SVZ (Kriegstein and Alvarez-Buylla, 2009). Recent analysis has found that the gap junction proteins Connexin43 (*Gja1*) and Connexin30 (*Gjb6*) play key roles in adult hippocampal neurogenesis and *Gjb6* has been demonstrated to be a marker of quiescent NSCs (Leibmann *et al.,* 2013; Llorens-Bobadilla *et al.,* 2015). Interestingly we find that both Connexin43 ($p_{adjusted2v1}$ = 6.23E-10, $p_{adjusted2v3}$ = 8.09E-06) and Connexin30 ($p_{adjusted2v1}$ = 5.08E-10, $p_{adjusted2v3}$ = 1.78E-07) are highly and specifically expressed in Group2 (Figure 3.28 b and c). This led us to conclude that Group2 cells may be a subset of quiescent NSCs, *i.e.* B-cells, with Group1 representing another subset—as it precedes Group2 in our Pseudotime analysis. Thus, it seems

as though our analysis identifies two subgroups of B-cells within the V-SVZ. Group2 markers are summarised in Table 3.2.

We completed this by analysing Group3 cells for their expression of neuroblasts markers and compared it to the expression from the Group26 cluster (from StriatumSeq1). We found that the neuroblast marker *Sox11* was, once again, specifically upregulated in Group3 ($p_{adjusted3v1}$ = 2.39E-10, $p_{adjusted3v2}$ = 1.64E-06; Figure 3.29a). We also found the neuroblast marker *CD24a* to be similarly upregulated in Group3 ($p_{adjusted3v1}$ = 3.41E-05, $p_{adjusted3v2}$ = 0.02; Figure 3.29b). The α-tubulin subunit Tuba1a has been shown to play a key role in hippocampal neurogenesis (Keays *et al.,* 2010). In our analysis, we find it to be specifically and highly expressed in Group3 ($p_{adjusted3v1}$ = 0.001, $p_{adjusted3v2}$ = 1.37E-05; Figure 3.29c). Doublecortin (*Dcx*), another canonical marker of migrating neuroblasts, was found to be significantly upregulated in Group3 ($p_{adjusted3v1}$ = 2.33E-05, $p_{adjusted3v2}$ = 0.01). We found that *βIII-tubulin* was also significantly upregulated in Group3 ($p_{adjusted3v1}$ = 3.56E-05, $p_{adjusted3v2}$ = 0.009). Finally, the transcription factor Dlx5 has been shown to play an important role in neurogenesis; its absence leads to defects in the neurogenic process (Perera *et al.,* 2004). We find that *Dlx5* is significantly upregulated in Group3, and virtually absent in Group1 and Group2 ($p_{adjusted3v1}$ = 0.0001, $p_{adjusted3v2}$ = 0.02). This expression profile provides substantial support for the hypothesis that the Group3 cluster is comprised of cells that are migrating neuroblasts. Group3 markers are summarised in Table 3.2.

In summary, the gene expression analysis led to the following conclusions: (1) Group1 is a previously unidentified subgroup of B-cells; (2) Group2 is comprised of B-cells that are either more differentiated than Group1 or just a different subgroup of B-cells; (3) Group3 represents migrating neuroblasts. Thus, our Penk-tdTomato^{+} labels V-SVZ stem cells at three stages of differentiation.

Table 3.2

Gene Markers		
Group1	**Group2**	**Group3**
Kcnj13	*Slc1a3*	*Sox11*
Prlr	*Gja1*	*CD24a*
Folr1	*Gjb6*	*Tuba1a*
Igfbp2		*Dcx*
Sfrp1		*ßIII-tubulin*
Htr2c		*Dlx5*

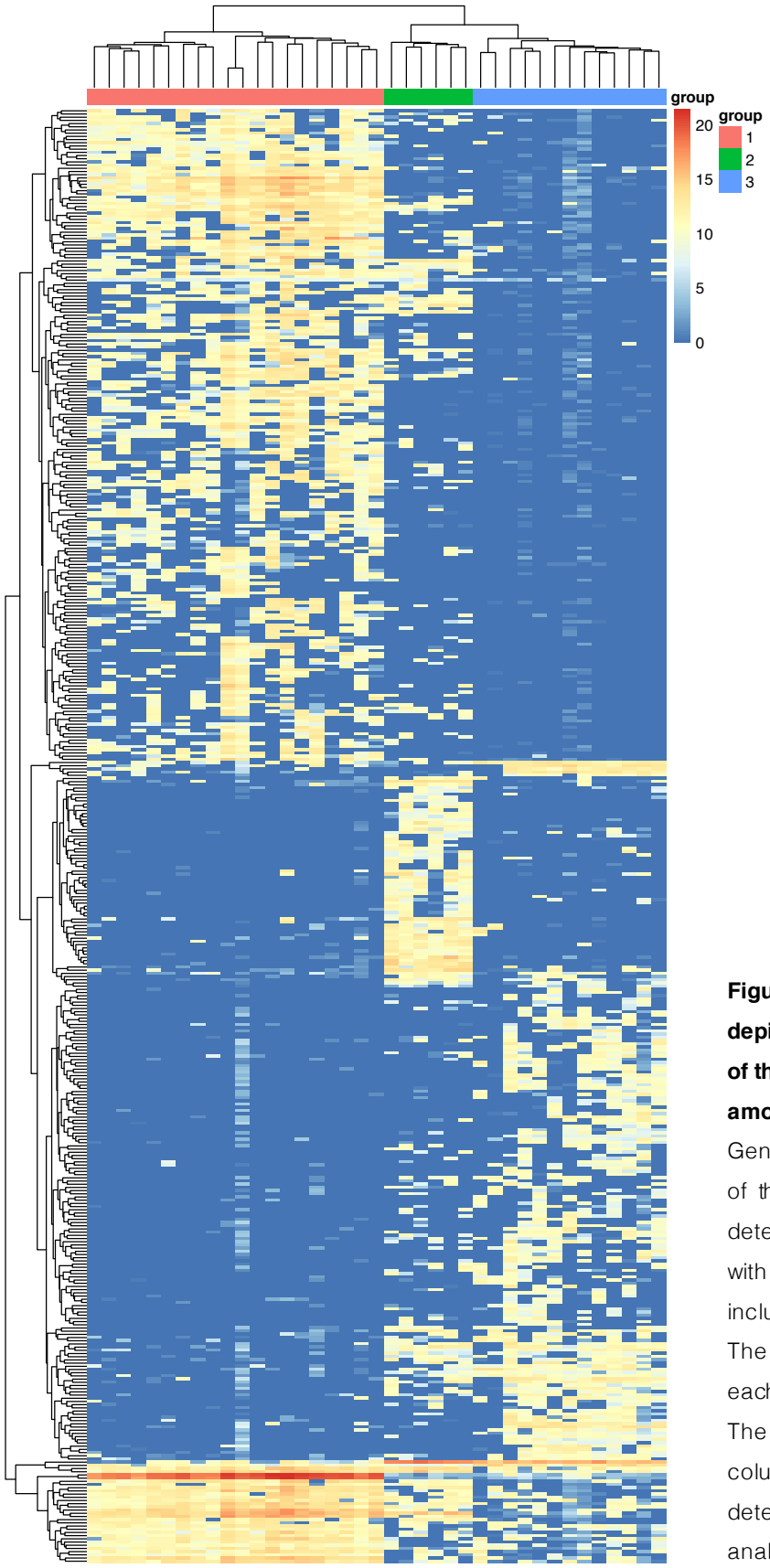


Figure 3.26. A heatmap depicting the expression levels of the most variable genes amongst the 39 cells.

Genes were selected on the basis of their differential expression as determined by SCDE. Only genes with $p_{adjusted} < 0.001$ were included, each presented in rows. The relative expression level of each gene is indicated by colour. The cells are ordered into columns. The group identity (as determined by previous clustering analysis) is indicated by colour.

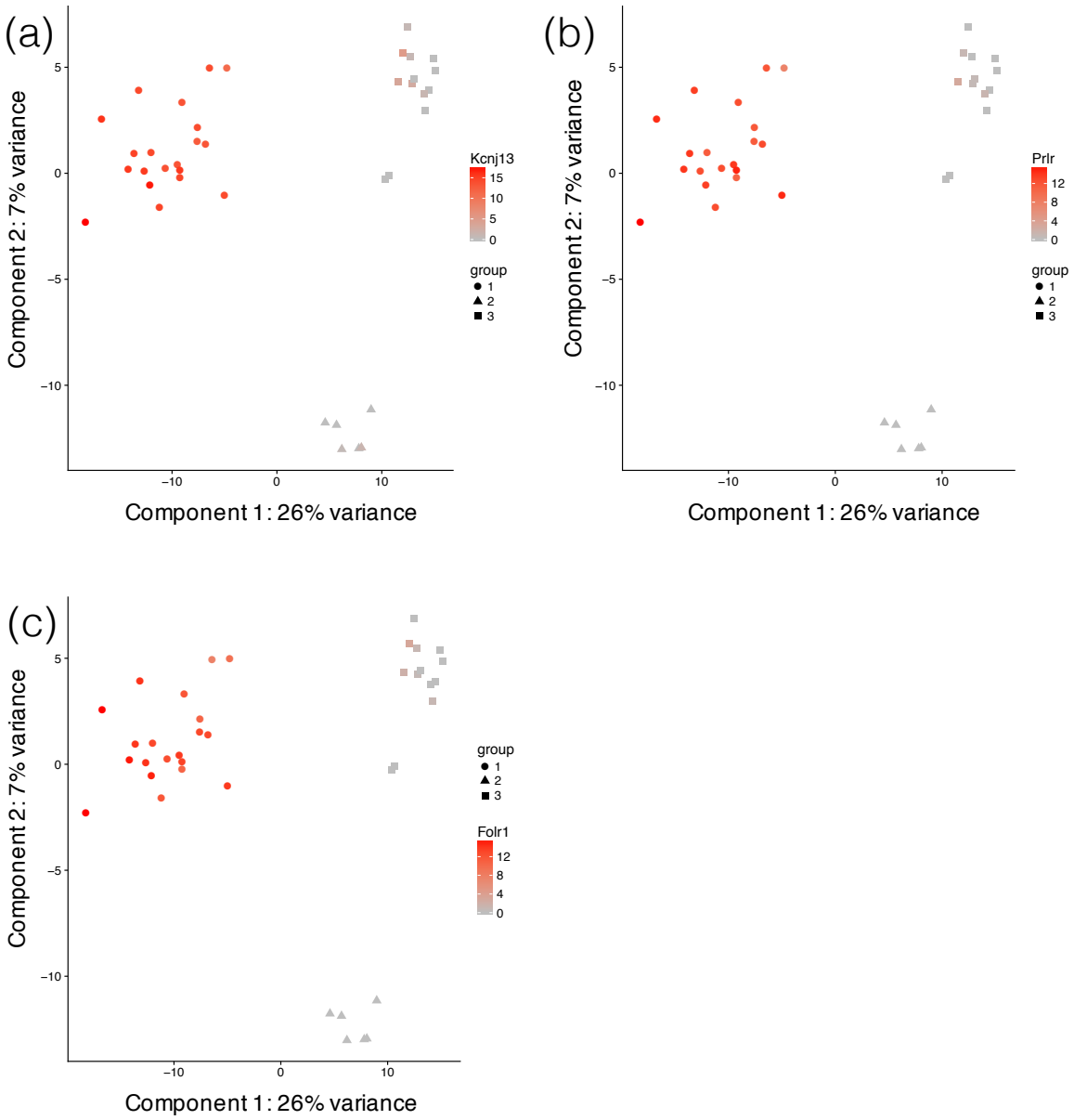


Figure 3.27. The selective expression of quiescent NSCs markers in Group1. The PCA plot is generated using the 500 most variable genes and the expression level for the gene of interest in each cell is projected. **(a)** Depicts the expression of *Kcnj13* in each of the sequenced cells. SCDE analysis finds it is significantly upregulated in Group1 ($p_{adjusted1v2}$ = 5.08E-10, $p_{adjusted1v3}$ = 2.39E-10; average expression: Group1 = 9660.4, Group2 = 0.272, Group3 = 2.142). **(b)** Depicts the expression of *Prlr* in each of the sequenced cells. It is significantly upregulated in Group1 ($p_{adjusted1v2}$ = 6.56E-06, $p_{adjusted1v3}$ = 2.46E-10; average expression: Group1 = 3394.8, Group2 = 0, Group3 = 0.502). **(c)** Depicts the expression of *Folr1* in each of the sequenced cells. Its expression is virtually restricted to Group1, and this difference is highly significant ($p_{adjusted1v2}$ = 4.07E-07, $p_{adjusted1v3}$ = 2.39E-10; average expression: Group1 = 3976.9, Group2 = 0, Group3 = 0.532).

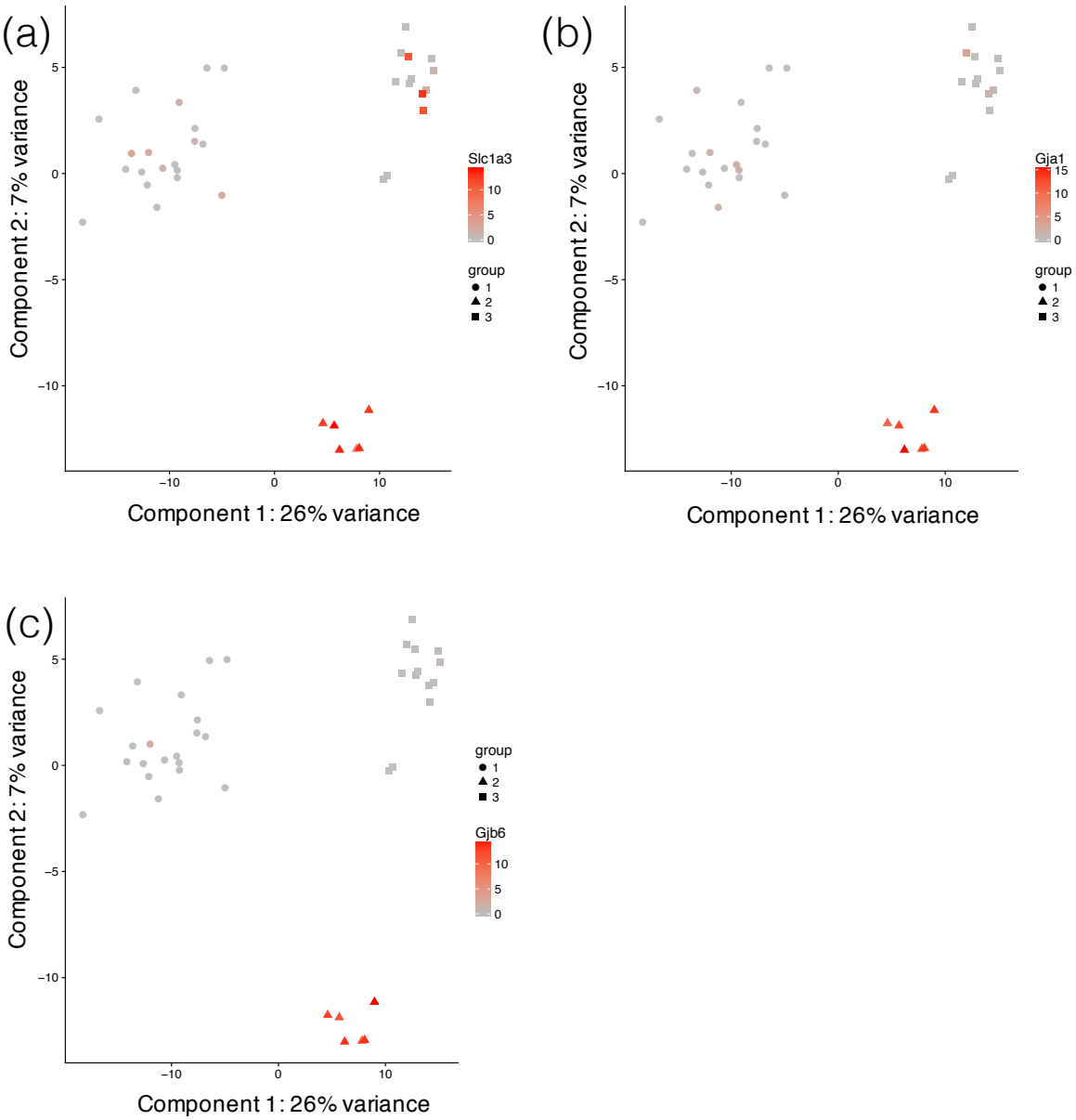


Figure 3.28. The selective expression of quiescent NSCs markers in Group2. The PCA plot is generated using the 500 most variable genes and the expression level for the gene of interest in each cell is projected. **(a)** Depicts the expression of *Slc1a3/GLAST* in each of the sequenced cells. SCDE analysis finds it is significantly upregulated in Group2 ($p_{adjusted2v1}$ = 1.96E-07, $p_{adjusted2v3}$ = 0.1; average expression: Group1 = 0.387, Group2 = 7256.8, Group3 = 506.4). **(b)** Depicts the expression of *Gja1/Connexin43* in each of the sequenced cells. It is significantly upregulated in Group2 ($p_{adjusted2v1}$ = 6.23E-10, $p_{adjusted2v3}$ = 8.09E-06; average expression: Group1 = 0.188, Group2 = 11687.0, Group3 = 0.436). **(c)** Depicts the expression of *Gjb6/Connexin30* in each of the sequenced cells. Its expression is virtually restricted to Group2, and this difference is significant ($p_{adjusted2v1}$ = 5.08E-10, $p_{adjusted2v3}$ = 1.78E-07; average expression: Group1 = 0.105, Group2 = 7983.6, Group3 = 0.021).

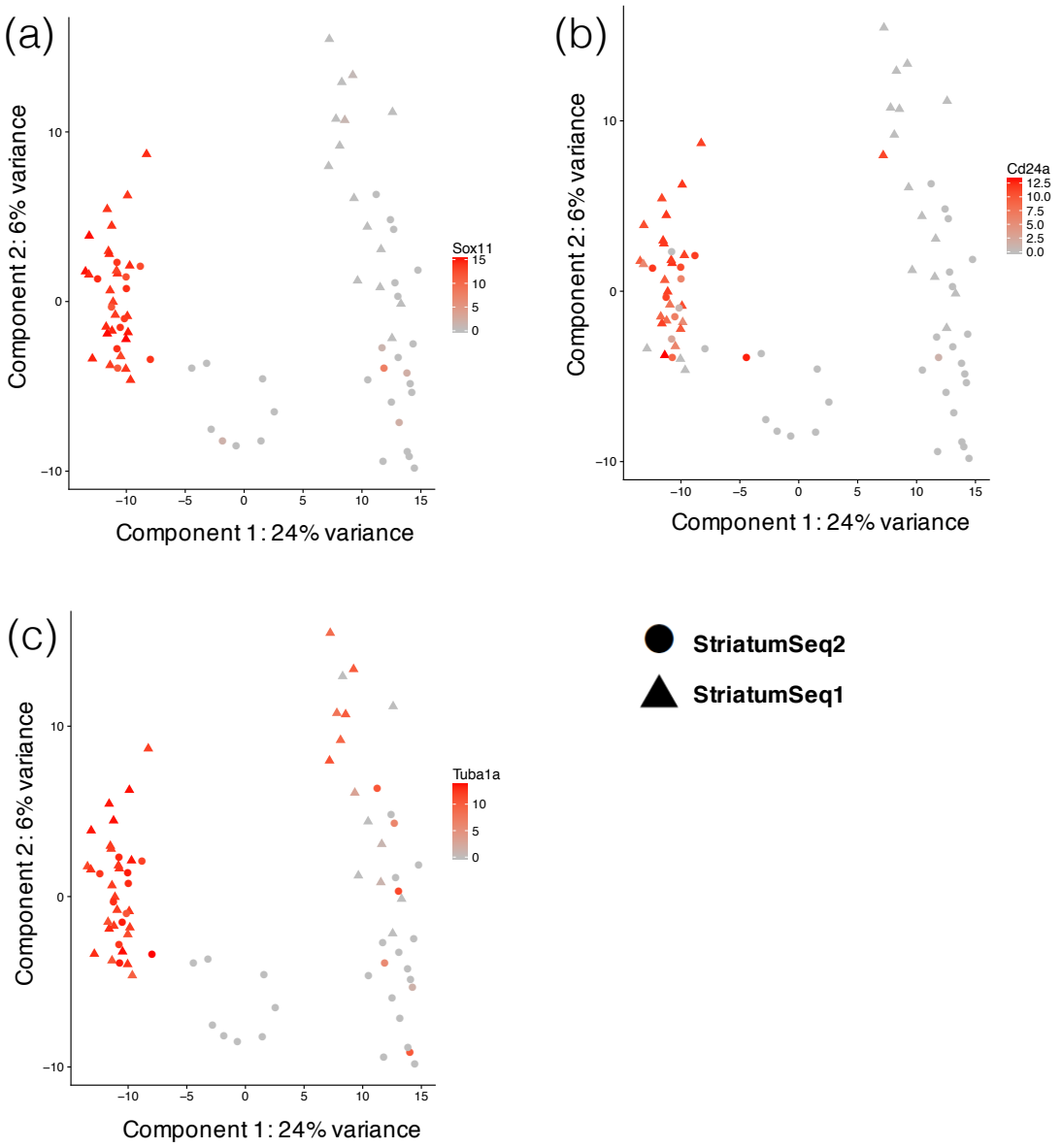


Figure 3.29. The selective expression of neuroblast markers in Group3/Group26 of both experiments. The PCA plot is generated using the 500 most variable genes and the expression level for the gene of interest in each cell is projected. The data from both experiments are combined to show the conserved and restricted expression of the neuroblast markers **(a)** Depicts the expression of *Sox11* in each of the sequenced cells. SCDE analysis finds it significantly upregulated in Group3/Group26 ($p_{adjusted3v1}$ = 2.39E-10, $p_{adjusted3v2}$ = 1.64E-06; average expression: Group1 = 3.875, Group2 = 0.204, Group3 = 5975.5). **(b)** Depicts the expression of *CD24a* in each of the sequenced cells. It is significantly upregulated in Group3/Group26 ($p_{adjusted3v1}$ = 3.41E-05, $p_{adjusted3v2}$ = 0.02; average expression: Group1 = 0.027, Group2 = 0, Group3 = 627.3). **(c)** Depicts the expression of *Tuba1a* in each of the sequenced cells. Its expression is virtually restricted to Group3/Group26, and this difference is significant ($p_{adjusted3v1}$ = 0.001, $p_{adjusted3v2}$ = 1.37E-05; average expression: Group1 = 134.2, Group2 = 0, Group3 = 3303.7).

3.18 Gene ontology analysis on the three clusters.

We next performed gene ontology (GO) enrichment analysis on the sets of differentially expressed genes to determine possible functional consequences for the up- or down-regulation of identified genes. This was performed using the GOsummaries package (Kolde and Vilo, 2015). We compared the clusters against each other to functionally determine the possible identities of the cells in the three clusters (Figure 3.30). We first compared Group1 to Group2 (Figure 3.30a). We found that 349 genes were upregulated and 171 genes were downregulated in Group1 when compared to Group2. The wordcloud indicates that the processes relatively upregulated in Group1 are, therefore, constituents of ribosomes and lysosomes. The upregulation of both of these pathways is expected in B-cells. Electron micrographs performed by the Alvarez-Buylla group found that a subgroup of B-cells have high densities of ribosomes (Obernier *et al.,* 2018). A recent analysis of neural stem cell activation by the Brunet group found that quiescent NSCs contained large lysosomes (Leeman *et al.,* 2018). Meanwhile, the processes relatively upregulated in Group2 are related to cellular homeostasis—as would be expected of a subgroup of quiescent NSCs. We then compared Group1 to Group3 (Figure 3.30b). We found that 504 genes are upregulated, while 296 genes are downregulated in Group1 when compared to Group3. Processes related to endoplasmic reticulum function are relatively upregulated in Group1. The Leeman *et al.* (2018) analysis also found endoplasmic reticulum genes to be highly expressed in quiescent NSCs compared to more active NSCs. We found that processes related to neuronal development are upregulated in Group3—providing further evidence that Group3 cells are neuroblasts. We finally compared Group2 to Group3 (Figure 3.30c). 127 genes are upregulated and 193 genes are downregulated in Group2 when compared to Group3. As seen before, we found that cellular homeostasis processes are upregulated in Group2 and neuronal development genes are upregulated in Group3. This analysis provided further evidence for the hypothesis laid

out in the previous section—that Group1 is a novel subset of B-cells, Group2 is a previously identified subset of B-cells, and that Group3 represents a cluster of neuroblasts.

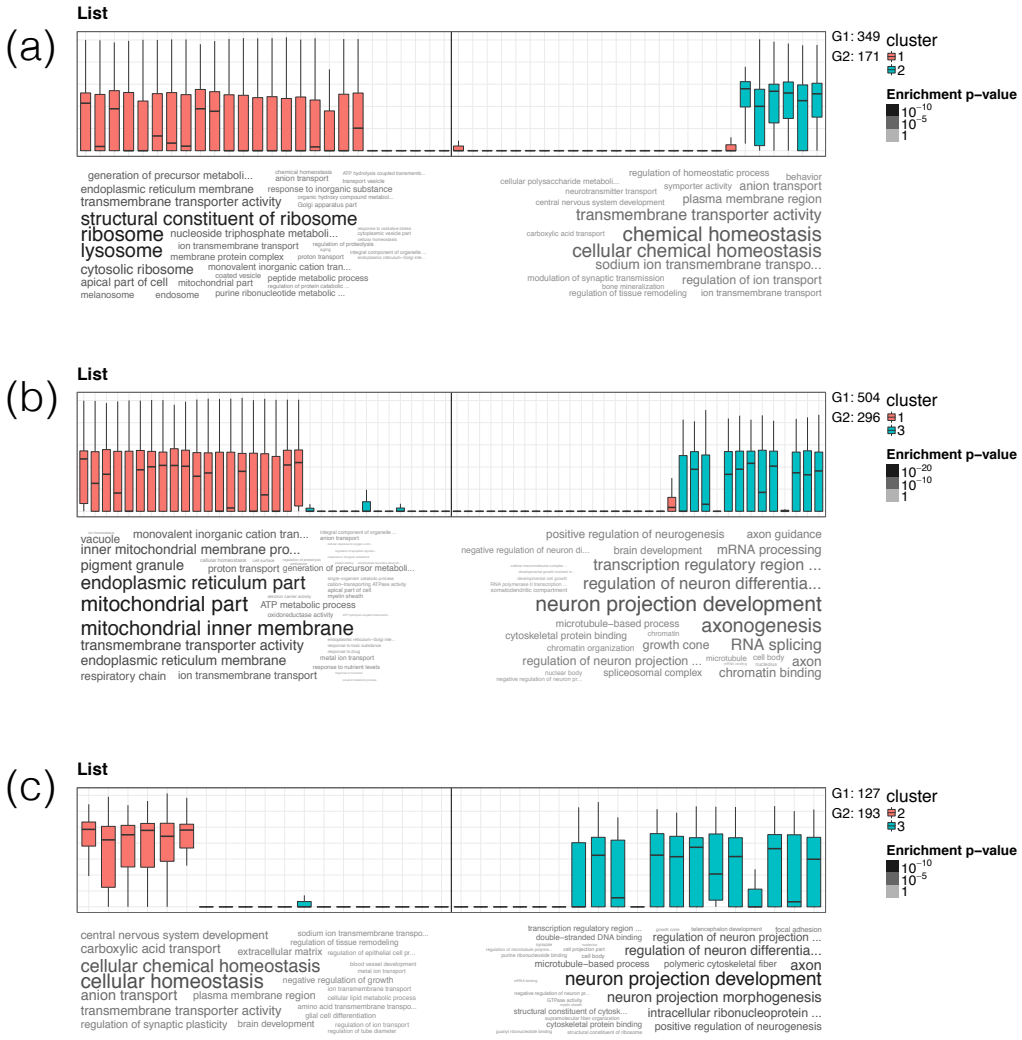


Figure 3.30. Gene ontology analysis reveals specific, stereotyped processes to be upregulated in each of the three clusters. Gene ontology analysis was performed using differentially expressed genes ($p_{adjusted} < 0.05$). G1 = upregulated, G2 = downregulated genes in the cluster in the left when said cluster (on the left) is compared to the cluster on the right. The size and darkness (see legend on the right of each panel) of the words are used to indicate the strength of enrichment of processes in each of the clusters. Each box in the box plot represents one cell from that cluster, the identity of the cluster is indicated in the legend on the right. Plotted are the average log-expression values of each cell for the genes that make up the set of processes enriched in the word cloud below. **(a)** Comparison between Group1 and Group2. **(b)** Comparison between Group1 and Group3. **(c)** Comparison between Group2 and Group3.

3.19 Comparing the StriatumSeq datasets to a published V-SVZ NSC single cell sequencing dataset.

We then compared our StriatumSeq datasets to published reports of sequenced V-SVZ NSCs to both validate our findings and to determine whether Group1 was indeed a novel subgroup of B-cells, as we had hypothesised. The Martin-Villalba group used antibodies against Prom1 (CD133) in combination with GLAST to isolate and sequence NSCs of the V-SVZ (Llorens-Bobadilla *et al.,* 2015).

A key difference between this study and ours is that we rely on a more reliable genetically expressed marker and not on the binding of antibodies. Furthermore, our Pseudotime analysis indicated that GLAST is highly enriched in NSCs that are slightly more advanced pseudotemporally (*i.e.* Group2) than the Group1 cluster we have identified. Thus, by its very design, the Llorens-Bobadilla *et al.,* study excludes the novel subgroup of B-cells we have identified. When we analysed their raw dataset for the expression of *Kcnj13* (a strong marker for Group1) we found that it was expressed at very low levels in only 3 cells across their whole dataset. Similarly, *Prlr* is expressed at barely detectable levels in their dataset. Finally, the 5-HT_{2C} receptor gene is not detected at all in their sequencing experiment. This manual analysis indicated that our Group1 cluster was indeed unique.

We performed a deeper analysis by re-analysing the heatmap of gene expression we previously generated against the Llorens-Bobadilla *et al.* (2015) gene set. In the Llorens-Bobadilla *et al.* (2015) paper, they identify classes of genes which they state defines cells into cell types. When we generated a heatmap of our cells using their gene classes, we found that Group1 and Group2 cells largely fall in their “Class II and III”—which define quiescent NSCs. Whereas, Group3 falls in “Class V” and “Class VII”—which define neuroblasts (Figure 3.31a). When we added the cells from StriatumSeq1 into this analysis, we found that the same pattern is recreated (Figure 3.31b). Thus, although the markers for Group1 cells are unique and novel, comparison to the

Llorens-Bobadilla *et al.* (2015) does support our general hypothesis for the cell dynamics between the clusters.

We then performed a deeper, automated comparison analysis against a more recently published report of sequenced V-SVZ NSCs to determine whether Group1 was a novel subgroup of B-cells. The Brunet group used the genetically-encoded reporter GFAP-GFP along with antibodies against Prom1 and EGFR to isolate and sequence quiescent and activated NSCs from the V-SVZ (Dulken *et al.,* 2017). We wanted to perform a direct comparison between our single cell RNAseq dataset and the dataset they had produced. To achieve this, we simply used their raw data and did exactly the same downstream analysis we performed on our own dataset. Following the analysis pipeline, we plotted the cells from StriatumSeq2 and the Dulken *et al.* dataset onto a PCA plot (Figure 3.32a). We found that our Group1 cells (a putative novel subgroup of B-cells) does not overlap with any of the cells they identify as quiescent NSCs, and is therefore unique. Interestingly, we find that it is Group2 (which we find to be slightly more differentiated than Group1, as per Pseudotime analysis) that overlaps with the Dulken *et al.* quiescent NSC group. Finally, Group3 shows some overlap between their NPC (neural progenitor cell) and activated NSC clusters—indicating that they are indeed terminally differentiating neuroblasts as we hypothesised. We then plotted all the cells from StriatumSeq1 and 2 along with the Dulken *et al.* dataset (Figure 3.32b). This confirmed the results from the first PCA plot. We found that Group15 (from StriatumSeq1) and Group1 (from StriatumSeq2) remain unique in this plot. Group26 (from StriatumSeq1) cells overlap with the NPC population. We also performed Pseudotime analysis on the combined dataset (Figure 3.32c). As expected, we found that Group1 and Group15 precede the Dulken *et al.* (2017) quiescent NSCs on this pseudotemporal differentiation pathway. Thus, we concluded that Group1 and Group15 cells represented a novel group of cells resident within the V-SVZ niche, and because they expressed markers of quiescent stem cells and

predicted markers for B-cells, we further concluded that they represent a novel subgroup of B-cells.

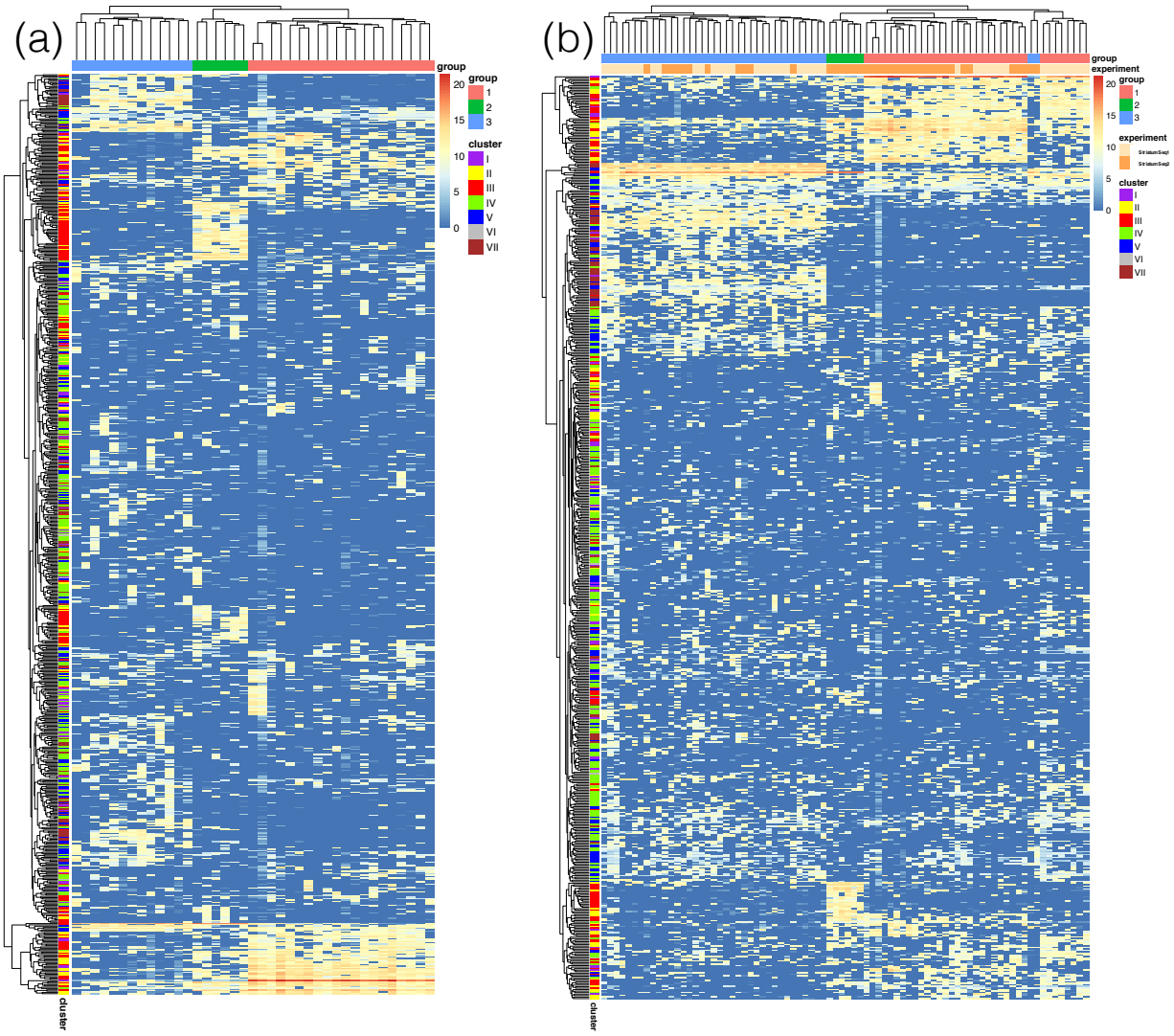


Figure 3.31. A heatmap comparing gene expression in StriatumSeq2 and StriatumSeq1 to a previously published report of sequenced V-SVZ NSCs. The classes identified by Llorens-Bobadilla *et al.* (2015) are listed on the right and coloured. Classes II-III represent quiescent NSC genes. Classes V-VII represent neuroblast genes. The genes are presented in rows according to the classes to which they belong. Cells from **(a)** StriatumSeq1 or from **(b)** both StriatumSeq1 and StriatumSeq2 are listed in columns and coloured according to cluster and experiment.

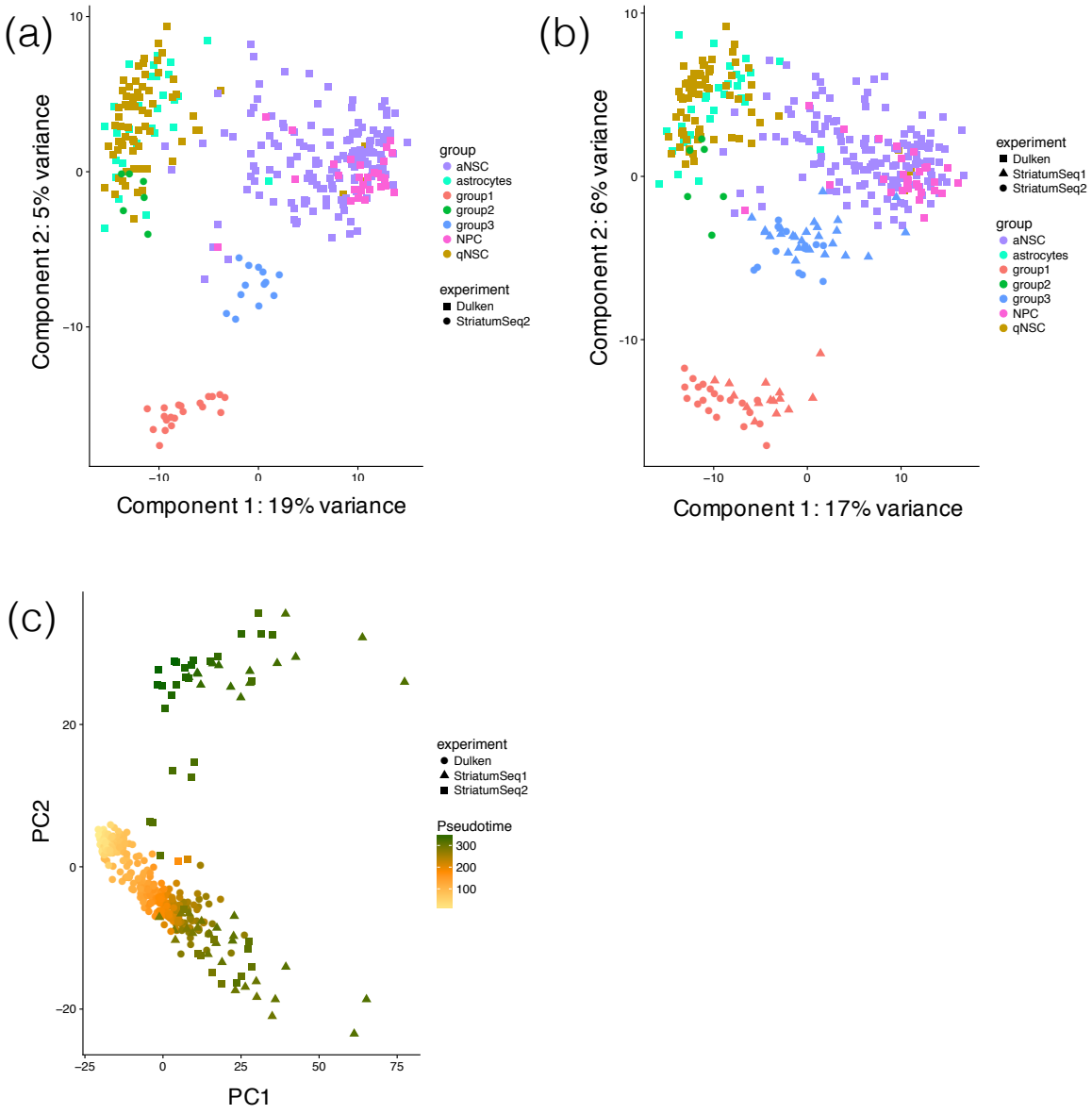


Figure 3.32. Direct comparison of StriatumSeq1 and 2 to a recently published report reveals that the Group1/Group15 cluster represent a novel group of cells. The raw data from Dulken *et al.* (2017) was computationally processed as for StriatumSeq1 and 2. The PCA plots are generated using the 500 most variable genes resulting from this analysis. The assignations "qNSC, aNSC, NPC, astrocytes" are from Dulken *et al.* and are used to label their dataset. The assignations "Group1/2/3" are used to label cells from StriatumSeq1 and 2. **(a)** The cells from StriatumSeq2 and Dulken *et al.* are plotted, revealing Group1 to be a non-overlapping cluster. **(b)** The cells from StriatumSeq2, StriatumSeq1, and Dulken *et al.* are plotted, revealing Group1/Group15 to be a non-overlapping cluster. **(c)** Pseudotime analysis was performed on the cells from StriatumSeq2, StriatumSeq1, and Dulken *et al.* A group of cells corresponding to Group1/Group15 (dark green, on the top on the plot) lies at the pseudotemporal extreme, compared to all other cells.

3.20 Enkephalin gene expression in sequenced Penk-tdTomato$^+$ cells.

The cells sequenced for both experiments were isolated from IPNK mice, which are engineered to permanently label (with tdTomato) any cells expressing Enkephalin at any point in the developmental arc of a cell. The line was developed in the laboratory and extensively characterised and analysed by Dr. Dario Besusso and a previous D.Phil. student in the laboratory, Chinnavuth Vatanashevanopakorn. It was found that approximately 95% of cells, in IPNK brains, that expressed Enkephalin (as assessed by their immunoreactivity to the Penk antibody) were positive for the reporter—meaning that the reporter is a faithful proxy for Enkephalin expression in the mouse brain (Besusso *et al.,* 2013; Vatanashevanopakorn *et al.,* 2017). Thus, in our RNAseq data set we wanted to determine the expression level of the *Penk* gene in the Penk-tdTomato$^+$ cells isolated from the V-SVZ. To our surprise, we found that *Penk* expression was very minimally detected in the identified groups from both experiments—therefore, these cells may not express detectable levels of Penk protein, even though the *Penk* promoter must have been active at some point during their lifetime in order to activate the reporter gene. So, we performed an immunofluorescence experiment using IPNK brain sections to assess whether the V-SVZ Penk-tdTomato$^+$ cells expressed Penk protein (Figure 3.33). In line with the RNA sequencing data, we found that the Penk-tdTomato$^+$ cells of the V-SVZ do not express Penk. However, as expected, the Penk-tdTomato$^+$ cells of the striatal parenchyma do express Penk protein. This suggests that a subset of NSCs of the V-SVZ are defined by their expression of Penk early in their development (and not at the time of sequencing), and that this subset of NSCs represent the novel subgroup of B-cells.

In summary, in this chapter we showed that the IPNK mouse line leads to the labeling of three subgroups of cells at the V-SVZ germinal niche. These subgroups comprise stem cells at three separate stages of their differentiation arc. Furthermore, we show that one of these

groups—Group1—represents a putative novel subgroup of quiescent NSCs or B-cells. In the next chapter, we show how we validated these findings *in vivo.*

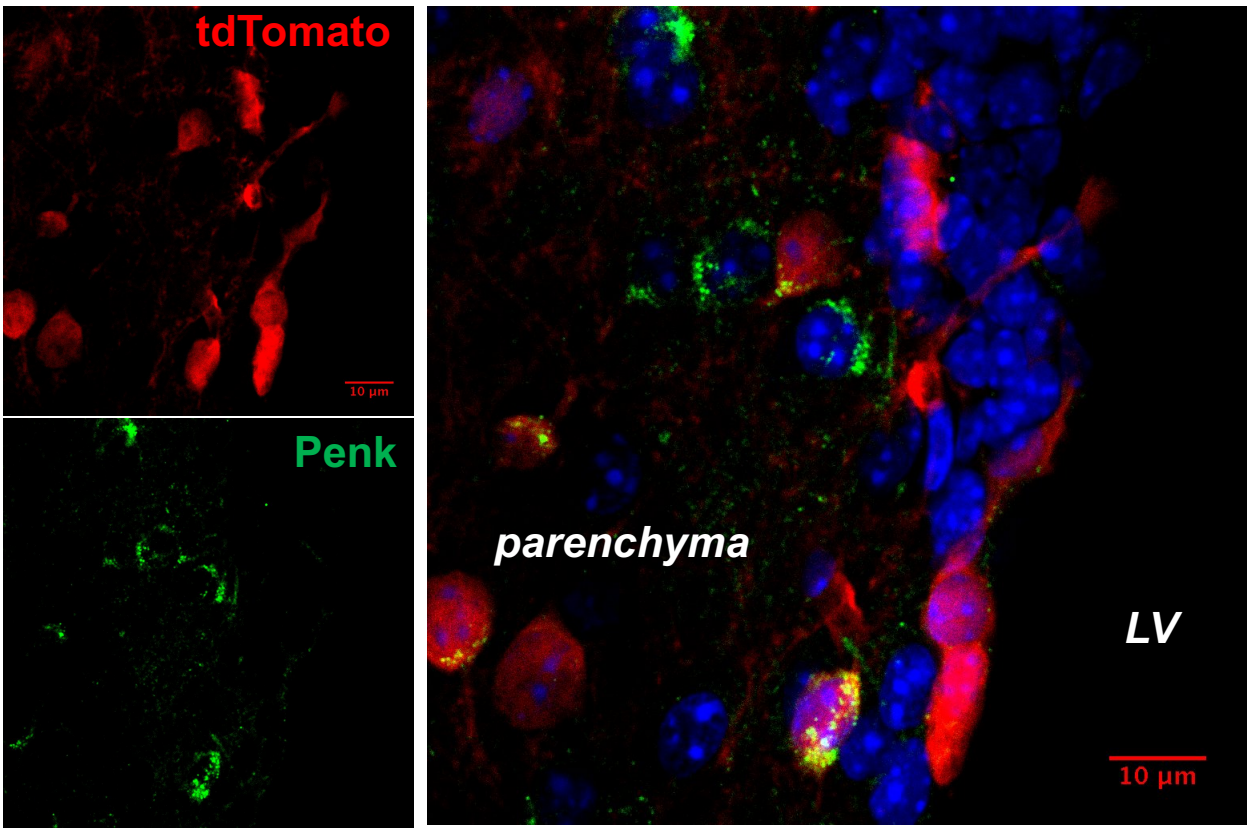


Figure 3.33 Penk-tdTomato⁺ cells residing in the V-SVZ stem cell niche do not express Penk protein. Representative IPNK mouse brain section (female; 2 months old) immunostained using an antibody against Penk and visualised using an upright confocal microscope. The V-SVZ is demarcated by the thin, dense, and continuous line of DAPI-stained nuclei. Penk-tdTomato⁺ cells are seen in the V-SVZ and the striatal parenchyma. However, staining for Penk protein is only found in the striatal parenchyma and not in the cells of the V-SVZ.

Chapter 4 – Results II: The *in vivo* validation of the bioinformatics data.

4.1 An immunocytochemical analysis of Penk-tdTomato$^+$ cell types in the V-SVZ germinal niche.

The bioinformatics data indicated that the clusters seen in StriatumSeq1 and StriatumSeq2 were NSCs of the V-SVZ at different stages of differentiation. This finding needed to be validated *in vivo*. Thus, we began by first performing immunochemical analyses to determine the number and distribution of Penk-tdTomato$^+$ cells within the V-SVZ germinal niche. As we wanted to analyse the distribution of the cells throughout the whole of the lateral ventricle, we systemically selected six 30 μm sections spanning the entire LV (repeated in three different 2/3-month-old IPNK brains). The number of Penk-tdTomato$^+$ cells were counted from the lateral walls of the LVs from both hemispheres. We found that the highest numbers of tdTomato$^+$ cells was near bregma. To estimate the number of Penk-tdTomato$^+$ cells throughout the LV, we used a simple calculation (see ***Materials and Methods***) to determine that there were approximately 15,000 Penk-tdTomato$^+$ cells lining the lateral wall of the LV (Table 4.1).

We then performed a series of colocalisation analyses to determine the cell type composition of Penk-tdTomato$^+$ cells in the V-SVZ. First, we determined whether the Penk-tdTomato$^+$ cells displayed any ependymal cell signatures. This was important, because ependymal cells are an important part of the V-SVZ stem cell niche but are not neural stem cells. In our sequencing data (for both experiments) we did not detect any expression of *CD133* (*Prom1*) in any of the groups. CD133, although used by some groups to isolate NSCs, is a strong marker for ependymal cells (Llorens-Bobadilla *et al.,* 2015; Chaker *et al.,* 2016; Dulken *et al.,* 2017). A recent study of the ependymal cells of the V-SVZ found that they are defined by the expression

of *Foxj1*—a gene involved in ciliogenesis (Shah *et al.,* 2018). In line with the hypothesis that Penk-tdTomato$^+$ cells are not ependymal cells, we did not detect the expression of *Foxj1* in any of the groups in either experiment. Thus, the bioinformatics indicated that the Penk-tdTomato$^+$ cells were not ependymal in nature. We wanted to confirm this using immunocytochemistry on IPNK brain sections. The analysis was performed in a similar way to the analysis outlined above, except that here we only analysed the number of Penk-tdTomato$^+$ cells that were also CD133$^+$ (Figure 4.1 and Figure A1). Across the three brains, we found that the average extent of colocalisation throughout the LV is approximately 0.5%. Thus, both our bioinformatics analysis and immunocytochemical experiments indicated that Penk-tdTomato$^+$ cells at the V-SVZ are not ependymal cells.

Since their initial identification as “astrocytes”, the neural stem cells of the V-SVZ have been visually and functionally identified by their expression of the astrocyte marker GFAP (Doetsch *et al.,* 1999; Dulken *et al.,* 2017), we hypothesise that Group1, Group2, and Group15 represent quiescent NSCs and thus we would expect to detect the expression of *GFAP* in these groups. However, when we analysed the RNAseq datasets for *GFAP* expression we could not detect it. This may be explained by the following reason. To perform their RNA sequencing analysis, Dulken *et al.* used the GFAP-GFP reporter line. However, when we analysed their dataset we also could not find the expression of *GFAP* in their data—even though the sequenced cells must have expressed *GFAP* for them to be sorted and then sequenced. Prof. Brunet hypothesises that this is because *GFAP* is expressed at relatively low levels and thus the transcript copy numbers are “below the threshold for detection in single cell RNAseq” (Personal e-mail communication, 2018). This could explain why we also do not detect *GFAP* expression in our putative NSC groups, even though they express a host of other quiescent stem cell markers. In line with this, in a recent sequencing experiment of dentate gyrus NSCs performed by the Linnarsson group, where they also used a GFAP-GFP reporter line, *GFAP* expression could not

be detected in a large proportion of GFP$^+$ cells (Hochgerner *et al.,* 2018). Thus, we wanted to determine whether Penk-tdTomato$^+$ cells express GFAP protein *in vivo* and to what extent (Figure 4.1; Figure 3.2). The analysis performed was identical to that performed for CD133 immunostaining. We found that throughout the lateral wall of the lateral ventricle, the average colocalisation between Penk-tdTomato and GFAP was approximately 36%. This indicates that roughly a third of the Penk-tdTomato$^+$ cells at the lateral wall of the LV express GFAP, a marker of quiescent NSCs.

The bioinformatics data also clearly indicated that Penk-tdTomato$^+$ labels a group of cells that are likely neuroblasts. The analysis indicated that the cells in this group specifically upregulated the expression of *Sox11*. Next, we wanted to determine whether Sox11 protein colocalised with Penk-tdTomato$^+$ cells of the V-SVZ (Figure 4.1; Figure 4.9). The analysis of three IPNK brain sections indicated that approximately 22% of the Penk-tdTomato$^+$ cells express Sox11 *in vivo*. Thus, in line with the bioinformatics data, approximately a quarter of the Penk-tdTomato$^+$ cells at the V-SVZ express a marker of terminally differentiating neuroblasts.

Finally, we wanted to confirm whether some of the Penk-tdTomato$^+$ cells were actively proliferating—as the cell cycle analysis indicated (Figure 3.25). We hypothesised that the percentage of proliferating cells should be similar to Sox11-expressing cells. To label proliferating cells we used an antibody against Ki67. Ki67 is a widely-used marker of cell proliferation and its expression can be detected in all the active phases of the cell cycle (Scholzen and Gerdes, 2000). In this analysis, we found that approximately 22% of Penk-tdTomato$^+$ cells express Ki67 (Figure 4.1 and Figure A1). This is in line with the extent of Sox11 expression and further provides evidence for the hypothesis that Penk-tdTomato$^+$ also labels a subset of terminally differentiating neuroblasts in the V-SVZ.

Thus, this initial *in vivo* validation indicated that Penk-tdTomato$^+$ colocalises with markers of neural stem cells and neuroblasts at the V-SVZ, as predicted by the bioinformatics data.

Extensive work has shown that stem cells of the V-SVZ give rise to neuroblasts that travel along the rostral migratory stream to populate the olfactory bulb (Merkle *et al.,* 2007). In line with the evidence that some Penk-tdTomato^{+} cells of the V-SVZ are quiescent NSCs, we find Penk-tdTomato^{+} cells in the olfactory bulb colocalising with Dcx—indicating they are undergoing neurogenesis (*data not shown*). Upon immunostaining we found that some Penk-tdTomato^{+} cells localise with canonical olfactory bulb interneuron markers—Tyrosine Hydroxylase, Somatostatin, Parvalbumin, Calbindin, and Calretinin (*data not shown*). This indicates that Penk-tdTomato likely labels quiescent NSCs of several different subgroups, since Merkle *et al.* (2014) and Fuentealba *et al.* (2015), showed that B-cells are lineage-restricted and sub-divided during embryogenesis. Finally, we also find cells in the olfactory bulb that are Penk-tdTomato^{+} but negative for the protein Penk—as in Figure 3.33. This data is not shown, as the focus of this thesis is the analysis of the stem cells at the V-SVZ. Future work will be performed to understand the contribution of Penk-tdTomato^{+} stem cells to olfactory bulb neurogenesis and its functional consequences, therein.

Penk-tdTomato+ Cells	*Mouse 1*	*Mouse 2*	*Mouse 3*
Section 1	95	127	71
Section 2	89	99	93
Section 3	300	111	183
Section 4	81	160	198
Section 5	67	78	88
Section 6	99	85	93
Total (6 sections)	731	660	726

Table 4.1. Counts of Penk-tdTomato+ cells in the lateral wall of the lateral ventricle of three IPNK mice. Three IPNK brains were cryosectioned (30µm) and six sections chosen to span the LV. The number of Penk-tdTomato$^+$ cells were counted (also DAPI$^+$) in the lateral wall of the lateral ventricles from both hemispheres.

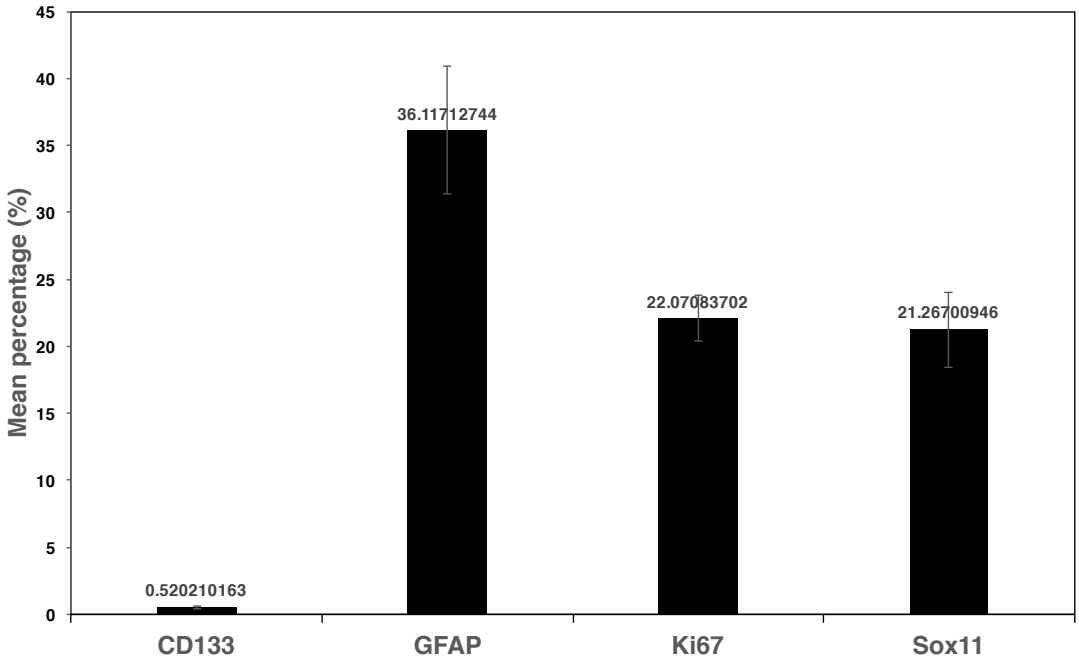


Figure 4.1. Penk-tdTomato labels at least two different cell types at the V-SVZ. The analysis was done as described in Table 4.1, except that each Penk-tdTomato$^+$ cell was assessed for the colocalisation of the stated antibodies (thus, DAPI$^+$/Penk-tdTomato$^+$/"*antibody*"$^+$). An average for the six sections of each mice was obtained, and then the average across all three mice was graphed along with the standard error for each antibody.

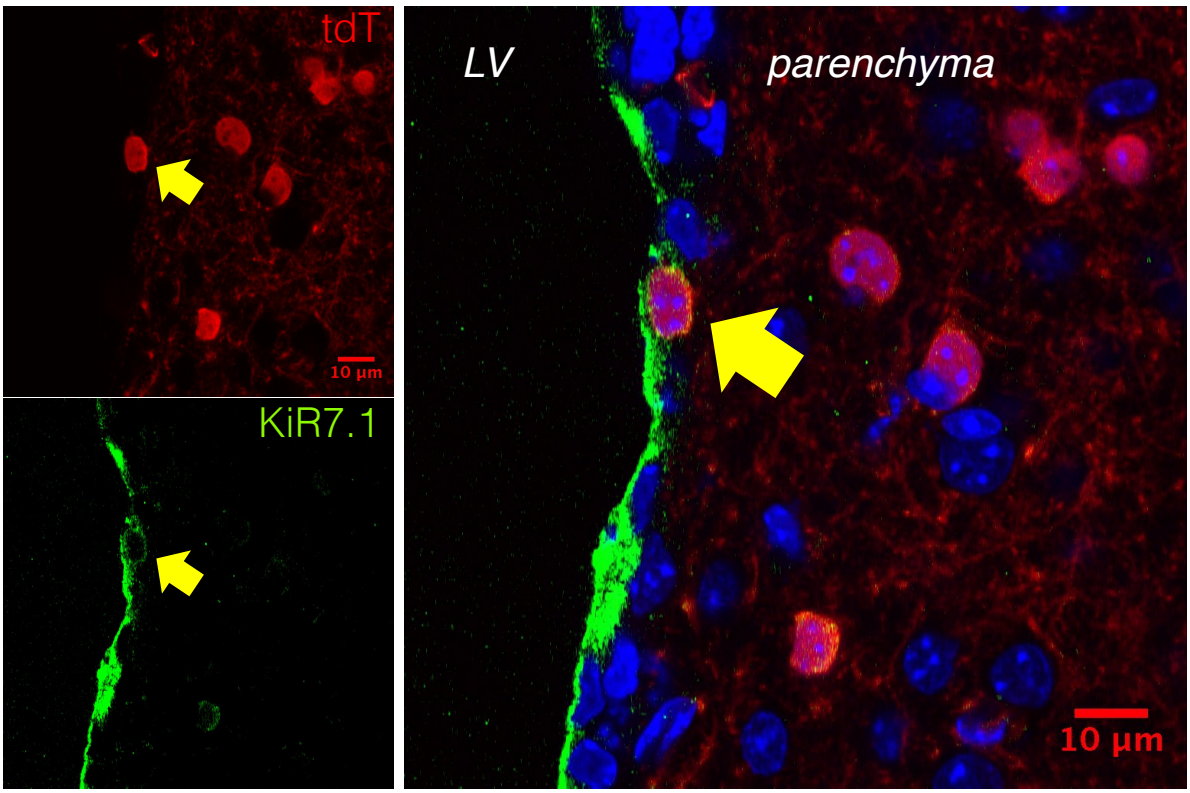
4.2 *In vivo* validation of novel and specific markers of the Group1 cluster.

The bioinformatics analysis predicted that each of the three clusters could be identified *in vivo* by their expression of specific marker genes. To determine whether this was the case we immunostained IPNK brains with antibodies against the proteins that the genes encode for, and analysed whether Penk-tdTomato$^+$ cells were immunopositive for said protein markers.

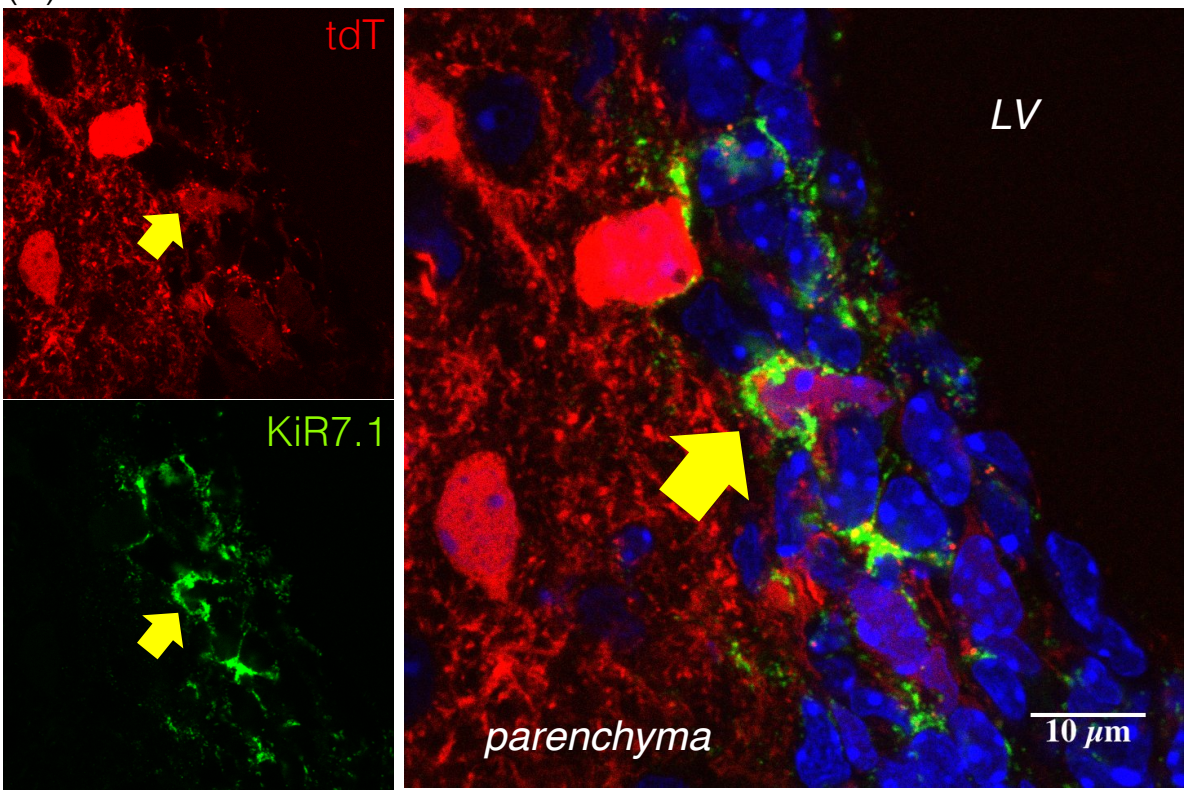
We began by analysing the Group1 markers. One of the strongest markers for Group1 was the *Kcnj13* gene, which encodes the KiR7.1 inwardly rectifying potassium channel. We stained an IPNK brain section with an antibody against KiR7.1 and visualised the stained sections using a high-powered confocal microscope and found that, indeed, Penk-tdTomato$^+$ colocalises with KiR7.1 (Figure 4.2a & b). We saw that the staining was restricted to the V-SVZ niche, as the signal is almost undetectable in the striatal parenchyma (Figure 4.2c). This confirms the hypothesis that *Kcnj13* is a specific marker of a subset of NSCs that reside in the V-SVZ niche. We had hypothesised that *Kcnj13* is a marker for a subset of B-cells (V-SVZ neural stem cells). According to this hypothesis, some Penk-tdTomato$^+$/KiR7.1$^+$ cells should be immunopositive for the B-cell marker GFAP. Figure 4.3 shows that we were able to find such cells in the V-SVZ. This is a key finding for this study, as it recapitulates, *in vivo,* a finding made several years ago in neurospheres—that GFAP$^+$ "early" neurospheres express KiR7.1, but not Dcx$^+$ "late" neurospheres (Pruss *et al.,* 2011).

We performed this analysis also for the prolactin receptor (Prlr), which was shown by the bioinformatics analysis to be a specific marker for Group1 cells. Figure 4.4 shows that strong staining for Prlr can be seen at the V-SVZ, and that some of these cells are also Penk-tdTomato$^+$. Similar to KiR7.1, the staining is largely restricted to the V-SVZ niche. Thus, I was able to validate the bioinformatics finding by showing that the Group1 markers Kcnj13 and Prlr specifically mark a subset of Penk-tdTomato$^+$ cells in the V-SVZ niche.

(a)


tdT
10 μm
KiR7.1
LV
parenchyma
10 μm

(b)


tdT
KiR7.1
LV
parenchyma
10 μm

(c)

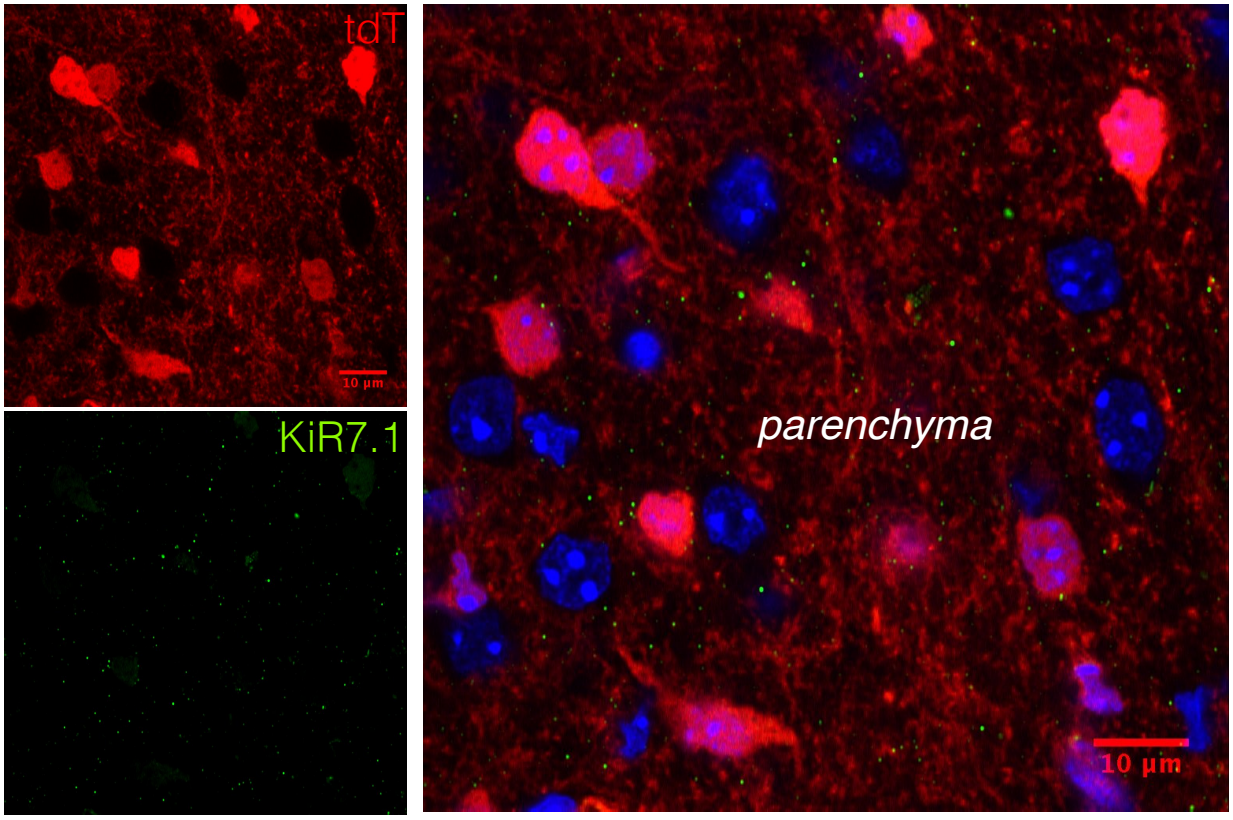


Figure 4.2. The Group1 marker Kcnj13/KiR7.1 colocalises with Penk-tdTomato in some of the the cells of the V-SVZ (a & b) A 2-month old IPNK brain was cryosectioned at 30µm and immunostained with an antibody against KiR7.1 and visualised using a high-powered confocal microscope. The ***yellow*** arrows point to cells that are Penk-tdTomato$^+$ and KiR7.1$^+$ at the V-SVZ. Note that not all Penk-tdTomato$^+$ cells are KiR7.1$^+$, and not all KiR7.1$^+$ are Penk-tdTomato$^+$. **(c)** The striatal parenchyma is visualised. No discernable staining for KiR7.1 can be seen, even though Penk-tdTomato$^+$ MSNs are visible.

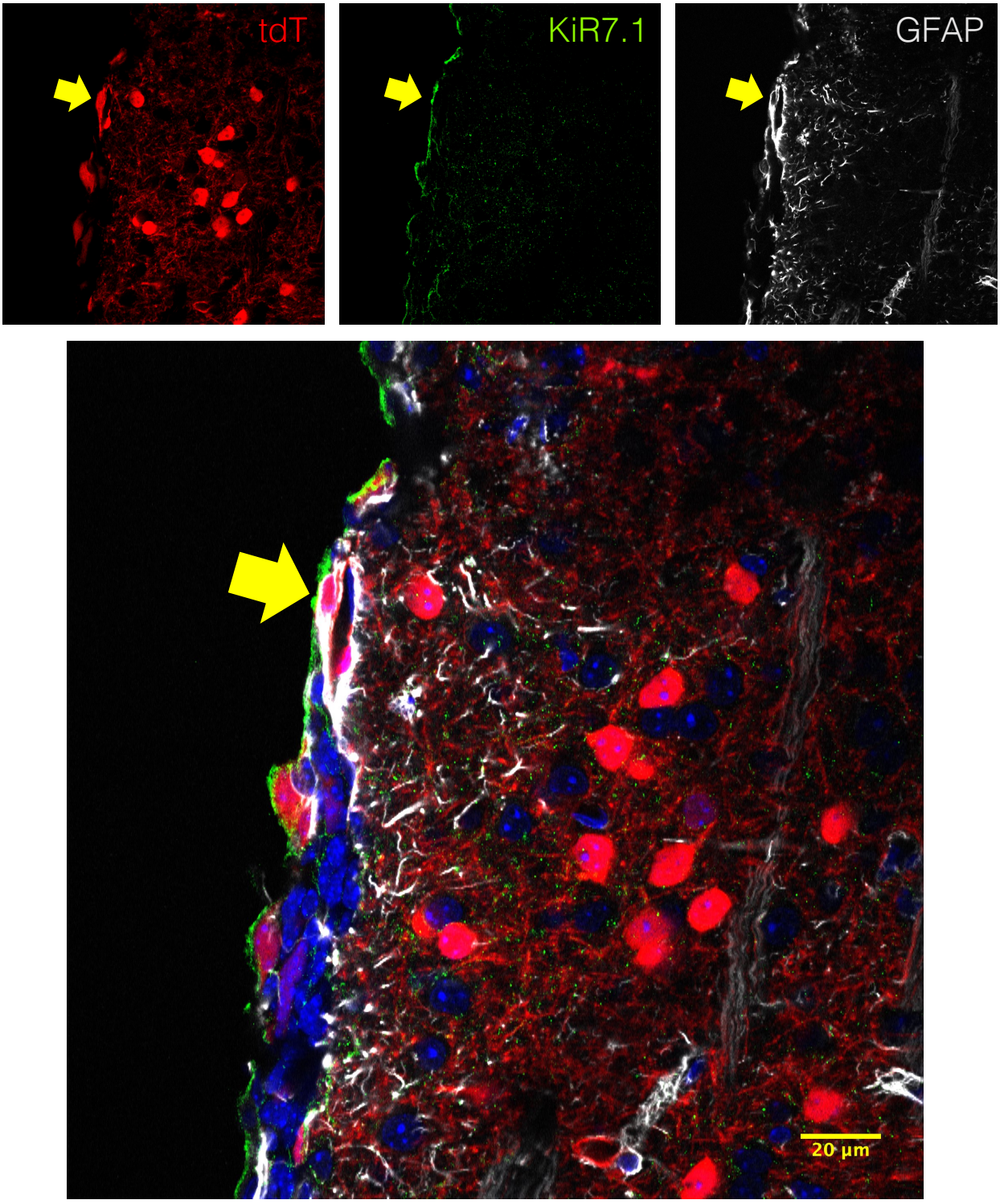


Figure 4.3. The Group1 marker Kcnj13/KiR7.1 colocalises with some Penk-tdTomato⁺ cells and GFAP in the cells of the V-SVZ. A 2-month old IPNK brain was cryosectioned at 30µm and immunostained with an antibody against KiR7.1 and GFAP and visualised using a high-powered confocal microscope. The ***yellow*** arrow points to a cell that is Penk-tdTomato⁺ and KiR7.1⁺ and GFAP⁺ at the V-SVZ.

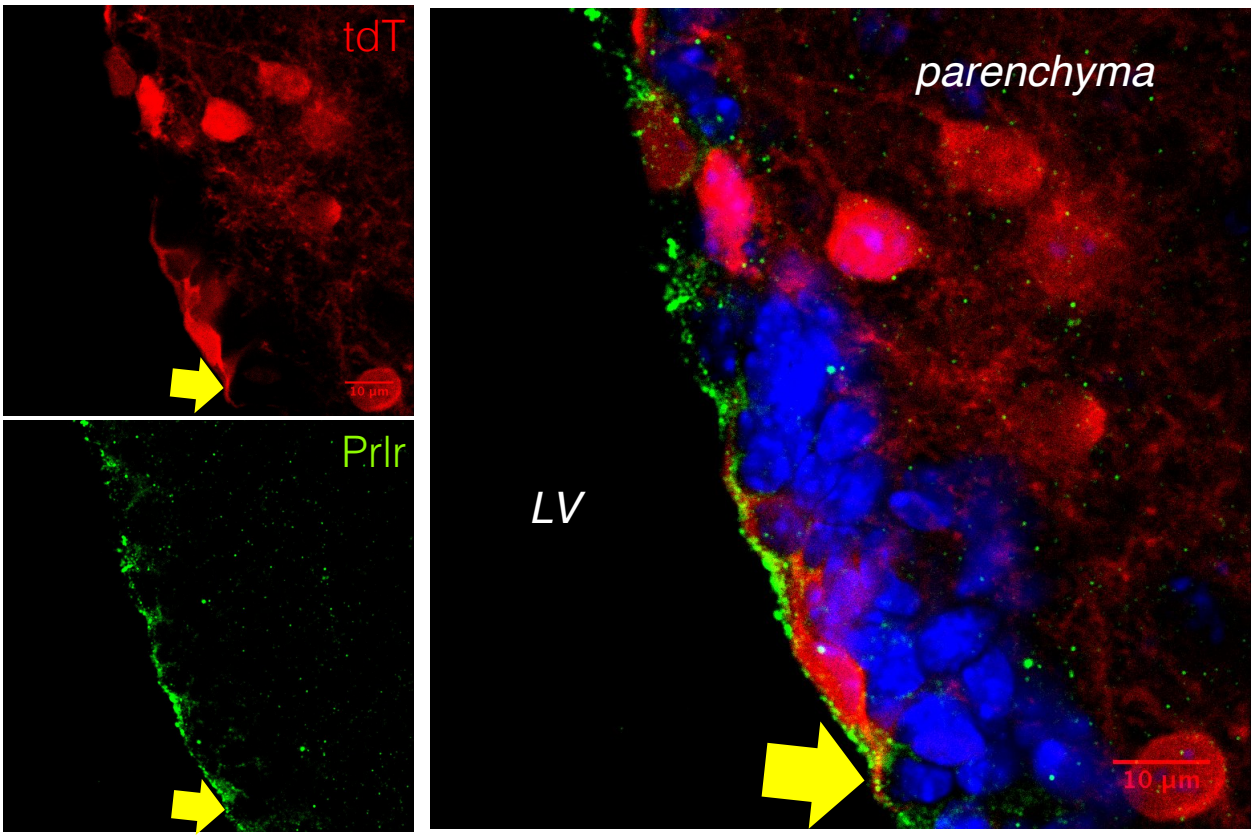


Figure 4.4. The Group1 marker Prlr colocalises with Penk-tdTomato in some of the cells of the V-SVZ. A 2-month old IPNK brain was cryosectioned at 30µm and immunostained with an antibody against Prlr and visualised using a high-powered confocal microscope. Prlr signal is found on the outermost layer of the V-SVZ and only stains Penk-tdTomato$^+$ cells at this layer. Note the particularly strong staining of the processes of Penk-tdTomato$^+$ cells ***(yellow*** arrow)**.** Prlr does not stain the cells of the striatal parenchyma.

4.3 *En face* staining of the subventricular zone using wholemounts.

As discussed in detail in the ***Introduction***, the NSCs of the V-SVZ are heterogeneous in their function, their morphology, and their spatial distribution within the germinal niche. Only the most quiescent NSCs contact the lateral wall, proper, of the lateral ventricle. The other NSCs and neuroblasts do not, and are pushed closer to the striatal parenchyma. Thus, one would expect that if Group1 represented a subset of B-cells, then a specific marker of Group1 should be seen to contact the wall—in other words, we should be able to detect the expression of the marker protein on the wall. The V-SVZ is usually viewed in the sagittal or coronal planes, and is thus only seen as a sliver. Though challenging, it is possible to isolate, stain, and mount just the lateral wall of the lateral ventricle. This preparation has been termed a wholemount, and allows one to visualise the lateral wall *en face* (Mirzadeh *et al.,* 2008; Mirzadeh *et al.,* 2010). Figure 4.5a shows a wholemount stained with Dcx, with the focal point just underneath the wall. The stereotypical "tubular" chains of migrating Dcx^+ neuroblasts previously visualised by other groups can be seen, indicating that we were able to properly isolate and stain intact wholemounts (Mirzadeh *et al.,* 2008; Mirzadeh *et al.,* 2010). A further concern was background or nonspecific staining, due to the unusual preparation. Figure 4.5b shows secondary antibody controls. We found that there is no strong background or nonspecific staining, with the exception of some staining of blood vessels in the 488 nm channel. Satisfied with the preparation, we proceeded to stain wholemounts with an antibody against KiR7.1 to validate the expression of this Group1 marker in cells of the lateral ventricular wall. Figure 4.6a shows an IPNK wholemount stained with KiR7.1. Strong staining for KiR7.1 can be seen in a localised area of the wholemount. The staining pattern suggests that the plasma membrane of some Penk-$tdTomato^+$ cells in this area stain positively for KiR7.1. Figure 4.6b shows, in higher resolution, another wholemount area isolated from a different IPNK brain. The Penk-$tdTomato^+$ cells contacting the wall show a flat, "stellate" morphology previously seen in electron micrographs of B-cells (Doetsch *et al.,* 1997; Doetsch *et al.,* 1999). Furthermore, some

of these Penk-tdTomato$^+$ cells colocalise with KiR7.1. The wholemount staining analysis confirms that: (1) KiR7.1 is a marker of a subset of putative B-cells that contact the wall of the lateral ventricle; (2) Penk-tdTomato$^+$ and KiR7.1$^+$ colocalise; (3) some Penk-tdTomato$^+$ cells at the wall show the stereotypical morphology of B-cells. In conclusion, this provided further evidence in support of the hypothesis that Group1/Group15 represent a novel subset of B-cells that can be identified by the expression of Penk-tdTomato$^+$ and *Kcnj13*.

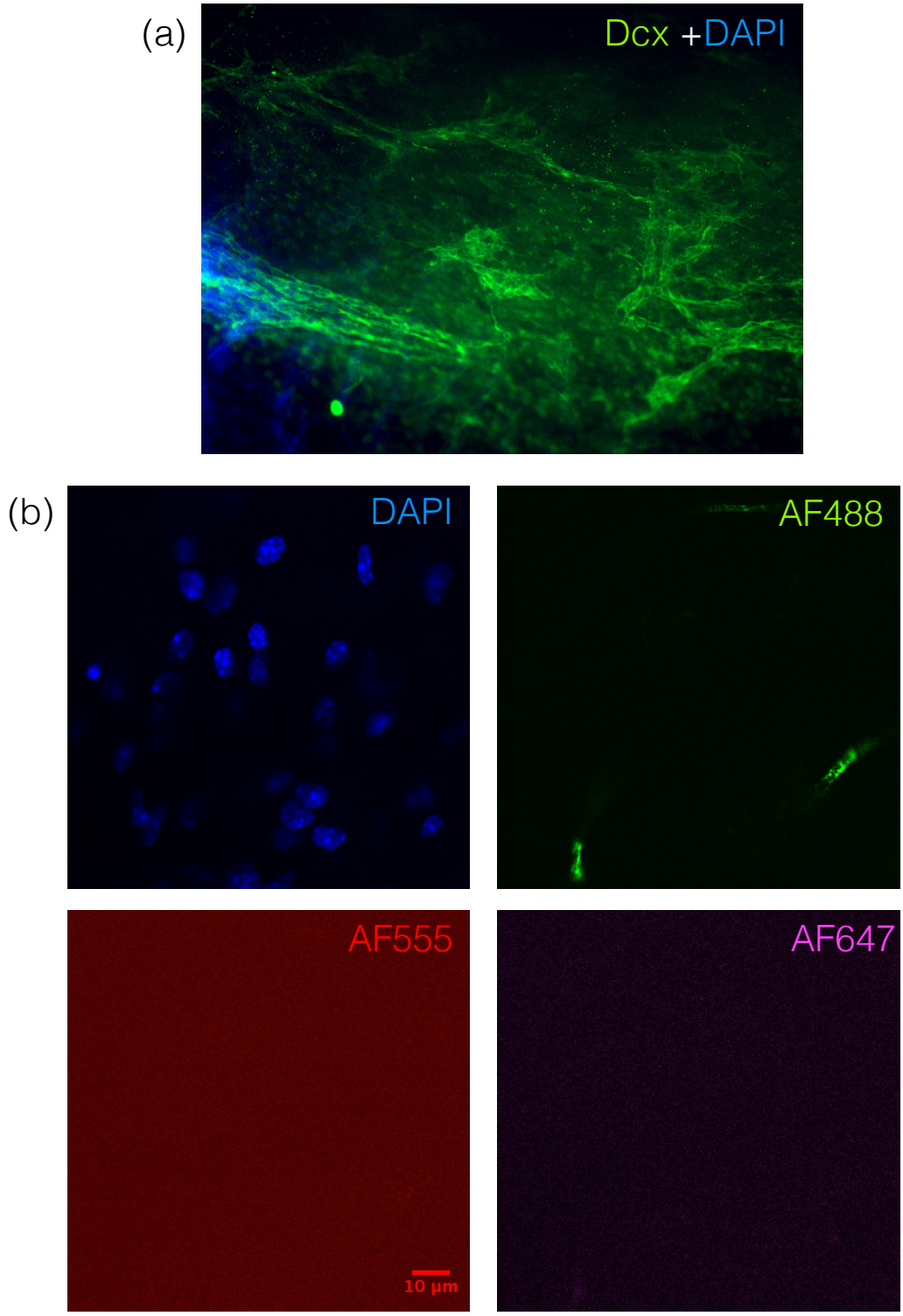


Figure 4.5. Wholemount preparations of the lateral wall of the lateral ventricle. **(a)** A WT wholemount of the lateral wall of the LV was stained with an antibody against Dcx and DAPI. The characteristic tubular chains of migrating neuroblasts are visible indicating an intact wholemount. **(b)** Secondary antibody controls for a WT wholemount. There is no non-specific binding of secondary antibodies in the channels, there is however some autofluorescence in the 488nm channel arising from blood vessels.

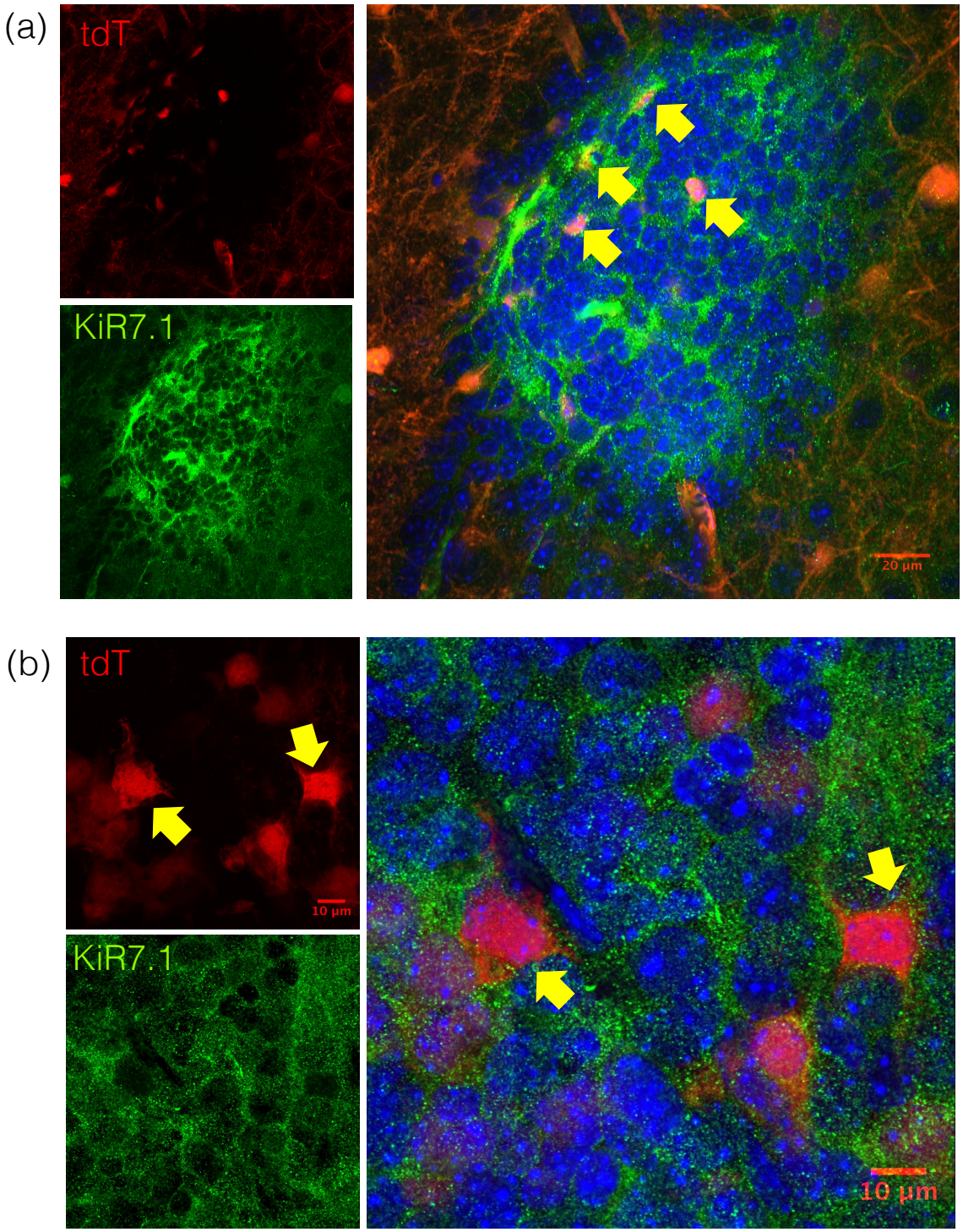


Figure 4.6. Wholemount staining reveals some Penk-tdTomato⁺ cells at the lateral wall colocalising with KiR7.1. (a) An IPNK wholemount preparation was immunostained with an antibody against KiR7.1. Intense staining of KiR7.1 is limited to a cluster of cells at the wall, some of these colocalise with Penk-tdTomato (see arrows). **(b)** A higher resolution image of a different part of the wholemount showing the stellate nature of Penk-tdTomato⁺ cells at the wall and colocalisation with KiR7.1 (see arrow).

4.4 Validation of specific markers of the Group2 cluster.

Next, we validated some of the markers of the Group2 cells. After analysing the bioinformatics data, we hypothesised that Group2 represented another subgroup of quiescent NSCs—one that was slightly more differentiated, pseudotemporally, than the Group1 cells. The analysis found that *GLAST*, a marker used to isolate quiescent NSCs, was specifically upregulated in this cluster. Thus, we immunostained IPNK brain sections with antibodies against the GLAST protein to determine whether it would colocalise with the Penk-tdTomato$^+$ cells of the V-SVZ. Confirming the bioinformatics data and in line with our hypothesis, we found Penk-tdTomato$^+$/GLAST$^+$ cells at the lateral wall of the LV (Figure 4.7). We note that the cell indicated with the yellow arrow denoting colocalisation seems to be in the second layer of cells, not in the outermost layer of the V-SVZ. This is interesting because the Group1 markers seem to label cells at the outermost layer. Thus, this may represent a spatial difference, along with the molecular differences, between Penk-tdTomato$^+$ Group1 and Group2 cells. Similar to the Group1 markers, we found that the staining is strongest in the germinal niche and drastically reduced in the striatal parenchyma. Note that though staining for GLAST can be seen in the striatal parenchyma (since it is also a strong marker for astrocytes), here the signal does not colocalise with Penk-tdTomato.

The bioinformatics analysis also indicated that Group2 was specifically upregulated for two gap junction proteins that are critical for neurogenesis, those being Connexin43 (*Gja1*) and Connexin30 (*Gjb6*). Immunochemical analysis of an IPNK brain section indicates that similar to GLAST, we could find Penk-tdTomato$^+$ cells at the V-SVZ that stain positively for Connexin43 and Connexin30 (Figure 4.8 a and b). Thus, I was able to validate, *in vivo*, three markers of the Group2 cluster. The presence of these stem cell markers colocalizing with some Penk-tdTomato$^+$ cells at the V-SVZ, further supports the hypothesis that Group2 is a subset of quiescent NSCs.

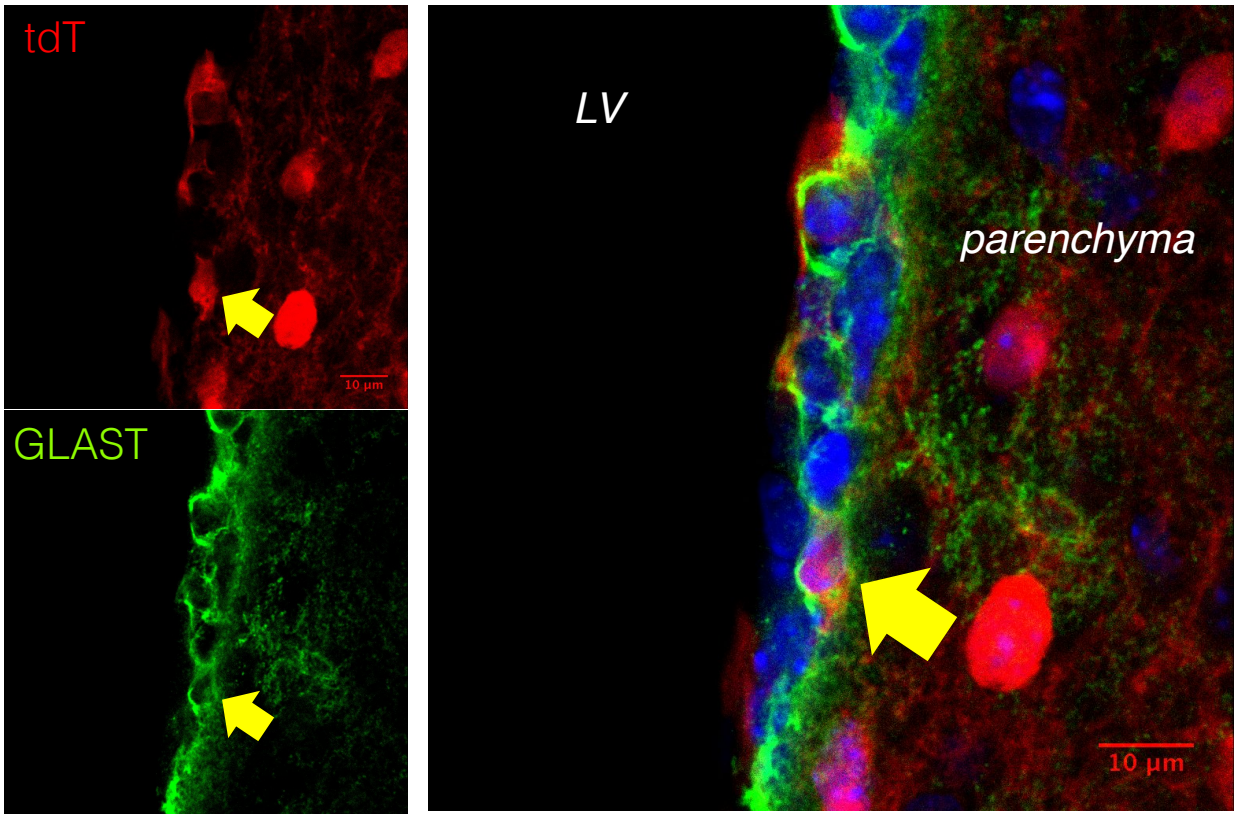


Figure 4.7. The Group2 marker GLAST colocalises with some Penk-tdTomato⁺ cells of the V-SVZ. A 2-month old IPNK brain was cryosectioned at 30µm and immunostained with an antibody against GLAST and visualised using a high-powered confocal microscope. The ***yellow*** arrow points to a cell that is Penk-tdTomato⁺ and GLAST⁺ at the V-SVZ.

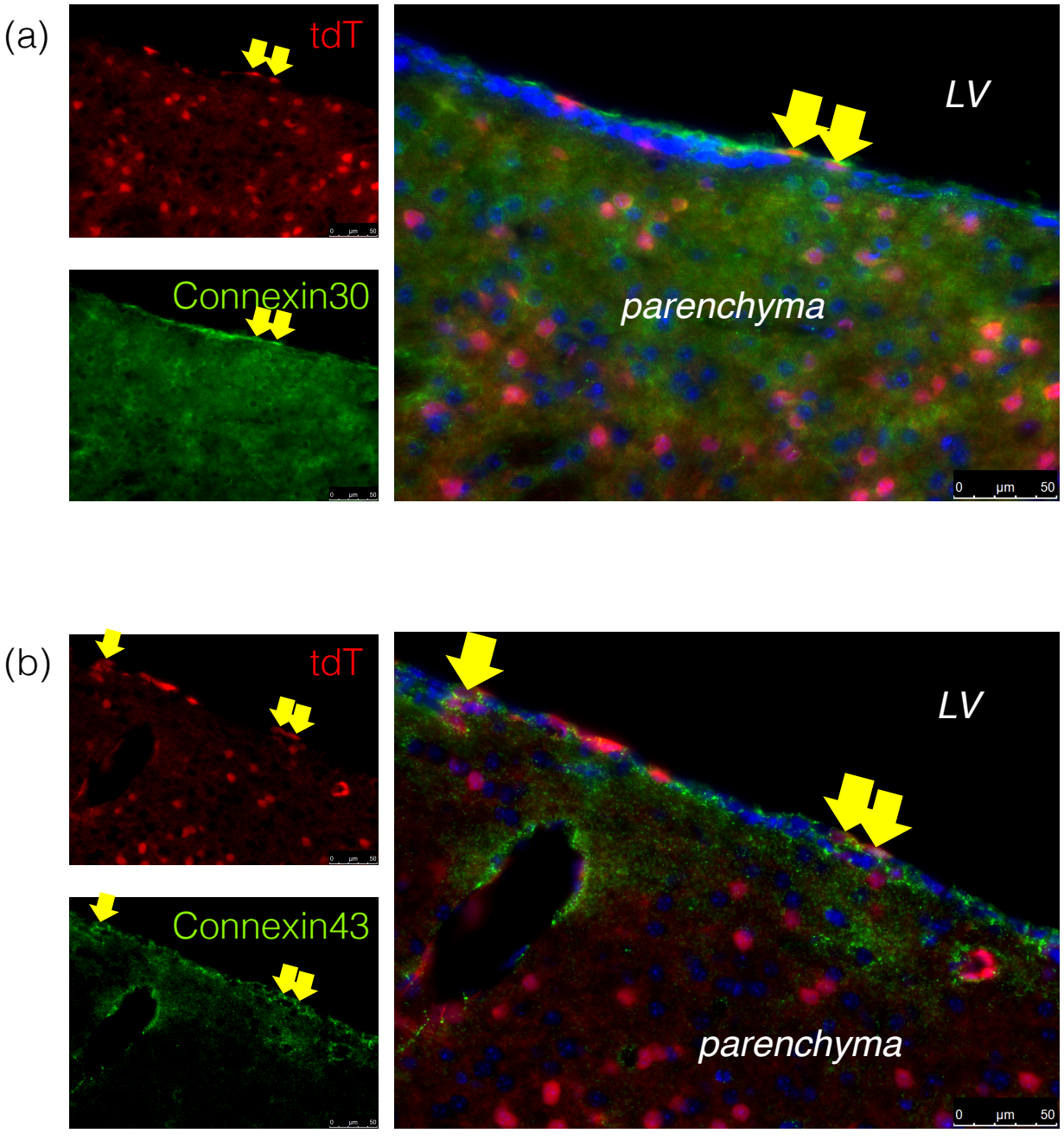
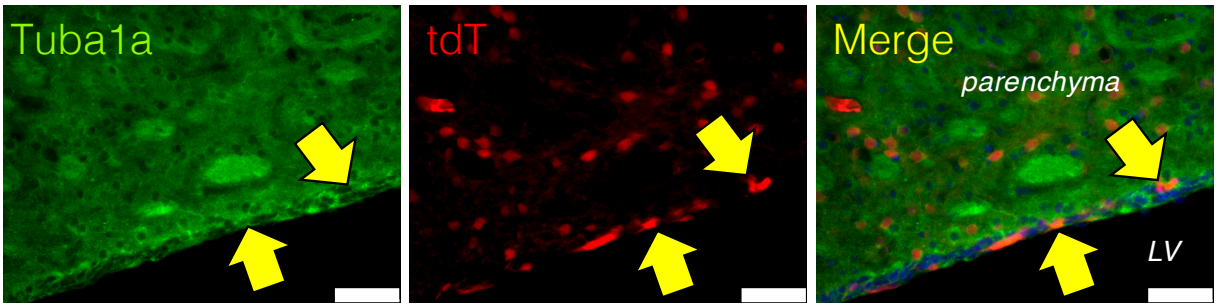


Figure 4.8. The Group2 markers Connexin30 and Connexin43 colocalise with Penk-tdTomato in the some of the cells of the V-SVZ. A 3-month old IPNK brain was cryosectioned at 30µm and immunostained with an antibody against Connexin30 **(a)** and Connexin43 **(b)** and visualised using an epifluorescence microscope. The ***yellow*** arrows point to cells that are Penk-tdTomato^{+} and either Connexin30^{+} or Connexin43^{+} at the V-SVZ.

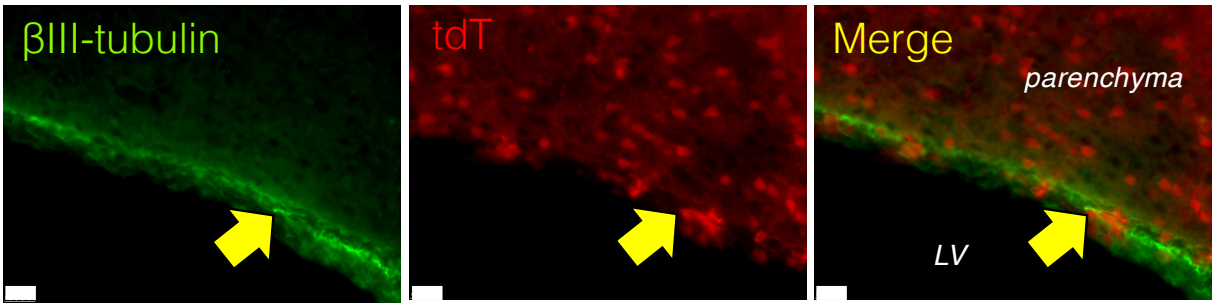
4.5 Validation of specific markers of the Group3 cluster.

The bioinformatics data and the quantitative immunochemical analysis clearly indicated that Penk-tdTomato$^+$ labels neuroblasts. In Figure 4.1 we had validated the expression of *Sox11* by its colocalising with some Penk-tdTomato$^+$ cells in a quantitative manner, as was predicted by the bioinformatics analysis. Figure 4.9 shows a confocal image of one such section. We next completed the analysis by staining for five other neuroblast markers that were specifically upregulated in the Group3 cluster (Figure 4.1). I began by validating the expression α-tubulin subunit *Tuba1a*. When I immunostained an IPNK brain section against an antibody for Tuba1a, I was able to find Penk-tdTomato$^+$/Tuba1a$^+$ cells at the V-SVZ (Figure 4.10a). *βIII-tubulin* is a marker of neuroblasts and specifically upregulated in Group3. Similar to Tuba1a, I was able to validate its presence in Penk-tdTomato expressing cells of the V-SVZ (Figure 4.10b). The bioinformatics analysis showed that *CD24a* was a marker for Group3 cells. Staining for this protein in IPNK brains confirmed that it can be found in some Penk-tdTomato$^+$ cells in the germinal niche (Figure 4.10c). *Dcx* is another canonical marker of migrating neuroblasts that was found to be a specific marker of Group3 cells. Figure 4.10d shows that this can be validated *in vivo*. Finally, I also validated the expression of the marker gene *Dlx5 in vivo*. In line with the other Group3 markers, I found that it is expressed in some Penk-tdTomato$^+$ cells of the V-SVZ (Figure 4.10e). In summary, I was also able to validate the *in vivo* expression of markers for the Group3 cluster. Furthermore, the combined expression of these extensively studied markers indicates that the Penk-tdTomato$^+$ cells of the V-SVZ expressing them are terminally differentiating and migrating neuroblasts.

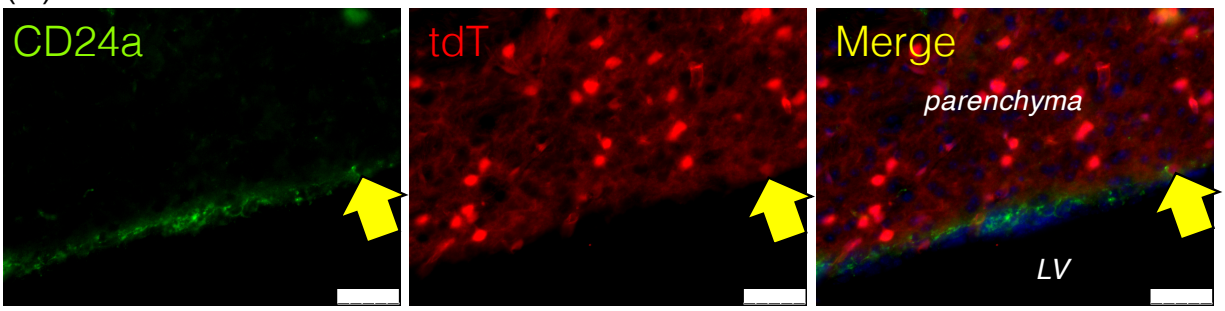
(a)


Tuba1a
tdT
Merge
parenchyma
LV

(b)


βIII-tubulin
tdT
Merge
parenchyma
LV

(c)


CD24a
tdT
Merge
parenchyma
LV

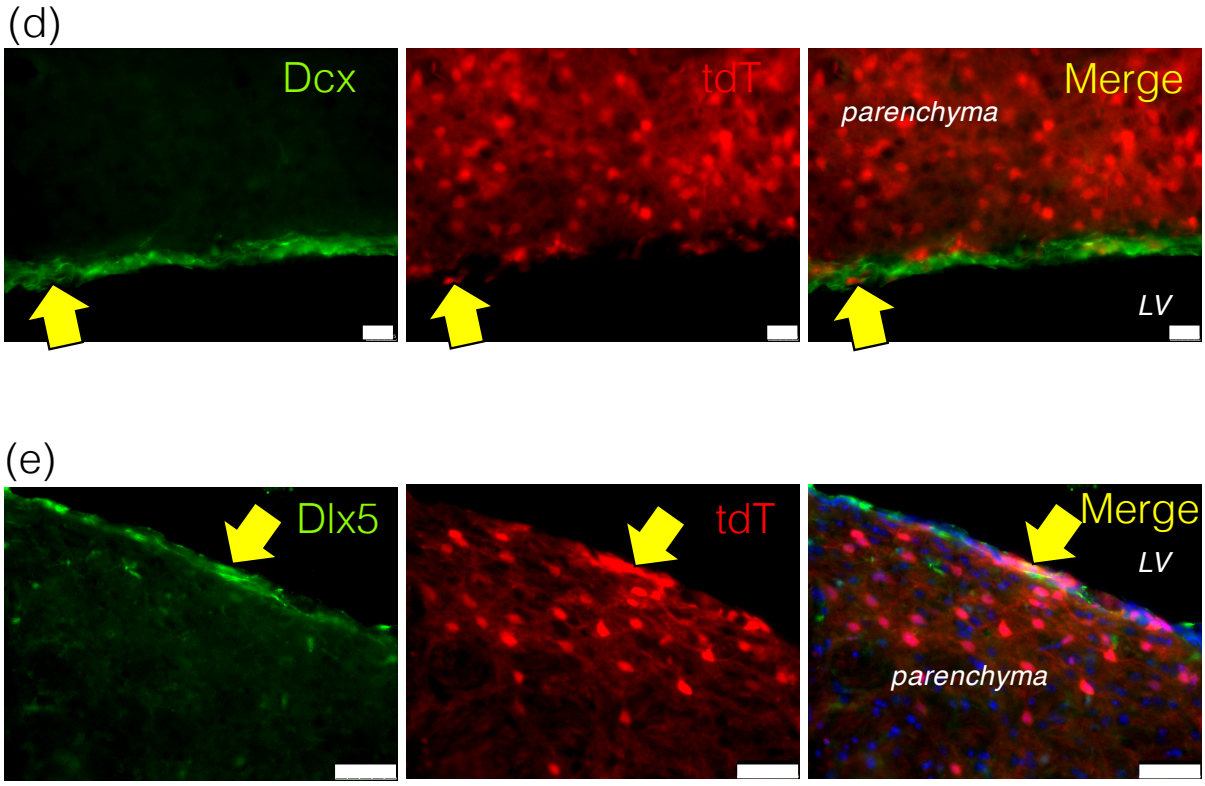


Figure 4.10. The Group3 markers colocalise with Penk-tdTomato in some of the cells of the V-SVZ. A 2-month old IPNK brain was cryosectioned at 30µm and immunostained with antibodies against the protein markers indicated in the leftmost panels (**Tuba1a**, **βIII-tubulin**, **CD24a**, **Dcx**, and **Dlx5**, respectively). The ***yellow*** arrows points to cell(s) that are Penk-tdTomato^{+} and positive for the respective marker at the V-SVZ. Scale bars included in each panel, scale bar = 25µm in βIII-tubulin and Dcx panels, scale bar = 50µm in Tuba1a, CD24a, and Dlx5 panels.

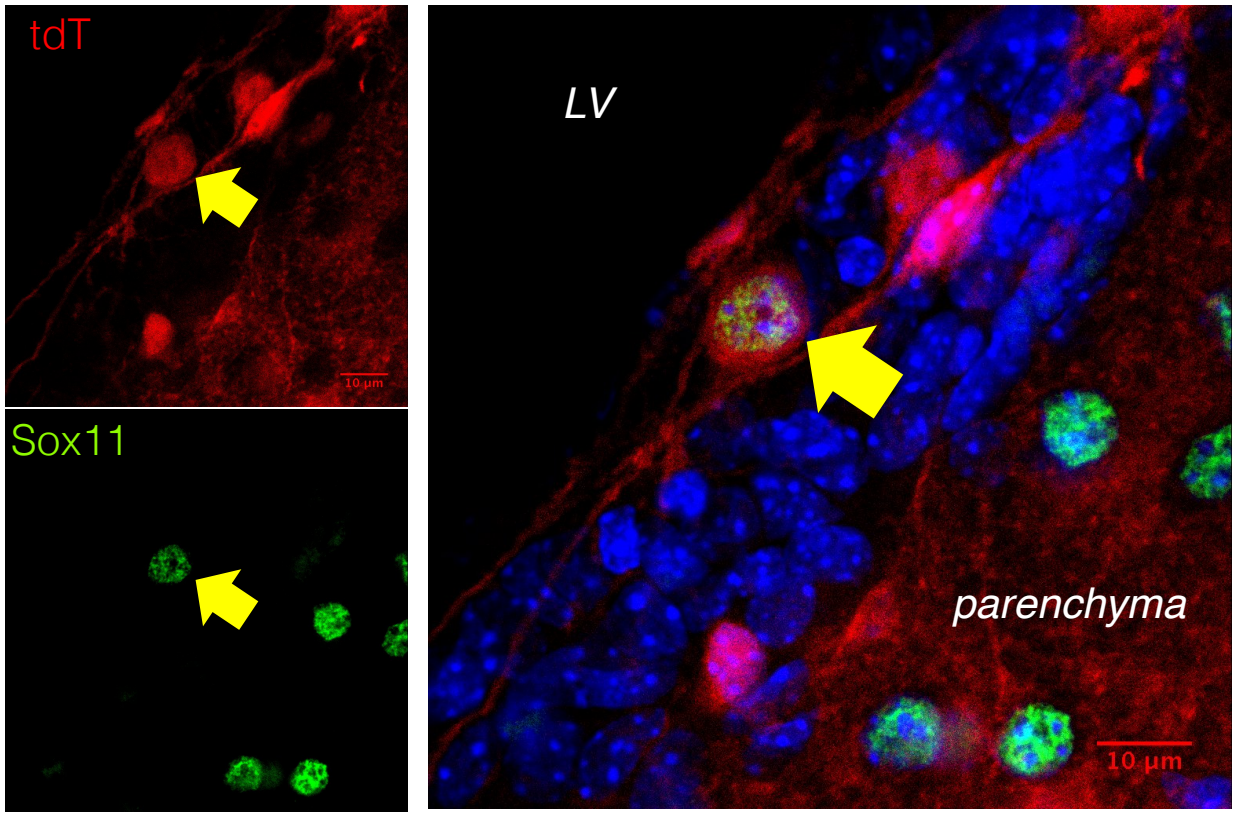


Figure 4.9. The Group3 marker Sox11 colocalises with Penk-tdTomato in some of the cells of the V-SVZ. A 2-month old IPNK brain was cryosectioned at 30µm and immunostained with an antibody against Sox11 and visualised using a high-powered confocal microscope. The ***yellow*** arrow points to a cell that is Penk-tdTomato^{+} and Sox11^{+} at the V-SVZ.

4.6 Three-dimensional visualisation of Penk-tdTomato$^+$ cells at the V-SVZ.

We also wanted to visualise the Penk-tdTomato$^+$ cells within the germinal niche, particularly the ones that make direct contact with the lateral wall of the lateral ventricle, in three dimensions in an intact brain to better understand their spatial distribution. To achieve this, we used the CLARITY method developed by the Diesseroth group (Tomer *et al.,* 2014). This method allows the experimenter to transform an intact brain into a transparent "hydrogel-tissue hybrid". Since the tissue is made transparent one can use modified confocal or lightsheet microscopes to visualise signals coming from genetically encoded fluorescent reporters. We processed an IPNK brain and "cleared" it for 15 hours in preparation for confocal imaging (Figure 4.11a). The brain was then imaged and post-processing was performed to generate a 3D image of the brain. Figure 4.11b shows snapshots of a movie generated using this method. At the lateral wall we can clearly see patches of Penk-tdTomato$^+$ cells that are unevenly distributed throughout the lateral wall. This uneven distribution is suggested by the Penk-tdTomato$^+$ cell clusters in the coronal sections. Furthermore, the patches are also similar to the patches of KiR7.1 staining seen in the wholemount (Figure 4.6). As expected, when the striatal parenchyma is imaged strong signal for Penk-tdTomato$^+$ cells can be detected in the MSNs.

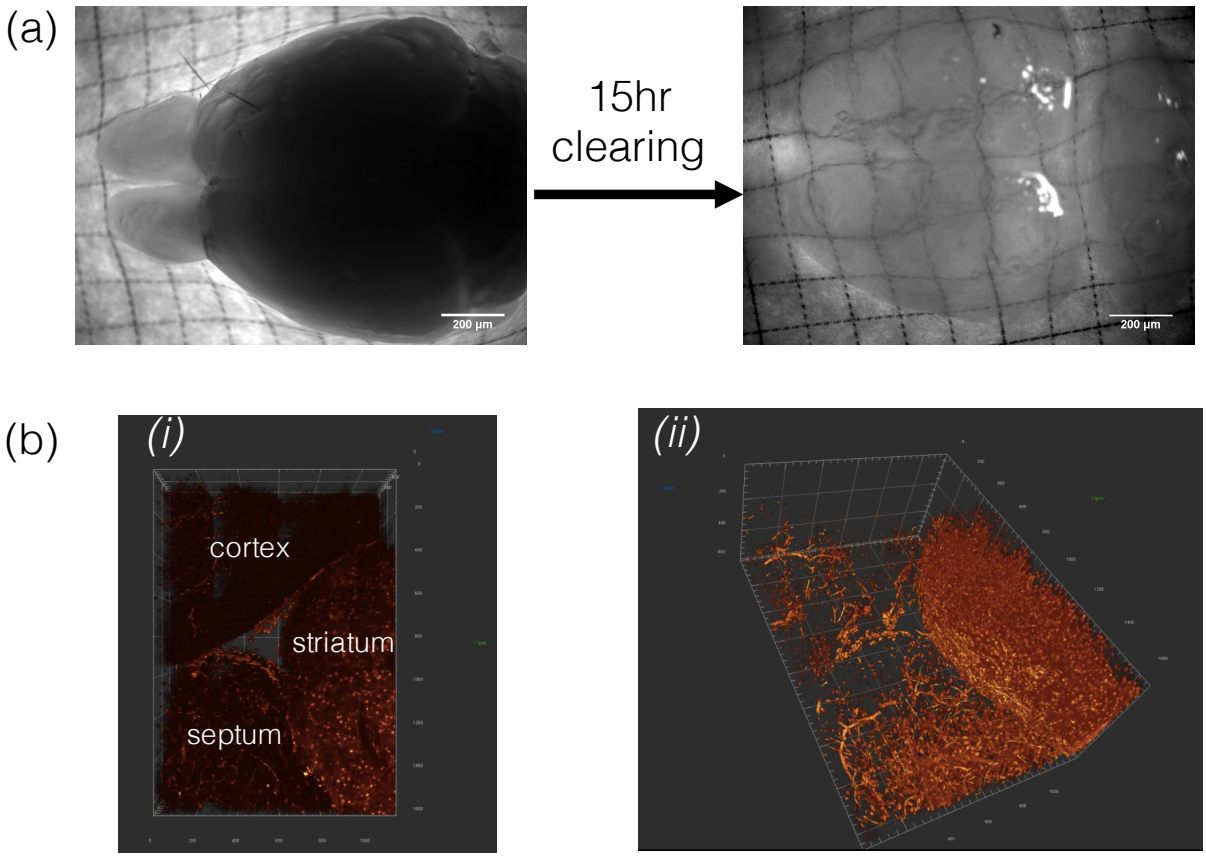


Figure 4.11. CLARITY Imaging. (a) Shows the result of the electrophoretic clearing process—the brain is made transparent and amenable to fluorescence imaging of deep, intact tissues. **(b)** Shows several stills from the movie generated by the confocal images obtained from CLARITY imaging. ***(i)*** Shows a global layout of the image. ***(ii)*** Shows how the 3D image was tilted to look *en face* at the lateral wall. ***(iii & iv)*** The ***blue*** arrow points to patches of Penk-tdTomato^{+} cells at the wall. ***(v)*** Shows the striatal parenchyma and the round MSNs therein.

4.7 Using novel cell-surface markers to functionally isolate Group1 cells.

Having validated the expression of the novel Group1 markers *in vivo*, we wanted to determine if this finding could be functionally useful. To unequivocally demonstrate that Penk-tdTomato^{+}/KiR7.1^{+} cells of the V-SVZ are, in fact, a subset of B-cells one would need to isolate them and experimentally probe the cells to understand their potency and physiology. We reasoned that since *Kcnj13* encodes the cell-surface KiR7.1 ion channel, we should be able to use an antibody against an extracellular motif of the channel to isolate Penk-tdTomato^{+}/KiR7.1^{+} cells by FACS. To determine successful isolation of the cells of interests we would perform high-throughput gene expression analysis using a panel of pre-determined genes.

To perform this, we purchased an antibody directed against an extracellular loop of KiR7.1 (amino acid residues 80-94), which was conjugated with an ATTO-488 fluorophore. This would enable us to select for cells expressing this protein (similar to the CD11b-APC we used for all the sorts). We isolated three different populations of cells to test for the specificity of the protein marker. The first population was the target population: Penk-tdTomato^{+}/Kcnj13^{+}/CD11b^{-}. The second population we collected was Penk-tdTomato^{+}/Kcnj13^{-}/CD11b^{-} to ensure that the antibody was specific (although the perfect experiment to prove this would have been to omit also the tdTomato fluorescent protein from the mouse striatum). Finally, we hypothesised that since *Kcnj13* was a specific marker for quiescent NSCs in the striatum, mature MSNs should not express this marker. Thus, to determine whether mature MSNs (isolated using the M-Peak gating paradigm described in the previous chapter) expressed *Kcnj13*, cells were isolated only by a Penk-tdTomato^{+}/CD11b^{-} signature. Further, we decided to perform the experiment using bulk-cell analysis due to the extreme rarity of the cells of interest and the difficulty in keeping the cells alive. Therefore, I collected three technical replicates of 10 cells for each of the three isolated cell types. Due to challenging nature of the isolation, we used the striata from three different IPNK mice to generate the required 9 bulk-cell samples.

Figure 4.12a shows the result of this experiment. The analysis was performed as described previously. In terms of *Kcnj13* expression, we find that it is indeed absent from the populations of mature MSNs. As expected, *Kcnj13* expression is highest in the Penk-tdTomato^{+}/Kcnj13^{+} populations and reduced in the Penk-tdTomato^{+}/Kcnj13^{-} population. The minimal expression in the Penk-tdTomato^{+}/Kcnj13^{-} could possibly be due to one or two Kcnj13^{+} entering this sorted population since this is a bulk-cell experiment and we still retain tdTomato fluorescent protein in these sample. In general, it shows that the KiR7.1 antibody is specifically able to isolate *Kcnj13*-expressing cells. Next, to investigate the nature of the three cell types we used sets of genes that were shown in the bioinformatics data to be specific for each of the three clusters (some of which we also validated by immunofluorescence), as well as genes that identify mature MSNs. The Penk-tdTomato^{+}/Kcnj13^{+} and Penk-tdTomato^{+}/Kcnj13^{-} populations show relatively high expression of Group1 markers, while they are largely absent in the MSNs. We also note that two Group1 markers, *CD59a* and *Igfbp2*, were more enriched in the Penk-tdTomato^{+}/Kcnj13^{-} population. This indicates that there may be heterogeneity in the Penk-tdTomato^{+} cells of V-SVZ, as expected. In line with the hypothesis that *Kcnj13* is a marker for Group1 cells, the Group2 markers *Gja1* (Connexin43) and *Itih3* are more enriched in the Penk-tdTomato^{+}/Kcnj13^{-} than in the Penk-tdTomato^{+}/Kcnj13^{+} population. This makes sense, as Penk-tdTomato is expressed in at least three cell types, including the Group2 cluster in the V-SVZ according to the bioinformatics data and the *in vivo* validation analysis (Figures 4.2, 4.3, 4.4, 4.7, 4.8, 4.9, 4.10). The Penk-tdTomato^{+}/Kcnj13^{+} and Penk-tdTomato^{+}/Kcnj13^{-} populations are depleted of the neuroblast marker *Sox11*. Finally, we confirm that M-Peak cells are mature MSNs as they show the expression of mature MSN markers such as *βIII-tubulin* and *Penk*. Also, confirming the sequencing data and the colocalisation analysis in Figure 4.1, we found that *CD133* is not expressed by any of the three cell populations. In summary, this experiment shows that the

newly identified markers can be used to isolate specific cell populations of the V-SVZ and that this method can be used for future experiments to perform functional analyses.

As outlined in the ***Materials and Methods*** and in our work, Nath *et al.* (2017), this method requires the experimenter to normalise the gene expression to a reference gene. In this experiment, we normalised the gene expression to the gene *Canx*. Although the data are quite definitive in showing that we can use an antibody against the Group1 marker to pull down a population that expresses markers expected for Group1 cells, we wanted to repeat the experiment and perform the normalisation using a different reference gene. The bioinformatics data indicated that the *Rn45s* gene, which encodes for the non-coding 45S Pre-Ribosomal 5 RNA, was highly and uniformly expressed across all three clusters. Thus, we repeated the experiment exactly as before except for the use of *Rn45s* as the reference gene instead of *Canx*.

Figure 4.12b shows the relative gene expression of the three populations as a heatmap. Confirming the first sorting experiment, we found that the expression of the two Group1 markers of interest—*Kcnj13* and *Prlr*—are expressed at the highest levels in the Penk-tdTomato^{+}/Kcnj13^{+} population and virtually absent in mature MSNs. The other Group1 markers—*Folr1*, *CD59a*, and *Igfbp2*—are more enriched in the Penk-tdTomato^{+}/Kcnj13^{-} populations. Indicating that, even within the Group1 cluster, there may be heterogeneity—with a *Kcnj13*-high and a *Kcnj13*-low population. We found that the Penk-tdTomato^{+}/Kcnj13^{+} population does not express Group2 markers, neuroblast or mature MSN markers—in line with the hypothesis that they are a group of quiescent NSCs. The Group2 genes are not highly expressed in the Penk-tdTomato^{+}/Kcnj13^{-} as in the previous experiment, but this population does express neuroblast markers. This indicates that Penk-tdTomato^{+}/Kcnj13^{-} population in this experiment may be mainly comprised of Penk-tdTomato^{+} neuroblasts. Finally, we found that the Penk-tdTomato^{+} M-Peak populations (i.e. mature MSNs) express the Group1 markers at very low levels or are completely depleted of them. They do, however, express neuroblast and mature MSN markers at the highest levels.

In conclusion, this set of experiments indicates that we were able to reliably isolate a cell population that corresponds to the Group1 cluster in the bioinformatics data using a protein marker for said cluster.

The experiments described in this chapter show that we were able to perform *in vivo* validation of the markers of the three clusters using different methods. Furthermore, our data show that the novel marker *Kcnj13*, and its encoded protein KiR7.1, can be used to pull down a population of cells that has a gene signature consistent with those of quiescent NSCs.

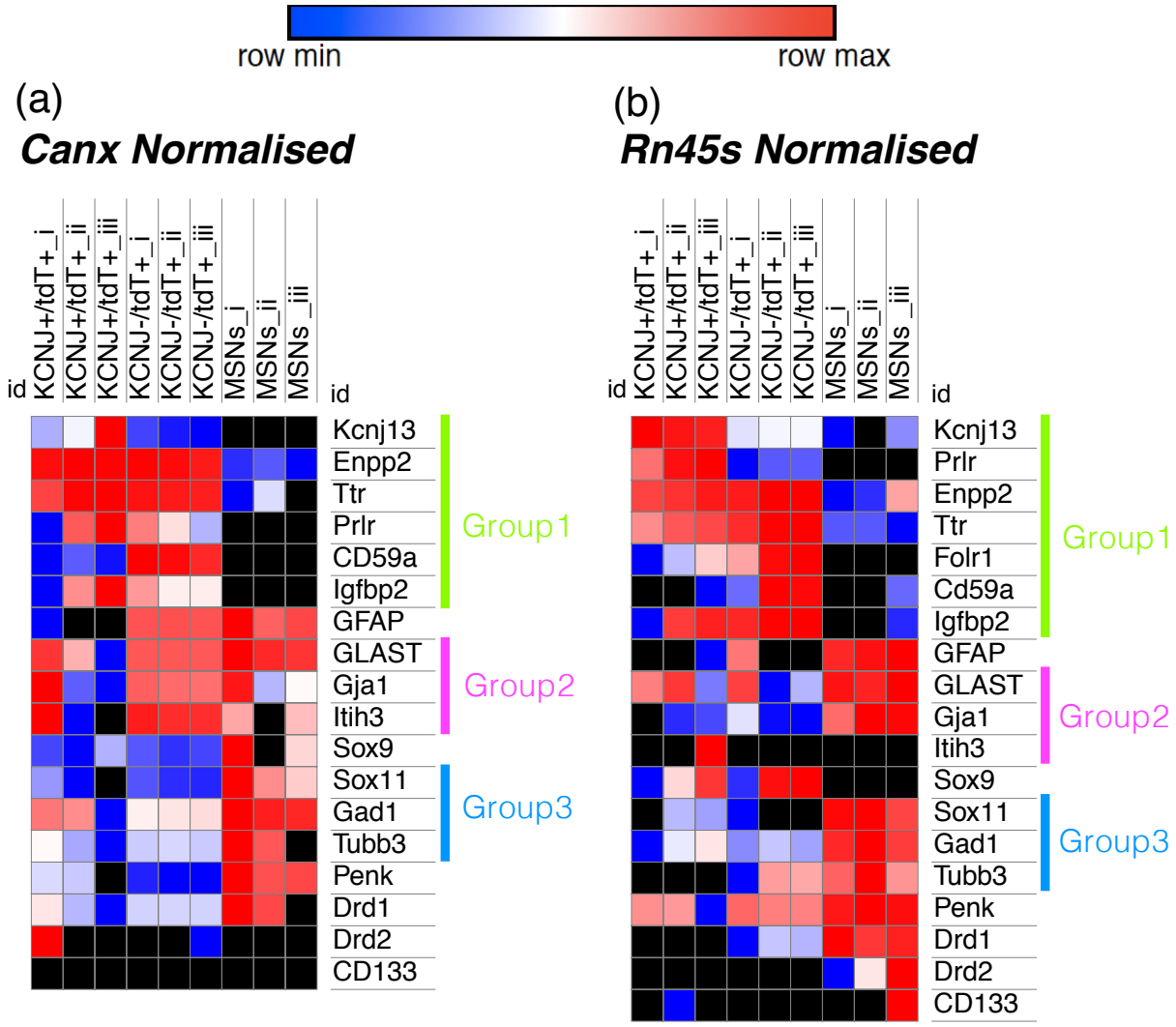


Figure 4.12. Kcnj13/KiR7.1 can be used to functionally isolate a group of cells with the molecular signature of Group1 cells. Striata were collected from IPNK mice to obtain bulk samples corresponding to the columns. Three technical replicates of 10-cell samples were collected for each isolated population, with the key population being Penk-tdTomato^{+}/KiR7.1^{+}. Using the Fluidigm system, the samples were assayed for their expression of the genes in each row. Group1/2/3 markers are indicated. For the relative expression analysis, the gene expression values were either normalised to the expression of Canx **(a)** or Rn45s **(b)**. High relative expression is indicated by red colouring and low by blue. Detailed information regarding the analytical method can be found in the ***Materials and Methods***. The heatmaps were generated using Morpheus (Broad Institute).

Chapter 5 – Results III: The *in vivo* functional analysis of Penk-tdTomato$^+$ cells in a partial progressive model of Parkinson's disease.

5.1 Establishing the partial progressive model of Parkinson's disease.

Extensive work has shown that the stem cells of the V-SVZ can be activated in response to different brain injury paradigms and that, in some instances, this activation can lead to neurogenesis (Arvidsson *et al.,* 2002; Llorens-Bobadilla *et al.,* 2015; Girasole *et al.,* 2018). The data presented in the two preceding chapters supports the hypothesis that Penk-tdTomato labels, among other things, a pool of quiescent neural stem cells. We, therefore, wanted to test whether this pool of Penk-tdTomato$^+$ cells could proliferate, migrate, and become neurogenic in response to a brain injury. This was of particular interest to us, because if successful, it could provide another means for therapeutic intervention for diseases such as Huntington's and Parkinson's.

Many different brain injury models have been employed to test the potential of the quiescent NSCs of the V-SVZ. The most commonly used injury paradigm has been the middle cerebral artery occlusion-induced stroke model (Arvidsson *et al.,* 2002; Llorens-Bobadilla *et al.,* 2015). This is however, a catastrophic and acute event, and can affect many different cell types. We wanted to test the potential of Penk-tdTomato$^+$ cells in a more refined injury model targeting the cells of the striatum specifically. To achieve this, we chose the 6-hydroxydopamine (6-OHDA)-mediated unilateral lesioning of the dorsal striatum to generate the partial progressive model of Parkinson's disease (Smith and Heuer, 2012; Bagga *et al.,* 2015; Girasole *et al.,* 2018). Aside from the fact that this model can be used to specifically target a cell population in a particular location, the unilaterality means that each animal can be used as its own within-subject control (Smith and Heuer, 2012). This makes for clean and straightforward effect comparisons. When the

toxin 6-OHDA is injected unilaterally directly into the striatum, it is taken up by the dopamine transporter and causes cell death by inducing oxidative stress and generating cytotoxic species (Simola, 2007). Since the dopaminergic cells of the substantia nigra project to the striatal parenchyma, 6-OHDA uptake leads to their death and results in the partial progressive model of Parkinson's disease. Although this is one of the more specific and least-toxic models of brain injury, the challenging surgery can lead to unintended complications in the mice.

To perform the experiment, we aimed to use 6 mice that would be unilaterally injected with 6-OHDA so that the experiment would be internally controlled and so we could compare the ipsilateral (injected) striatum to the contralateral (un-injected) striatum. We first began by doing a single stereotactic injection of 6-OHDA (750nl, 6mg/ml) in the dorsal striatum at position AP: -0.1mm from bregma; ML: 2.0mm from bregma; DV: -2.25mm from bregma. Four animals were injected with 6-OHDA and two animals were sham-injected with vehicle. The animals were then perfused 7 days post-injection. The extent of the lesion was analysed by immunostaining for Tyrosine Hydroxylase (TH)—a key enzyme in the dopamine synthesis pathway. In the literature, the loss of TH staining is understood to be a proxy for the death of dopaminergic cells. When I analysed the brains for TH staining I found that our protocol did not lead to a lesioning. We reasoned that this could be due to two different reasons—first, that the volume of 6-OHDA was not sufficient and second, that the toxin was not given long enough to exert an effect. To address these issues, we revised our protocol by following a recently published report wherein 6-OHDA is simultaneously injected into two locations (rostral and caudal, but still unilateral) and where the animals are perfused 21 days post-injection (Girasole *et al.,* 2018).

Thus, to determine whether Penk-tdTomato^{+} cells would proliferate, migrate, and/or become neurogenic in response to a brain injury we proceeded to inject four IPNK mice with 6-OHDA (750nl, 6mg/ml) into two sites in the dorsal striatum: rostral site: 0.6 mm in front of Bregma; 2.3 mm lateral, 2.5 mm ventral; caudal site: -0.1 mm behind Bregma, 2.4 mm lateral, 2.4 mm

ventral. The experimental protocol is depicted in Figure 5.1a. The animals also received intraperitoneal injections of BrdU to label proliferating cells. These injections were given every four days and the animals were finally perfused on Day 21 for immunochemical analyses. I began by analysing the efficacy of the 6-OHDA toxin, by immunostaining for TH. Figure 5.1b shows a representative tiled image of one such section from an injected brain. The 6-OHDA leads to a massive loss of TH staining, but this loss of TH is restricted to the injected (ipsilateral) side, leaving the un-injected (contralateral) side undamaged.

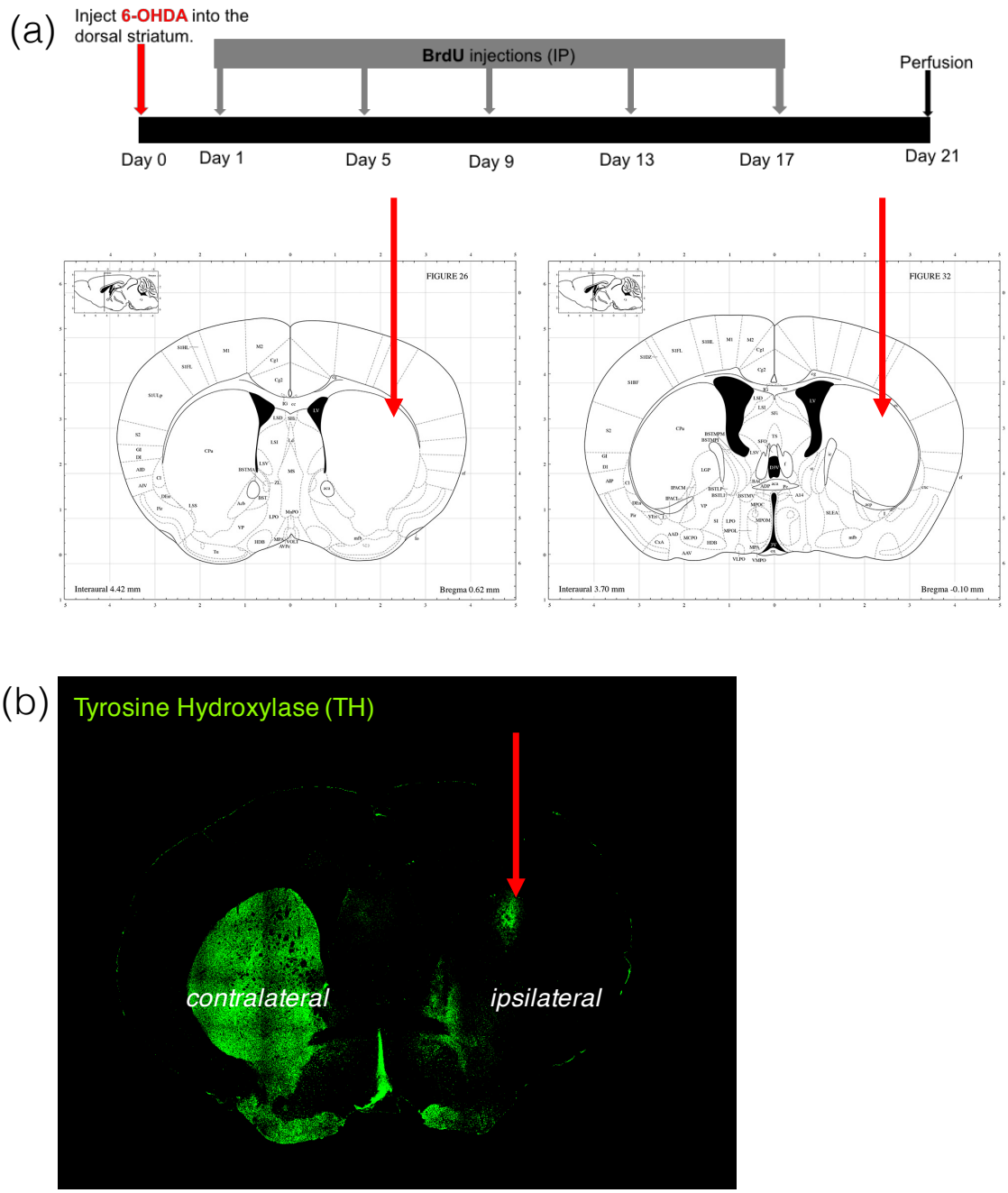


Figure 5.1. The experimental protocol for generating a partial progressive model of Parkinson's disease in IPNK mice. (a) A graphical summary of the experimental protocol. Following 6-OHDA stereotactic injection (750 nl; 6 mg/ml) into the two sites (rostral site: 0.6 mm in front of Bregma; 2.3 mm lateral, 2.5 mm ventral; caudal site: -0.1 mm behind Bregma, 2.4 mm lateral, 2.4 mm ventral) indicated by the atlas images, BrdU (200 mg/kg) is injected intraperitoneally every four days until the animal is perfused 21 days after the initial 6-OHDA injection. (Atlas images: *Paxinos and Franklin. The Mouse Brain in Stereotaxic Coordinates.)* **(b)** A tiled image of 6-OHDA-injected brain that was cryosectioned and immunostained with an antibody against TH. The red arrow indicates the location of the injection. TH staining is absent on the injected side and unaffected on the un-injected hemisphere.

5.2 The partial progressive model of Parkinson's disease causes a massive increase in cell proliferation in the affected striatum.

First, we analysed whether loss of dopaminergic cell support and subsequent striatal cell loss leads to an increase in cell proliferation in the striatal parenchyma. To perform this analysis, I began by locating the brain section with the largest TH lesion and selecting two sections (+/- 30μm) adjacent to it. For each brain, three sections were prepared for BrdU immunostaining and the number of $BrdU^+$ (and $DAPI^+$, to ensure that they were nuclei and not cell debris) cells in the striatal parenchyma were counted. Figure 5.2a shows a representative image of a brain section immunostained for BrdU. Note the massive increase in $BrdU^+$ immunostaining in the ipsilateral side compared to the contralateral side. We then quantified the number of cells in the three sections, for both sides, for all four animals. Figure 5.2b shows the data for each of the four animals, while Figure 5.2c shows the average number of $BrdU^+$ cells in the ipsilateral and contralateral sides across all four animals (ipsilateral = 901.5, contralateral = 23.5; p = 0.0003, unpaired t-test). Thus, this experiment confirms that cell proliferation is massively increased in the striatal parenchyma in response to a 6-OHDA-induced lesion. I then wanted to determine whether Penk-$tdTomato^+$ cells were contributing to this increase in cell proliferation.

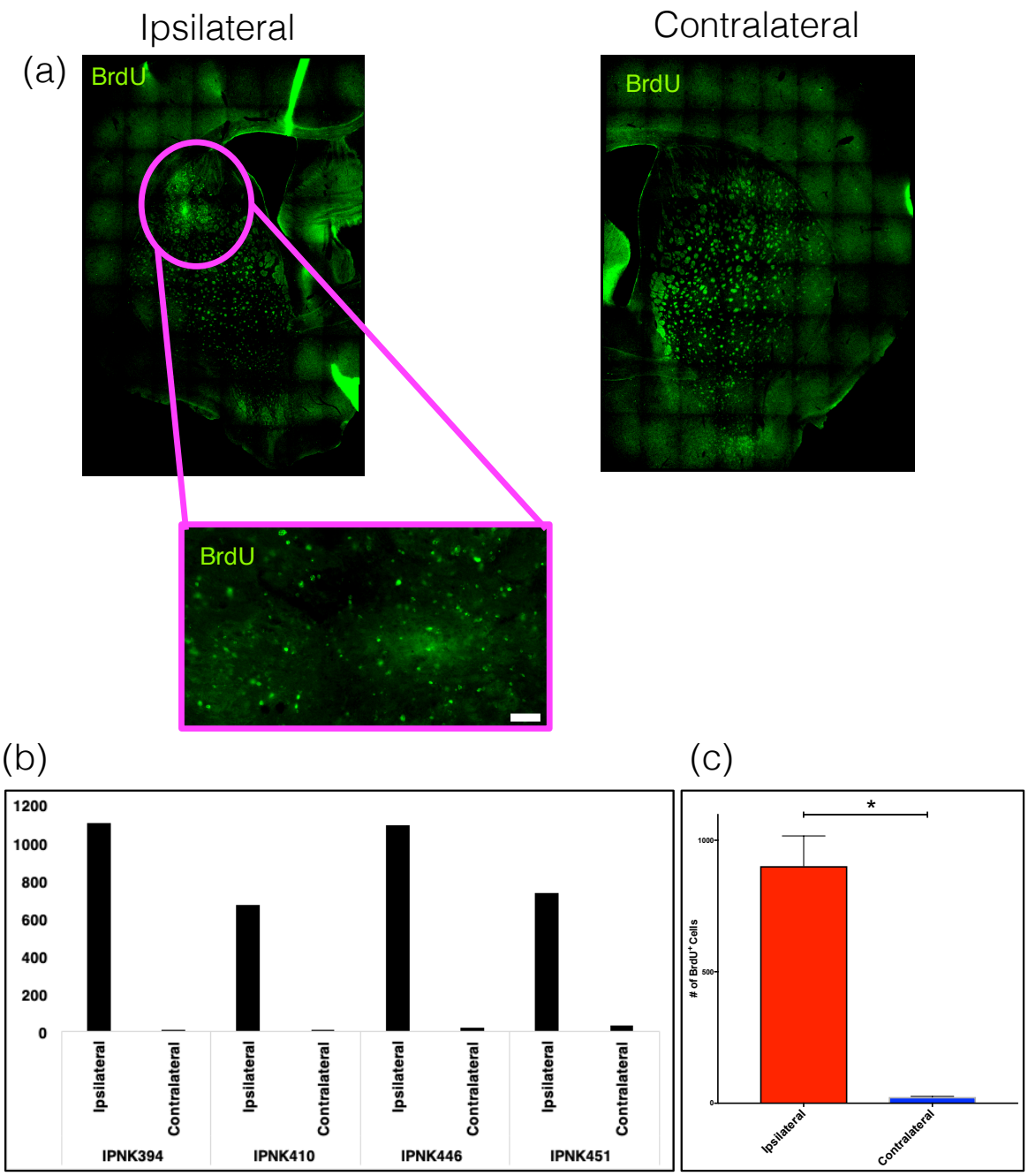


Figure 5.2. 6-OHDA injection leads a to massive and specific increase in cell proliferation. (a) A 6-OHDA injected brain cryosectioned and stained with an antibody against BrdU. The inset shows a massive increase in $BrdU^+$ cells in the ipsilateral hemisphere. **(b)** Sections were systematically analysed to determine the area with the largest lesion. One section was collected from this area and two more sections (+/- 30um) were selected for the analysis. $BrdU^+$ (conditional to the colocalisation with DAPI) were manually counted on the ipsilateral and control contralateral sides. The sum for each animal was then graphed, for the respective hemispheres. **(c)** The sums were then averaged from each biological replicate to obtain an average value for each side. 6-OHDA injections results in a massive and significant (ipsilateral = 901.5, contralateral = 23.5; p = 0.0003, unpaired t-test, N = 4) increase in BrdU staining, in comparison to the contralateral side. Scale bar = 25µm.

5.3 The partial progressive model of Parkinson's disease causes an increase in Penk-tdTomato$^+$ cell proliferation in the affected striatal parenchyma.

The method of analysis for this experiment was identical to that outlined above except for the crucial fact that here I also quantified the number of DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells in each of the sections for each of the animals. Figure 5.3a shows a representative image of a brain section immunostained for BrdU, overlaid with the Penk-tdTomato signal. Figure 5.3b shows the data for each of the four animals, while Figure 5.3c shows the average number of DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells in the ipsilateral and contralateral sides across all four animals (ipsilateral = 17.75, contralateral = 2.25; p = 0.02, unpaired t-test). Thus, there is a significant increase in the number of Penk-tdTomato$^+$ cells in the striatal parenchyma that are positively labeled for the proliferation marker BrdU in response to 6-OHDA injection, relative to the un-injected side. I note, however, that while the difference is significant, the number of DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells is approximately 50-fold lower than the total number of DAPI$^+$/BrdU$^+$ cells. This indicates that Penk-tdTomato$^+$ cells are part of a larger proliferation response to the 6-OHDA injection. The fact that we can find some DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells in the un-injected hemisphere could mean that, even in steady-state conditions, Penk-tdTomato$^+$ cells proliferate within the mature striatal parenchyma (see below for further analysis). In summary, this data indicates that even 21 days after 6-OHDA injection, Penk-tdTomato$^+$ cells continue to proliferate within the striatum as a specific response to the injury.

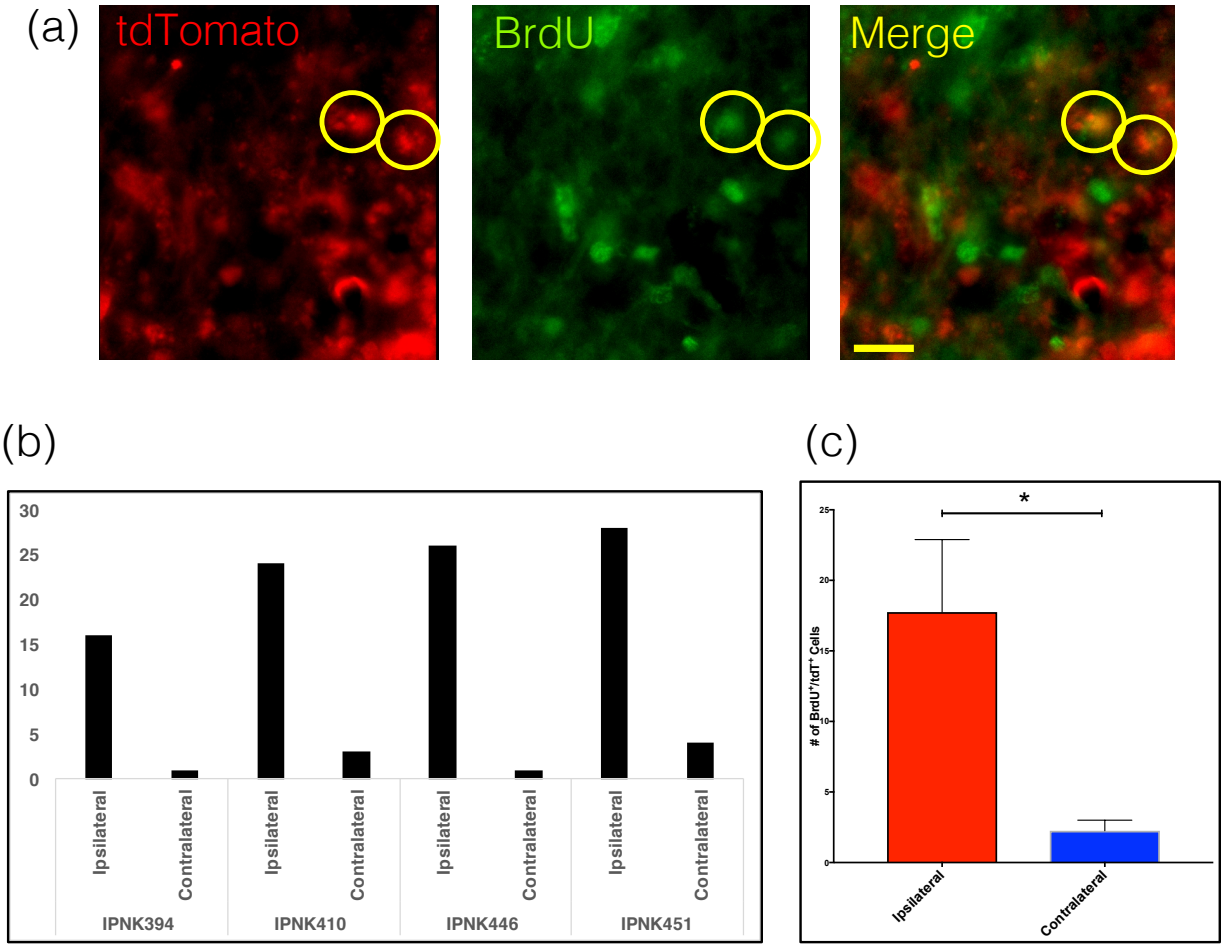


Figure 5.3. 6-OHDA injection leads to a significant increase in the proliferation of Penk-tdTomato⁺ cells. (a) A 6-OHDA injected brain cryosectioned and stained with an antibody against BrdU, with endogenous tdTomato signal also visualised. A zoomed-in image of an ipsilateral hemisphere. ***Yellow*** circles indicate two BrdU⁺/Penk-tdTomato⁺ cells. Note that most cells are not BrdU⁺/Penk-tdTomato⁺ cells. Scale bar = 20 µm. **(b)** The number of DAPI⁺/BrdU⁺/Penk-tdTomato⁺ cells in the three sections for each animal for the respective hemispheres. **(c)** The number of DAPI⁺/BrdU⁺/Penk-tdTomato⁺ cells were manually counted, summed, and averaged. 6-OHDA injections results in a massive and significant increase in DAPI⁺/BrdU⁺/Penk-tdTomato⁺ staining, in comparison to the contralateral side (ipsilateral = 17.75, contralateral = 2.25; $p = 0.02$, unpaired t-test).

5.4 The proliferating Penk-tdTomato$^+$ cells in the affected striatum do not significantly upregulate the expression of neuroblast markers.

The analysis of stroke models indicates that V-SVZ stem cells give rise to proliferating cells that also express markers of immature and developing neurons, such as NeuN and βIII-tubulin (Arvidsson et al., 2002). Thus, we also wanted to determine whether the proliferating Penk-tdTomato$^+$ cells expressed neurogenic markers. I began by assessing whether DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells also stain positively for the marker NeuN—which is widely considered to be a pan-neuronal marker (Wolf *et al.,* 1996; Arvidsson *et al.,* 2002). This analysis was technically challenging to perform, as it required the use of four different channels on the epifluorescence microscope. Figure 5.4a shows a representative image of how this analysis was done. Figure 5.4b shows the quantification in graphical form across all the animals and Figure 5.4c shows the average across all animals. It is evident from the analysis that there is not a significant upregulation of the number of DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$/NeuN$^+$ cells in response to the 6-OHDA-induced injury, relative to steady-state levels (ipsilateral = 2.25, contralateral = 1.75; $p = 0.4$, unpaired t-test). Thus, while there is an increase in the number of proliferating Penk-tdTomato$^+$ cells, the majority of these cells do not express the pan-neuronal marker NeuN. However, we reasoned that the absence of NeuN expression in Penk-tdTomato$^+$ cells may be due to this marker being expressed by neurons that are more mature than early neuroblasts.

Thus, we used a marker from our analysis that we know to be expressed in Penk-tdTomato$^+$ neuroblasts—βIII-tubulin. For each of the animals, I chose three new sections and re-performed the immunochemical analysis using antibodies against BrdU and βIII-tubulin. Figure 5.5a and b show that I see a similar increase in DAPI$^+$/BrdU$^+$ and DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells to that seen in the previous analysis. Figure 5.5c shows the number of DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$/βIII-tubulin$^+$ cells across all the animals and Figure 5.5d shows the average number

across all the animals. The results support the findings of the NeuN analysis—the proliferating Penk-tdTomato$^+$ cells do not significantly upregulate the expression of neurogenic markers.

In summary, the analysis of the partial progressive model of Parkinson's disease shows that although it does lead to a massive increase in proliferation and a significant increase in the number of proliferating Penk-tdTomato$^+$ cells in the striatal parenchyma, at 21 days post-injection those proliferating Penk-tdTomato$^+$ cells are not in the neuronal lineage. Further experiments will be required to understand exactly how DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells arrive at the affected site within the lesioned striatum.

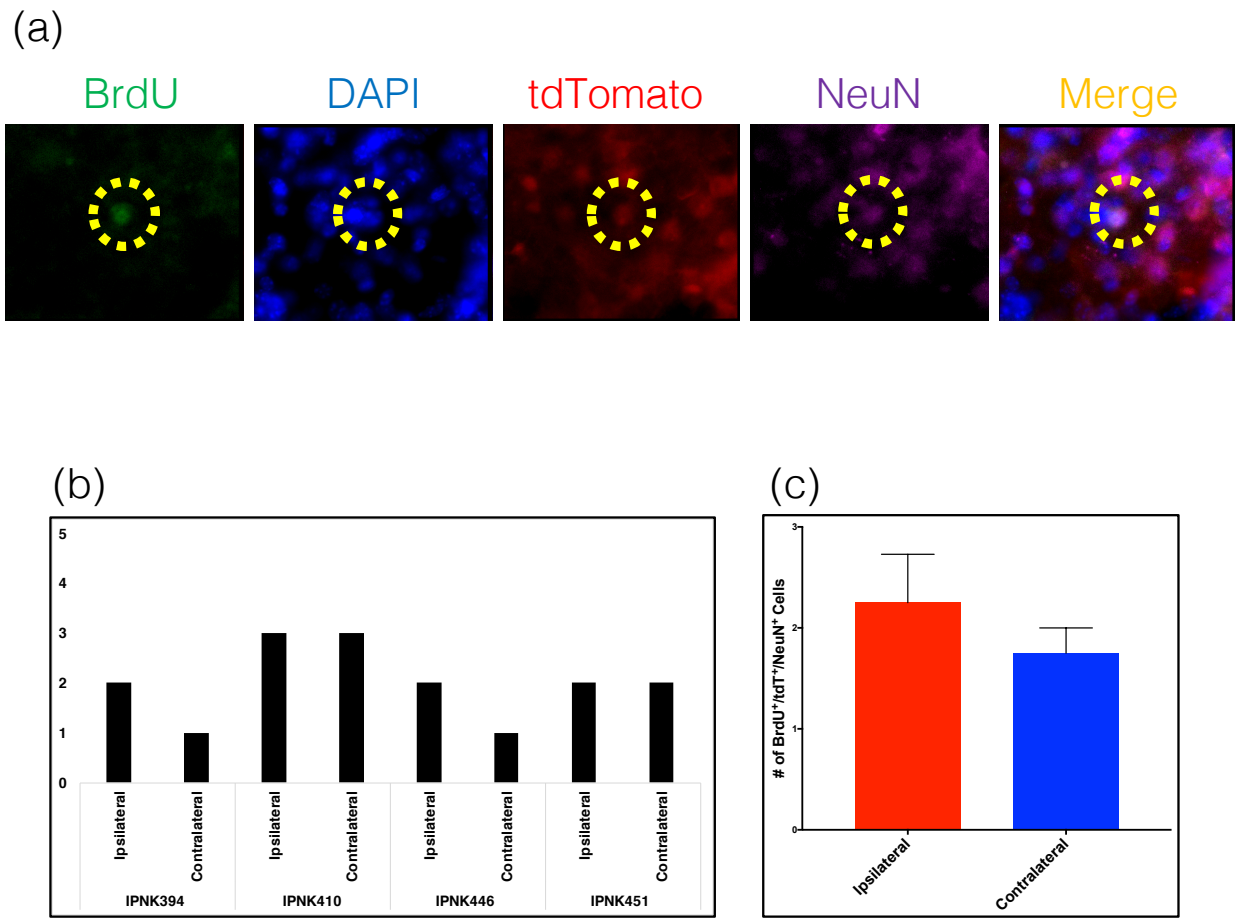


Figure 5.4. 6-OHDA injection does not lead to a significant increase in the expression of neuroblast markers in Penk-tdTomato⁺ cells 21 days post-6-OHDA. (a) A 6-OHDA injected brain cryosectioned and stained with an antibody against BrdU, NeuN, with endogenous tdTomato signal also visualised. From left to right, the panels show how the quantification was performed. **(b)** The number of DAPI⁺/BrdU⁺/Penk-tdTomato⁺/NeuN⁺ cells in the three sections for each animal for the respective hemispheres. **(c)** The number of DAPI⁺/BrdU⁺/Penk-tdTomato⁺/NeuN⁺ cells were manually counted, summed, averaged. 6-OHDA injections does not result in a significant increase in DAPI⁺/BrdU⁺/Penk-tdTomato⁺/NeuN⁺ staining, in comparison to the contralateral side (ipsilateral = 2.25, contralateral = 1.75; $p = 0.4$, unpaired t-test).

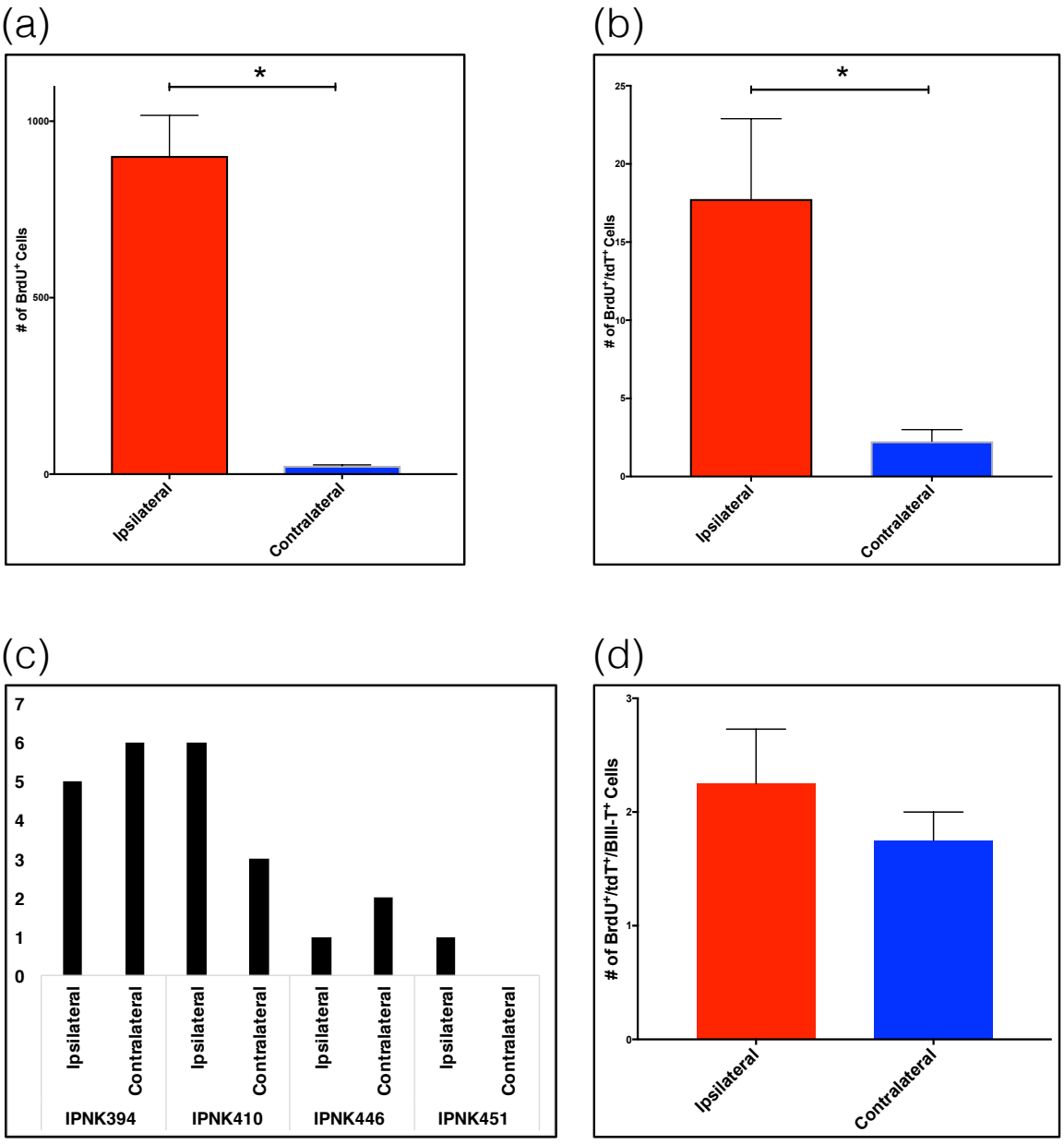


Figure 5.5. Analysis of neurogenesis using the marker βIII-tubulin. (a) A repeat of the analysis performed in Figure 5.2. There is a significant increase in BrdU+ cells in response to 6-OHDA injection (ipsilateral = 826.8, contralateral = 15.25; p = 0.00001, unpaired t-test) **(b)** A repeat of the analysis performed in Figure 5.3. There is a significant increase in BrdU+/Penk-tdTomato+ cells in response to 6-OHDA injection (ipsilateral = 26.0, contralateral = 5.5; p = 0.0005, unpaired t-test) . **(c)** The number of DAPI+/BrdU+/Penk-tdTomato+/βIII-tubulin+ cells in the three sections for each animal for the respective hemispheres. **(d)** The number of DAPI+/BrdU+/Penk-tdTomato+/βIII-tubulin+ cells were manually counted, summed, and averaged. 6-OHDA injections does not result in a significant increase in DAPI+/BrdU+/Penk-tdTomato+/βIII-tubulin+ staining, in comparison to the contralateral side (ipsilateral = 3.25, contralateral = 2.75; p = 0.8, unpaired t-test) 21 days post-injection.

5.5 Sham injection does not result in a specific increase of proliferating Penk-tdTomato$^+$ cells or of proliferating Penk-tdTomato$^+$ cells expressing neurogenic markers.

The data in the previous set of experiments clearly showed that the un-injected contralateral side is unaffected by the injection of 6-OHDA in the ipsilateral side. Furthermore, the use of the contralateral side as a within-subject control is common within the field and one of the major advantages of the 6-OHDA model of partial progressive Parkinson's disease (Smith and Heuer, 2012). However, we wanted to directly test whether the process of the stereotactic injection itself could have led to the increase in proliferating Penk-tdTomato$^+$ cells. Thus, two animals were injected with the 6-OHDA vehicle (0.02% ascorbic acid) and were taken 7 days post-injection. One animal experienced complications due to the stereotactic surgery and had to be put down. The remaining mouse was analysed in the two ways outlined above.

The results for the DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$/NeuN$^+$ and DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$/βIII-tubulin$^+$ analyses are shown in Figure 5.6. The mechanical injury that results from the insertion of the stereotactic needle does lead to a small proliferation response—as seen by the approximately two-fold increase in BrdU$^+$ cells in the ipsilateral sham injection hemisphere (Figure 5.6a). The injection does not, however, affect the number of DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells—as they are roughly equal in both hemispheres (Figure 5.6b). Neither are the number of DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$/NeuN $^+$ and DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$/βIII-tubulin$^+$ cells affected by the sham injection (Figure 5.6c). This analysis indicates that while the mechanical damage induced by the insertion of the needle does induce some cell proliferation, it does not affect the proliferation of DAPI$^+$/BrdU$^+$/Penk-tdTomato$^+$ cells.

In conclusion, the results from the analysis of the partial progressive model of Parkinson's disease in IPNK mice indicates that Penk-tdTomato$^+$ cells are able to proliferate in the striatal parenchyma.

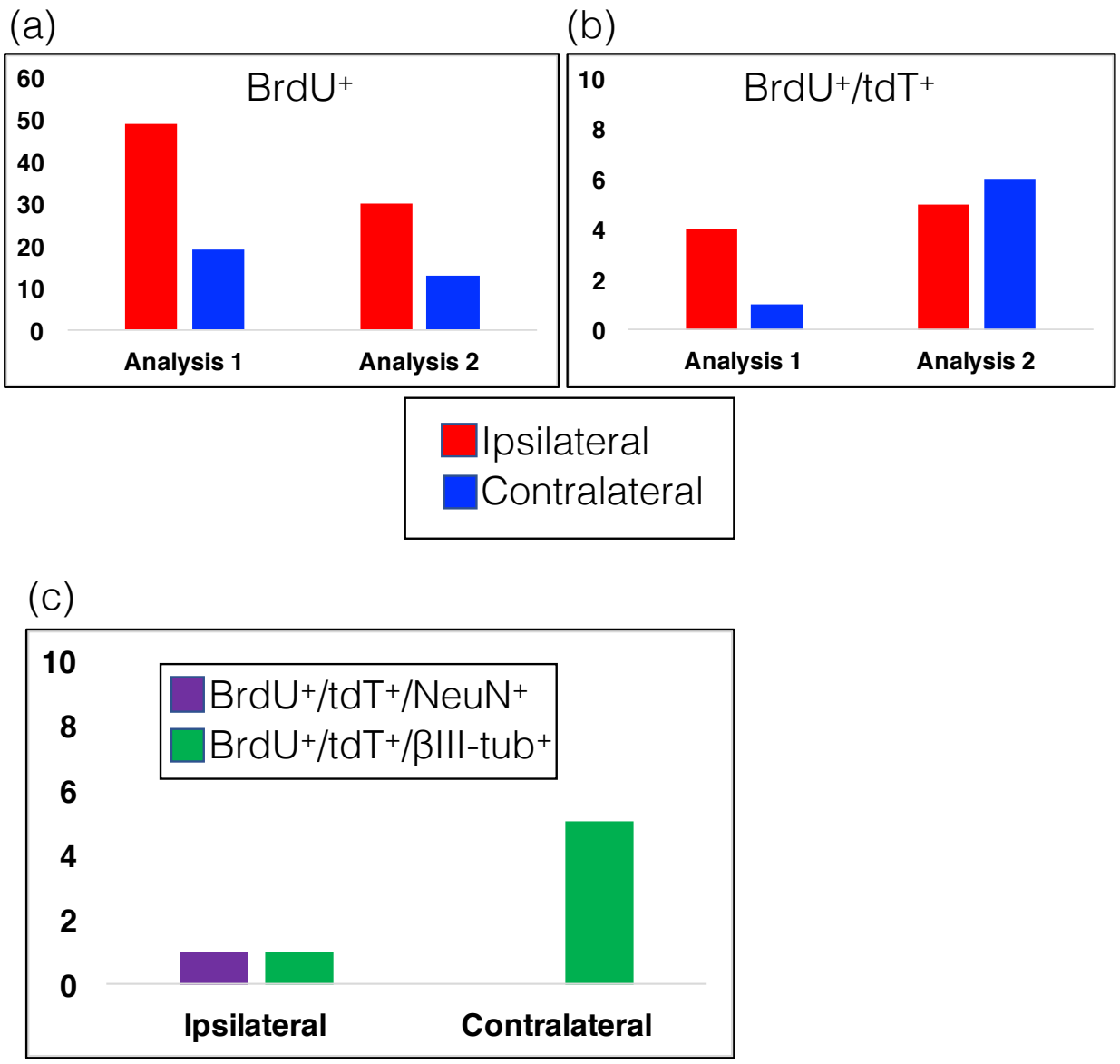


Figure 5.6. The sham injection does not affect Penk-tdTomato+ cell proliferation or neurogenesis. (a) Quantification of the number of DAPI+/BrdU+ cells in the ipsilateral and contralateral sides. **(b)** Quantification of the number of DAPI+/BrdU+/Penk-tdTomato+ cells in the ipsilateral and contralateral sides. **(c)** The number of DAPI+/BrdU+/Penk-tdTomato+/NeuN+ and DAPI+/BrdU+/Penk-tdTomato+/βIII-tubulin+ cells for the respective hemispheres.

Chapter 6 - Discussion

6.1 The relationship between Penk and tdTomato expression in IPNK mice.

The mouse model used in this thesis is the BAC-Penk-Cre$^{tg/+}$; Rosa26-tdTomato$^{tg/+}$(IPNK) strain, that consists of the BAC-Penk-Cre line crossed to the Rosa26-tdTomato$^{tg/+}$ line. The BAC-Penk-Cre line was developed in the Minichiello group and was characterized and published by Besusso *et al.* (2013) and Vatanashevanopakorn *et al.* (2017). When crossed to the reporter line Rosa26-tdTomato$^{tg/+}$, it will permanently label any cells in which the Pre-Proenkephalin (Penk) promoter was activated. The active promoter causes the expression of Cre which then removes the stop cassette, flanked by loxP sites, that otherwise prevents the expression of tdTomato from the Rosa26 locus. Once this occurs, that cell, and the cells derived from it, will continue to express the tdTomato reporter. In our study, we find that tdTomato is expressed in morphologically diverse cell types of the V-SVZ. However, when IPNK brains are immunostained to detect Penk protein expression, no such expression can be detected in the V-SVZ (Figure 3.33). In line with this finding, we also do not detect Penk transcripts in either sequencing experiments. The lack of mRNA and protein in Penk-tdTomato^{+} cells of the V-SVZ, strongly suggests that, while these cells must have expressed Penk at some point during their lifetime, they have ceased the expression of Penk.

This finding opens up several intriguing questions, the first being the temporal dynamics of the expression, *i.e.* when do they start expressing Penk and when do they stop? To attempt to address this, we sought to find an embryonic time point where we could detect no expression of tdTomato and work forwards from there. To this end, we obtained IPNK embryos that were E11.5, E12.5, and E14.5 and cryosectioned them in preparation for microscopic analysis (Fig A2). We found, however, tdTomato expression at E11.5 in the embryonic lateral ganglionic eminence—where V-SVZ stem cells are known to be derived from and the stage at which the earliest B-cells

are said to be generated (Fig A2; Fuentealba *et al.,* 2015). Based on the fact that previous work from the lab did not detect reporter expression at E11.0, the E11.5 reporter expression would indicate a possible starting point for Penk expression (Besusso *et al*., 2013). However, since the reporter used in that case (Z/EG) was different from the one used in this study we need to perform further stage analysis to confirm this. Future work will be needed to work further backwards in embryonic stages to find the earliest time point where tdTomato expression appears in IPNK embryos, once this stage is identified we can immunostain using an antibody against Penk to find the time point where they cease the expression of the protein. This work will help us to understand the temporal dynamics; but the larger question of why the cells express Penk at all, remains.

6.2 The functional significance of Penk expression in neural stem cells.

A seminal study by Alvarez-Buylla's group found that the B-cells (quiescent NSCs) of the V-SVZ originate in the embryonic brain (as early as E11.5, with a peak between E13.5-E15.5) and then remain quiescent until their activation later in life (Fuentealba *et al.,* 2015). Several lines of evidence presented in this thesis, lead us to conclude that Group1 and Group2 cells from StriatumSeq2 experiment and Group15 cells from StriatumSeq1 experiment are subsets of B-cells. If they are in fact B-cells it would also mean that they were generated as early as E11.5. As stated above, Penk-tdTomato expression can be detected in the LGE, in our hands, at E11.5 (Fig A2). Thus, it will be necessary to determine whether the expression of Penk by these B-cells works in a functional way—that is to say, does the expression of Penk during embryonic development define the specific set of neurons originating from these B-cells in the adult? One possibility, among others, would be lineage-restriction—so that Penk-tdTomato expressing B-cells preferentially generate D2 MSNs over D1 MSNs, since Penk expression is generally only seen in D2 MSNs in the adult.

Fuentealba *et al.* found that B-cells generated as early E11.5 were restricted in the type of olfactory bulb interneurons they generated in the adult mouse brain. We briefly discussed in the ***Results*** our finding that Penk-tdTomato^{+} cells colocalise with a wide-variety of olfactory bulb interneurons markers—Tyrosine Hydroxylase, Somatostatin, Parvalbumin, Calbindin, Calretinin (*data not shown*). This finding would seem to indicate however that Penk more generally labels a wide subset of B-cells—though further work will be required to fully understand the dynamics of Penk-tdTomato^{+} cell-mediate olfactory bulb neurogenesis. Future work using an inducible CRE system, and some of the novel markers identified herein, to perform lineage-tracing will allow us to determine more specifically what cell types are generated from Penk-tdTomato^{+} cells from the earliest embryonic times points.

It will be important to also understand why Penk-tdTomato^{+} NSCs cease the expression of Penk. The simplest reasoning would be that this is a consequence of entering a quiescent state—the cells stop expressing Penk, along with other proteins, to conserve energy and remain quiescent. But there could be other reasons that are functionally interesting. Work done in mouse embryos finds that progenitor cell proliferation in the developing SVZ relies on the permissive autocrine signaling of the endocannabinoid system (Mulder *et al.,* 2008). In an analogous manner, it could mean that Penk production and release by NSCs could be a mechanism by which control over differentiation and proliferation is exerted. Work done on adult zebra finches found that systemic administration of the opioid receptor antagonist, naloxone, led to a significant increase in cell proliferation (Khurshid *et al.,* 2010). However, work on mouse embryonic stem cells, found that the administration of Penk analogues diverts cells away from a state of self-renewal to a state of differentiation (Kim *et al.,* 2006). While other possibilities exist, the most parsimonious explanation could be that NSCs cease the expression and release of Penk in order to remain in a quiescent state. Testing this hypothesis would require forcing the expression of Penk (using

viral vectors) in quiescent NSCs in IPNK mouse brains and then analysing changes in differentiation and proliferation at some later time point.

6.3 Group1 cells likely represent a subset of B-cells.

The existence of B-cells was first shown in electron micrographs of the V-SVZ taken by Fiona Doestch and Arturo Alvarez-Buylla at the University of California San Francisco (Doetsch *et al.,* 1997). This and the set of functional experiments published two years later, showed definitively that the B-cells of the V-SVZ were the quiescent NSCs of the ventricular germinal zone (Doestch *et al.,* 1999). Although many studies have been performed on these cells, analysis has been hampered by the lack of a specific genetic or protein marker. Like the initial papers, many have relied on the co-expression of many different proteins, such as GFAP, GLAST, and CD133. This presents many problems as GFAP and GLAST are astrocytic markers, while CD133 is a marker of ependymal cells. As a result of this, many controversies have arisen within the field. Jonas Frisen's group has published reports stating that $CD133^+$ and $Foxj1^+$ ependymal cells have stem cell characteristics and become neurogenic upon stroke (Johansson *et al.,* 1999; Carlen *et al.,* 2009). However, Fiona Doetsch's group refutes this finding by stating that CD133 is not a specific marker (Chaker *et al.,* 2016). It must be noted, though, that the markers used by her group are also not unique for B-cells. A comprehensive RNAseq experiment recently performed on ependymal cells has indeed found that $CD133^+/Foxj1^+$ do not have stem cell function (Shah *et al.,* 2018). This highlights the need to find specific gene and protein markers for the quiescent NSCs of the V-SVZ so that they can be closely studied and unequivocally analysed.

Several groups have hypothesized, based on their functional analyses of V-SVZ quiescent NSCs, what unique proteins B-cells would likely express. Elegant studies by Samuel Weiss's group showed that the circulating hormone prolactin positively regulates neurogenesis at the V-SVZ (Shingo *et al.,* 2003; Mak *et al.,* 2008). This indicates that the stem cells of the V-SVZ must

be receptive to this signal, in other words—they must express the prolactin receptor (Prlr). This hypothesis has also been put forward by Alvarez-Buylla in a recent study of stem cell division dynamics (Obernier *et al.,* 2018). Thus, B-cells should specifically upregulate Prlr.

Another clue emerges from the finding that the V-SVZ is strongly innervated by serotonergic projections from the Raphe nuclei (Tong *et al.,* 2014). In this study, Alvarez-Buylla's group also finds that the B-cells of the V-SVZ express a serotonin receptor—namely the 5-HT_{2C} receptor. Thus, they show using electron micrographs and simultaneous immunofluorescence and electrophysiological analyses that B-cells must express the 5-HT_{2C} receptor (Tong *et al.,* 2014).

In a recent report by the same group, analysing stem cell division dynamics, they find, by examining electron micrographs, that B-cells have high densities of ribosomes within their cell bodies (Obernier *et al.,* 2018). Another recent study of aged quiescent NSCs of the V-SVZ by Anne Brunet's group found that they contained large lysosomes and had a transcriptomic profile showing upregulation of processes related to lysosomes (Leeman *et al.,* 2018).

As shown throughout the ***Results*** chapters, Group1 cells specifically upregulate the expression of Prlr and the 5-HT_{2C} receptor (Figure 3.27). The Gene Ontology analysis finds that Group1 cells specifically upregulate processes related to ribosomal and lysosomal pathways (Figure 3.30). Group1 also specifically upregulates the expression of *Sfrp1*, a gene identified by Fiona Doetsch's group to be specifically expressed by the quiescent NSCs of the V-SVZ, in a large RNAseq study (Mizrak *et al.,* 2019). Along with the morphological similarities of the Penk-tdTomato^{+} cells to electron micrographs of B-cells, we believe that this provides strong lines of evidence in support of the hypothesis that Group1 cells are a subset of B-cells. Furthermore, our RNAseq and validation analyses reveal novel and specific markers for these B-cells, thus opening the door for future lineage-tracing and *in vivo* manipulations.

6.4 Novel cell-surface markers for B-cells and their importance for the further characterization of B-cells.

Above, we provide an argument for the classification of Group1 cells as a specific subset of B-cells based on the specific expression of postulated genes, pathways, and morphologies. We now discuss novel cell-surface markers that can be used to study these cells and avoid controversies arising from different methodological approaches, as described in the previous section. The specific and significant expression of Prlr can be used to control and study the B-cells, as *Prlr*-CRE and *Prlr*$^{lox/lox}$ mouse tools already exist (Brown *et al.,* 2016; Kokay *et al.,* 2018). Furthermore, many small molecule antagonists of Prlr exist, thus allowing for even easier manipulation of Group1 cells at the V-SVZ (Bernichtein *et al.,* 2003). Our analysis, however, finds two more wholly novel cell-surface markers for B-cells. The first and most intriguing one is *Kcnj13*, a gene that encodes for the inwardly-rectifying potassium channel KiR7.1. Our functional validation, using immunochemistry and FACS, shows that this antibody can be used to localize cells at the V-SVZ and isolate cells which recapitulate the gene expression signature of Group1 cells. This channel has not been extensively studied in the brain. Current reports only describe its functions in the retinal pigment epithelium, where it is required for the normal physiological function of the retina (Shahi *et al.,* 2017). Interestingly, however, KiR7.1 has recently been reported to control cell synthesis, albeit outside of the central nervous system (Yin *et al.,* 2018). In particular, it is reported that KiR7.1 is a “critical modulator” for tracheal tubulogenesis through the modulation of AKT, a key regulator of actin. Though not direct evidence, it does indicate that KiR7.1 has the potential to control neurogenesis through similar mechanisms. Previously, a group attempted to determine whether neuroblasts of the V-SVZ, Dcx^{+} and Nestin^{+}, expressed different types of potassium channels using microarray analysis (Pruss *et al.,* 2011). Co-incidentally, the microarray they used did not include the gene for KiR7.1—*Kcnj13*. Nonetheless, they do report the KiR7.1 antibody (custom-made) staining at the walls of the lateral ventricle, but not in the Dcx^{+}

cells of the V-SVZ—confirming our findings. In a key finding, they demonstrate that KiR7.1 protein is detected (by Western blot) in cells of "early" neurosphere cultures that were GFAP$^+$, but not in "late" neurosphere cultures that were Dcx$^+$. This critical finding is in line with our finding that *Kcnj13* expression is restricted to Group1, and is not found in Group3 (which significantly upregulated Dcx expression). Furthermore, we show using antibody staining, that KiR7.1 and GFAP also colocalise in our hands at the V-SVZ in Penk-tdTomato$^+$ cells (Figure 4.3). We believe that this provides strong support for our finding that *Kcnj13* is a specific gene marker for B-cells at the V-SVZ. As described above, *Kcnj13^{flox}* mice are already in existence (Yin *et al.,* 2018). Thus, this model can be used alongside appropriate CRE lines to dissect the role of B-cells at the V-SVZ. The fact that KiR7.1 is a channel also presents other, easier, less invasive means of manipulation and control. A specific small molecule blocker of KiR7.1 exists—VU590 (Crankshaw *et al.,* 2017). In the future, it will be intriguing to analyse the effects of intracerebroventricular infusions of the VU590, and thus KiR7.1 blockade, on the stem cell dynamics of the V-SVZ.

A recent report by the lab of Magdalena Götz, reveals potential mechanistic reasons for why B-cells may specifically upregulate the expression of an ion channel (Petrik *et al.,* 2018). Their study shows that the epithelial sodium channel (ENaC) regulates stem cell proliferation at the V-SVZ in response to CSF fluid flow. We hypothesise that KiR7.1 may serve a similar role in B-cells—possibly monitoring CSF fluid flow and ionic content and regulating the activity of the stem cells in response to this.

Another novel finding in our RNAseq dataset was the expression of the Folate Receptor (*Folr1*) by Group1 cells. By extension, we hypothesise that B-cells specifically upregulate the expression of *Folr1*. The expression of Folr1 by neural progenitor cells is, in itself, not novel as other groups have shown it to be specifically expressed by midbrain dopamine neural progenitor cells during development and that it controls stem cell pool size during development (Gennet *et al.,* 2016; Mohanty *et al.,* 2016). However, we are the first to report its expression by the

prospective quiescent NSCs of the adult V-SVZ. This is again an attractive target, as Folr1 has been extensively studied in the cancer field. As such, inhibitors of Folr1 such as farletuzumab already exist, can be used to analyse the function of the receptor in the V-SVZ (Lin *et al.,* 2013).

Thus, *Kcnj13* and *Folr1* represent novel markers of putative B-cells at the V-SVZ and provide easy genetic and pharmacological means of intervention. Prlr and 5-HT_{2C}, as discussed above, are also attractive means to manipulate B-cells. In the future, we will use these tools in tandem to probe, genetically and pharmacologically, the role of these receptors, and the cells expressing these receptors, in V-SVZ neurogenesis.

6.5 The developmental arc of the three clusters.

Pseudotime analysis revealed that Group1 and Group3 cells were at the opposite ends of a pseudotemporal path. Gene expression and cell cycle analysis indicated Group3 cells were neuroblasts and thus further forward in the developmental arc than Group1 cells. The analysis also showed that Group1 preceded Group2 in the developmental arc. This leads to another key finding in our study. Group2 specifically upregulates the expression of *GLAST* and *Connexin30* and *Connexin43* (Figure 3.28). This is interesting because many groups use the expression of GLAST as a means of identifying the most quiescent NSCs of the V-SVZ. However, our analysis indicates that the *Kcnj13*$^+$, *Prlr*$^+$, *Folr1*$^+$, *Sfrp1*$^+$, *5HT$_{2C}$ receptor*$^+$ Group1 cluster precedes this quiescent NSC group, suggesting that Group1 may represent an even more quiescent pool of NSCs or B-cells. Although the evidence presented here does not directly address this, the genetic tools described in the previous sections can be used to address this in the future. We should be able to see cells that are positive for Group1 markers sequentially expressing Group2 markers, but not the other way around. This would provide evidence for Group1 developmentally preceding Group2 *in vivo*. However, another possibility is that Group1 and Group2 are just two separate

pools of B-cells that may give rise to different cell types when activated. Again, lineage-tracing studies will need to be performed to address this hypothesis.

6.6 Classifying the three clusters.

Above, I discussed, at length, the Group1 markers. Here, I will discuss the markers we used to classify the other two groups. When we performed StriatumSeq1 we did extensive analysis of the gene expression profiles to rule out the possibilities that the sequenced cells were medium spiny neurons (*Ctip2*$^-$*, Nkx2.1*$^-$*, Mash1*$^-$), interneurons (*PV*$^-$*, NPY*$^-$*, SST*$^-$*, NOS*$^-$*, Calb2*$^-$*,* and *ChAT*$^-$), astrocytes (*Aldh1L1*$^-$ and *S100β*$^-$), oligodendrocytes (*Sox10*$^-$*, NG2*$^-$*, PDGF-α*$^-$*, Olig2*$^-$) and ependymal cells *(CD133*$^-$*,* and *Foxj1*$^-$). We analysed the expression profile of Group1/Group15 to find that while they expressed some previously-identified stem cell genes such as *Sfrp1*, *Vcam1*, and *Igfbp2*, they also expressed the novel markers we identified. The expression of the novel markers and the Pseudotime analysis led us to conclude that Group1 was a subset of B-cells. The SCDE analysis identified *GLAST*, *Connexin30*, and *Connexin43* as being differentially upregulated in Group2 cells. As described previously, the upregulation of GLAST along with the gap junction proteins indicated to us that they were quiescent NSCs. While the Pseudotime analysis indicated that the they were likely to be more advanced on the pseudo-developmental path than Group1 cells.

Finally, we were able to confidently label Group3/Group26 cells as neuroblast due to the expression of many well-characterised genes such as *Sox11*, *Dcx*, and *βIII-tubulin*. This shows that our recombinant reporter allows us to capture Penk-tdTomato$^+$ stem cells that are at different developmental time points within the V-SVZ. In the future, it will be interesting to follow single Penk-tdTomato$^+$/Group1-marker$^+$ cells to determine whether in fact this developmental arc is maintained through the three groups identified in our sequencing experiments. It will also reveal the other transition states that must exist—most importantly, the activated NSC stage. Our

experiments are not able to pick up a cluster that could be classified as activated NSCs, for example by the specific upregulation of the widely-accepted marker Epidermal Growth Factor Receptor (*EGFR*). This could be due to the fact that cells at this activated "transition" state exist for a short period of time and are less likely to be represented in a population of sorted cells. It may also be the case that activated NSCs are slightly more sensitive to mechanical damage and thus become under-represented in the final sorted population of Penk-tdTomato^{+} cells.

6.7 The isolation and sequencing of the Group2 cluster in StriatumSeq2.

The analysis of StriatumSeq1 produced two clusters (Group15 and Group26), while the StriatumSeq2 produced three clusters (Group1, Group2, and Group3). When the cells from both experiments were clustered and projected onto a PCA plot, we found that the Group15 and Group1 cells overlap and that Group26 and Group3 cells overlap. Interestingly, however, Group2 remains unique (Figure 3.20). Thus, it is pertinent to ask why it is that Group2 cells are only identified in StriatumSeq2 and not in StriatumSeq1. We think that there could be several different reasons. The first and most likely, is the difference in ages between the animals used in the experiments. In StriatumSeq1 the animals were approximately 4 months of age, while in StriatumSeq2 they were 2 months of age. A rigorous and recently published review of neurogenesis experiments in rodents and humans illustrates the severe importance of age selection when performing experiments on cells involved in neurogenesis (Snyder, 2019). Though not quantifiable, there is a visually significant decline in the rate of neurogenesis between P60 and P120—when a meta-analysis of mouse neurogenesis studies is performed (Snyder, 2019). Since StriatumSeq1 was performed on ~P120 mice and StriatumSeq2 was performed on ~P60 mice, there may have been more cells actively proliferating and differentiating in the V-SVZ when it was isolated for sequencing in StriatumSeq2. The Pseudotime analysis performed in Figure 3.23 and 3.24 indicates that Group2 cells are more advanced in the pseudotemporal path than

Group1 cells, in other words they may be more differentiated Group1 cells. Furthermore, they seem to exist as a transition state between Group1/Group15 and Group3/Group26. Thus, when the rates of neurogenesis decline at P120 there may be fewer cells at this stage than when neurogenesis is occurring at higher levels in P60. To address this directly, we will need to repeat the sequencing experiment at several different time points to assess how the proportion Penk-tdTomato^{+} cells at the V-SVZ changes with respect to age.

Another factor that needs to be considered when making comparisons between the two experiments is that only males were sequenced in StriatumSeq1 and only females were sequenced in StriatumSeq2. Although, there is a great degree of overlap between the clusters it is something that must be considered when analysing the data in combination because a recent publication found differences in the amount of oligodendrocyte progenitors in the V-SVZs of male and female mice (Mizrak *et al.,* 2019). Although a definitive experiment analysing sex differences in neurogenesis in the V-SVZ has not been performed, a review by Brunet states “females tend to exhibit increased stem cell self-renewal, regeneration potential, and in some cases, proliferation” (Dulken and Brunet, 2015). This is line with our finding that Group2 cells are found only when females are sequenced. Females may be undergoing increased neurogenesis, thus increasing the chance of capturing a small number of cells in this transition state more easily.

The last—and in our opinion, the least likely—reason for the difference in sequenced clusters is the inclusion of cells with a microglial signature during the sequencing stage of StriatumSeq1. It may be that Group2 cells are more prone to stress, and thus phagoptosis (discussed below), than Group1 and Group3 cells and were then subsequently excluded in the analysis of StriatumSeq1. However, when we analyse the Group320 cells we do not find any evidence of cells with NSC signatures (*data not shown*). Thus, it is unlikely that Group2 cells were excluded in StriatumSeq1 due to microglial signature presentation.

Future sequencing experiments will need to be performed on IPNK mice to understand how age and sex affect the heterogeneous Penk-tdTomato^{+} cells at the V-SVZ.

6.8 Comparisons to previously published reports of sequenced V-SVZ cells.

During the process of classification, we used previously published reports to aid the analysis. Both studies, Llorens-Bobadilla *et al.* (2015) and Dulken *et al.* (2017), used the same deep sequencing method as our study (Smart-seq2) and similar numbers of cells to perform the bioinformatics analysis. They also used animals between the ages of 2-3 months, where we used 4-month-old animals for StriatumSeq1 and 2-month-old animals for Striatum Seq2. Here, I discuss how this was done and how it gave us further evidence that our Group1 cells were a novel group of cells at the V-SVZ. We began by comparing our dataset to the first complete and in-depth sequencing analysis of V-SVZ cells performed by Ana Martin-Villalba's group (Llorens-Bobadilla *et al.,* 2015). This study did not use genetically-encoded reporters for cell isolation, instead they relied on the use of antibodies to capture quiescent NSCs that were dually GLAST^{+} and CD133^{+}, or PSA-NCAM^{+} to isolate neuroblasts. This already creates a key point of difference between our analyses and theirs. We find that our Group2 cells are GLAST^{+}, not our Group1 cells (which precede Group2 cells in the Pseudotime analysis). This alone explains why they were not able to detect cells expressing our Group1 markers. To compare our datasets more broadly, we used a classification table generated by them to see how our clusters would be classified according to their system. They use their transcriptomic data from the different cell types to assign groups of genes signatures to the cell types (or classes) upregulating those genes. When we performed this comparison, we found that our Group1 and Group2 cells generally express genes belonging to Class II and III—which they state defines quiescent NSCs (Figure 3.31). The Group3 cluster exhibits gene expression characteristic of their Class V and VII, which define neuroblasts. Thus, this comparison gave us further evidence that Group1 and Group2 were quiescent NSCs and that

Group3 cells were neuroblasts. The use of GLAST to isolate the most quiescent NSCs in their dataset is key difference between their analysis and ours and shows why our Group1 cell cluster is novel.

We then performed a more thorough comparison to a more recently published report of sequenced V-SVZ cells by Anne Brunet's group (Dulken *et al.,* 2017). In this study, the group used a GFAP-GFP transgenic line along with an antibody against CD133 to isolate quiescent NSCs and an antibody against EGFR to isolate activated NSCs. In order to compare our dataset to theirs, we obtained the raw FASTQ files and performed the bioinformatics analysis exactly as we had performed on our own. This allowed us to project our cells in combination to theirs on the same PCA plot and thus determine if there was any overlap between our clusters and theirs. We found that Group2 cells overlapped with their quiescent NSC cluster—similar to the findings of the Llorens-Bobadilla *et al.* (2015) comparison (Figure 3.32a & b). We also found that Group3 cells overlapped with their activated NSC and neural progenitor cell clusters. Most interestingly, we found that our Group1/Group15 cluster was completely different from anything they had been able to isolate, and also preceded all their clusters in the Pseudotime analysis (Figure 3.32c). Thus, these comparisons provided us support for the ideas that Penk-tdTomato labels stem cells of the V-SVZ and that it is able to label a completely unique subset of quiescent NSCs. This comparison illustrates the utility of our recombinant reporter line. Importantly, since we do not rely on antibodies we are able to capture early quiescent NSCs that other groups have missed due to the reliance on published antibody markers and/or on reporter lines that miss the earliest NSCs.

6.9 Are Group1 cells choroid plexus cells?

The choroid plexus is a structure that exists in every ventricle, including the lateral ventricle, where its main function is the generation of cerebrospinal fluid, among other things (Lun *et al.,* 2015). This epithelial cell structure extends into the ventricular space and lies adjacent to

the ventricular walls and the V-SVZ. Recent work has also shown how released factors, such as NT-3, from the choroid plexus regulate the stem cells of the V-SVZ (Delgado *et al.,* 2014; Silva-Vargas *et al.,* 2016). As stated in this thesis, our bioinformatics analysis indicates that among the most specific markers of Group1 cells are the genes *Kcnj13*, *Prlr*, and *Folr1*. In the literature, there is evidence for these markers being expressed in choroid plexus cells. *In situ* hybridisation studies have found *Prlr* to be highly expressed in the choroid plexus cells of the lateral ventricle (Brown *et al.,* 2010). Early analyses of choroid plexus tumours found that they expressed the folate receptor, and has since been identified in many other tumours (Patrick *et al.,* 1997). Two recent studies used DropSeq to sequence tens of thousands of cells from different parts of the brain (Saunders *et al.,* 2018; Zeisel *et al.,* 2018). Although not discussed within their manuscripts, when we analysed their data on the web tools the two groups developed, we found that *Kcnj13* is classified as being enriched in a group they classify as choroid plexus cells—though the extent of that enrichment is difficult to gauge. *Kcnj13* is also found in many other cell types in their analysis, but not in any groups of cells they label as stem cells (Saunders *et al.,* 2018; Zeisel *et al.,* 2018). However, a slightly earlier transcriptome analysis of the lateral ventricular choroid plexus by the Doetsch group did not report the upregulation of either *Kcnj13*, *Prlr*, or *Folr1*—further complicating the matter (Silva-Vargas *et al.,* 2016). An even earlier report states that the folate receptor was detected in transit amplifying cells (C-cells), as well as in ependymal cells and choroid plexus cells (Lee *et al.,* 2012). Thus, we must consider if our Group1 cells, which significantly upregulate the expression of these three markers, are also choroid plexus cells. We believe there are several lines of evidence to indicate that, while the lateral ventricle choroid plexus may indeed express those markers, Group1 cells are quiescent NSCs of the V-SVZ and not epithelial cells of the choroid plexus. First, the way in which we perform the sub-dissection of the striatum is designed to eliminate all tissues that are not directly attached to the striatum. In this way, we make sure to collect only the striatal parenchyma and the lateral wall of the lateral ventricle—leaving behind

the medial wall, septum, and likely the choroid plexus. However, we cannot exclude the small possibility that the choroid plexus was included in some preparations because it somehow remained attached to the striatal tissue. In the future, we can perform further microdissections to collect only the V-SVZ prior to sequencing to ensure that we truly eliminate the choroid plexus from our analyses. Second, the immunochemical analyses we performed using KiR7.1 and Prlr both in coronal and wholemount sections of the lateral wall of the lateral ventricle indicate unequivocally that these markers are present at the V-SVZ—and thus that Group1 cells are part of the germinal niche (Figures 4.2, 4.3, 4.4, 4.6). In the future, we will look to perform fluorescence *in situ* hybridisation to more closely analyse the transcript localisation of the three markers. We believe this analysis will support our immunochemical analyses, and show that cells positive for Group1 markers are present at the V-SVZ. Third, our immunochemical localisation of the Group1 marker Prlr at the V-SVZ is strongly supported by two other separate analyses, which also report the strong localisation of this protein at the V-SVZ (Shingo *et al.* 2003; Mak *et al.,* 2008). Finally, Pruss *et al.* (2011) generated neurospheres from dissected SVZ tissue and found that KiR7.1 was expressed in "early" $GFAP^+$ neurospheres. This further supports the idea that "early" neurogenic cells of the V-SVZ express the Group1 marker *Kcnj13*. Thus, we believe that while choroid plexus cells do express some of the Group1 markers, Group1 cells themselves are not choroid plexus cells.

6.10 Microglial signature in StriatumSeq1.

Analysis of the StriatumSeq1 found the large Group320 cluster displaying a strong microglial signature. This was determined by their upregulation of four key microglial genes—*P2ry12, Csf1r, Tmem119*, and *Cx3cr1* (Figure 3.12; Butovsky *et al.,* 2014; Holtman *et al.,* 2015). Analysis of IPNK brain sections presented throughout the thesis definitively indicates that Penk-tdTomato, in the striatum and V-SVZ, is not expressed in cells with the morphological

characteristics of microglia. Why, then, do we detect the expression of microglial markers in sequenced striatal cells and why is it not detected in the NSCs? The reason for this could be that the mature, differentiated MSNs are more sensitive to the mechanical damage that results from the dissection, trituration, and sorting procedure (Saunders *et al.,* 2018). They may also be more prone to damage because they require more cell-support than V-SVZ cells—meaning that their health is more dependent on the activity of neighbouring cells and long-distance projections (Saunders *et al.,* 2019). Thus, when they are isolated they may become more prone to damage than their stem cell counterparts. Recent work has shown that "stressed-but-viable" neurons can be phagocytosed by microglia, in a process termed "phagoptosis" (Brown and Neher, 2014). What this could mean is that the cells derived from IPNK brains become stressed during the isolation procedure and are phagocytosed by microglia as a result. The cell+microglia body would still retain tdTomato fluorescence and would thus be sorted as "target cells" into the 96-well plate for further sequencing.

The initial clustering of StriatumSeq1 resulted in three clusters Group15, Group26, and Group320 (Figure 3.11). When the Group320 cells were removed from the analysis (due to their expression of microglial genes) the re-clustering resulted in the same Group15 and Group26 cells (Figure 3.13), indicating that they do not display a microglial signature, nor is their clustering dependent on Group320. While we were confident that the two clusters were not affected, we wanted to develop a method to eliminate microglial cells from the outset for future experiments. To achieve this, relied on the finding that all macrophages express the cell-surface protein CD11b (Zhang *et al.,* 2008). We reasoned that if were able to exclude $CD11b^{+}$ cells during FACS, we could obtain a purified neuronal population (Figure 3.19a). We confirmed the purification by means of high-throughput gene expression analysis (Figure 3.20). All future analysis was performed on $CD11b^{-}$ cells.

6.11 "M-Peak" cells display the signature of mature striatal neurons.

During the experiments described above, where we sought to validate the elimination of the microglial signature, we also found an interesting population of cells that we termed the "M-Peak" cells. In Figure 3.19b we see that the LIVE/DEAD fluorescence plot is trimodal. In our other experiments we collected both of the two left-most peaks, while discarding the right peak. However, when we back-gated to locate where the majority of the Penk-tdTomato^{+}/CD11b^{-} lie, we found that a large majority of them lie in the middle peak of the LIVE/DEAD plot. When we analysed the gene expression of these cells, shown in Figure 3.20, we found that they display the molecular characteristics of mature MSNs. Thus, this gave us a FACS-based method of isolating putative MSNs from whole-striatum cell suspensions. A question that arises from this is why putative MSNs would lie in the middle peak of the LIVE/DEAD plot. A simple answer would be that the plasma membrane composition of putative MSNs and stem cells differ, wherein the MSN membranes are slightly more permeable and thus allow for the diffusion of LIVE/DEAD dye into the cytoplasm. Regardless, this is a useful tool to isolate live MSNs from the same preparation as NSCs.

6.12 The use of whole-striatum lysates in place of micro-dissected V-SVZ tissues and alternative sequencing methods.

The tissue handling and cell-suspension preparation for both StriatumSeq1 and StriatumSeq2 were identical. We sub-dissected whole striata from both hemispheres and manually chopped them before trituration with a glass pipette. We designed the experiment in this way so that we would be able to compare, bioinformatically, mature MSNs and the stem cells of V-SVZ. This is a wholly unique approach and it is facilitated by our reporter line, where Penk-tdTomato is expressed by the mature MSNs (that actively express Penk protein) and by the cells of the V-SVZ (that no longer express Penk protein). The earliest attempts to sequence the stem

cells of the V-SVZ (using suites of combinatorial antibodies) by the labs of Magdalena Gotz and Fiona Doetsch relied on first making sections of the striatum and then micro-dissecting the V-SVZ (Beckervordersandforth *et al.,* 2010; Codega *et al.,* 2014). The studies that we directly compared our datasets against in this thesis also used the same dissection approach used by the Gotz and Doetsch laboratories and the same RNAseq method as our experiments—Smart-seq2 (Figures 3.31 and 3.32; Llorens-Bobadilla *et al.,* 2015; Dulken *et al.,* 2017). In all of these studies they were forced to perform the sub-dissection because the antibodies that were used to purify for stem cells are also expressed highly by other neurons, glia, and astrocytes.

Two new studies have been performed using DropSeq (and thus without antibody-mediated FAC sorting) of sub-dissected V-SVZ (Zywita *et al.,* 2018; Mizrak *et al.,* 2019). Discussed at length below, these new studies—possibly due to the lack of sequencing depth or crude isolation procedure—are not able to reliably pick up quiescent NSCs and thus cells expressing *Kcnj13* or *Prlr*, as our Group1 cells do. Other groups have used the DropSeq method to sequence many different brain tissues simultaneously, including the striatum but without sub-dissecting the V-SVZ (Zeisel *et al.,* 2018; Saunders *et al.,* 2018). These groups are able to find some cell types expressing *Kcnj13* and *Prlr*, but they are not able to reliably isolate the NSCs of the V-SVZ. This may, again, be due to low-sequencing depth or the isolation procedure itself.

I have discussed here the variety of published methods to illustrate the different findings employing those methods generates. A comparison shows that deep sequencing is required to pick up the quiescent stem cells of the V-SVZ, since all experiments employing DropSeq were unable to pick up quiescent NSCs. While we were able to pick up NSCs using our whole-striatum preparation, it is likely that we may able to isolate even more Penk-tdTomato^{+} NSCs in the future by performing a sub-dissection of the V-SVZ prior to cell suspension preparation.

6.13 The use of a relatively small number of cells to perform bioinformatics analysis.

The number of cells sequenced in StriatumSeq1 was 41, and 39 in StriatumSeq2. The deep sequencing we performed on two occasions showed that we could faithfully reproduce cell clusters using these comparatively small numbers of cells. However, the advent of DropSeq technology has meant that groups are now routinely analysing thousands of cells in single sequencing runs, albeit at drastically reduced sequencing depths (Zywita *et al.,* 2018; Zeisel *et al.,* 2018; Saunders *et al.,* 2018). The Doetsch group recently published a report where they sequenced more than 41,000 cells from lateral ventricle wholemounts (Mizrak *et al.,* 2019). As discussed above, the reduction in read depth meant that they were not able to isolate any cells displaying the transcriptomic profile of a quiescent NSC. Thus, sequencing depth is seemingly more important in being able to identify rare cell types than cell pool size. In line with this, earlier reports, using deep sequencing, were able to identify quiescent NSCs from sub-dissected V-SVZ tissues (Llorens-Bobadilla *et al.,* 2015; Dulken *et al.,* 2017). Their experiments, like ours, utilized much smaller pools of cells to perform the deep sequencing and bioinformatics analysis—with Llorens-Bobadilla *et al.* (2015) using as few as 45 cells for some experiments. Thus, we are confident that our cell pool sizes were sufficient to perform the bioinformatics analysis.

6.14 The partial progressive model of Parkinson's disease.

On the basis of several lines of evidence discussed above we believe that Penk-tdTomato labels a pool of stem cells as well as neuroblasts. We wanted to determine how these cells would respond to a traumatic stimulus. Being stem cells, we reasoned that the cells would initiate a proliferative response and that at least some of the cells would migrate from the V-SVZ towards the site of the injury and enter into the neurogenic lineage. Thus, our hypothesis was that brain injury would lead to a significant increase in proliferating Penk-tdTomato^{+} cells which would migrate to the site of injury and that some of the proliferating Penk-tdTomato^{+} cells would also express neurogenic markers.

A seminal paper by the group of Olle Lindvall showed that stroke can induce a strong neurogenic response in the brains of adult rats (Arvidsson *et al.,* 2002). For this study, the group used middle cerebral artery occlusion (MCAO) to induce a massive unilateral stroke in the rats, which caused a large lesion in the dorsal striatum. Five weeks after stroke induction, they found a 31-fold increase in the number of $BrdU^+/NeuN^+$ cells in the ipsilateral side, while the contralateral side was unaffected. They also showed that this was due mainly to the migration of neuroblasts from the V-SVZ to the injured area. Though we were interested in mouse neurogenesis, we based the theory, design, and analysis of our own experiment on the findings of this study.

Several years after the Arvidsson *et al.* (2002) paper, the group of Jonas Frisen used the same injury model in mice and showed using a Foxj1-CRE line that the neuroblasts that were generated following MCAO-induced stroke arose from $Foxj1^+/CD133^+$ ependymal cells (Carlen *et al.,* 2009). To this day, this remains quite a controversial finding as many groups believe that ependymal cells do not have stem cell capacity. A recent and thorough RNAseq experiment of the ependymal cells of the V-SVZ showed that they do not have stem cell capacity (Shah *et al.,* 2018). Nonetheless, this experiment did show that mice, like rats, exhibited a significant neurogenic response in response to brain injury. These findings of increased V-SVZ proliferation and neurogenesis have been confirmed by many independent studies (Zhang *et al.,* 2007; Kojima *et al.,* 2010; Magnusson *et al.,* 2014).

The proliferative and neurogenic response of SVZ cells to brain injury is not limited to MCAO model. It has also been shown through the use of quinolinic acid-induced striatal cell loss (Gordon *et al.,* 2007). Stroke induction in the cortex of mice using endothelin-1 injection led to a massive proliferative response and astrogenesis (Faiz *et al.,* 2015). The $Creb1^{Camkcre4}Crem^{-/-}$ mutant mice (CBCM) has been used to induce striatal degeneration (Luzzati *et al.,* 2011). In particular, they find that neuroblasts generated in the SVZ migrate to the site of the lesion, but that there is no change in the proliferative response. The study of Llorens-Bobadilla *et al.* (2015)

which I have discussed at length in this thesis, also assessed the degree of quiescent NSC activation, at the transcriptomic level, in response to brain injury. They used a severe stroke model generated by transient bilateral common artery occlusion. The group found that stroke led to a composition change in the population of NSCs at the V-SVZ: injury caused a relative increase in the number of activated NSCs and a decrease in the number of quiescent NSCs compared to controls, suggesting it led to the specific activation of the quiescent NSCs of the V-SVZ (Llorens-Bobadilla *et al.,* 2015). Thus, while V-SVZ cells become activated in many different injury paradigms the exact nature of the response does depend on the paradigm. This may be due to the temporal profile of the degeneration of striatal (and cortical) cells as well what suite of migratory, proliferative, and neurogenic cues/molecules are generated due to the injury (Aponso *et al.,* 2008).

Activation of V-SVZ stem cells is not limited to just models of acute brain injury. We were interested in chronic neurodegenerative diseases, as currently no cure or even viable long-term therapies for Huntington's and Parkinson's diseases exist. It has been clearly shown that there is an increase in proliferation and neurogenesis in the brains of adult humans with Huntington's disease (Curtis *et al.,* 2003). In the same study, they find that proliferation and neurogenesis increase in the striata of R6/2 mouse models of HD. In an elegant study using carbon^{14}-dating, Jonas Frisen's group showed that the newly generated neurons in humans become interneurons (Ernst *et al.,* 2014). Thus, there seems to be a clear link between HD and the activation of stem cells in the V-SVZ.

The role of neurogenesis in Parkinson's disease, however, remains a more complicated matter. Looking at human patients with PD, one study reported that there was a decrease in cell proliferation at the V-SVZ (Höglinger *et al.,* 2004). However, a more recent report found that the proliferative capacity of the V-SVZ remains unchanged in aged humans (van den Berge *et al.,* 2011). Similarly, in animal models of PD there are many conflicting reports regarding proliferation

and neurogenesis. The most common method of inducing a Parkinson's disease phenotype in animals is by striatal deafferentiation of dopaminergic projections through the injection of the toxin 6-hydroxydopamine (6-OHDA). As outlined in the ***Introduction***, the toxin 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP), can also be used to generate a model of PD. We, however, decided to use 6-OHDA in place of MPTP for several different reasons. The most important reason is the finding that MPTP can directly target neuroblasts for cell death thereby decreasing neurogenesis, thus confounding the analysis of the effects of Parkinson's on neurogenesis (He and Nakayama, 2009). The effects of MPTP can differ depending on species, gender, strain, and age of animal models, whereas 6-OHDA is said to have a more stable and reproducible effect (Smith and Heuer, 2012). There are also practical considerations regarding experimentation with a highly toxic substance such as MPTP. Its use requires stringent protocols, handling methods, facilities, and permission. Lastly, 6-OHDA creates a unilateral lesion, meaning that the un-injected contralateral side can be used as an internal control (Smith and Heuer, 2012).

6-OHDA is either injected directly into the medial forebrain bundle (MFB) to cause a severe phenotype (which can result in the death of the animal) or it is injected into the dorsal striatum to create the "partial progressive" model (which causes a less severe phenotype and is relatively safer for the health of the animals). One of the earliest reports on 6-OHDA injection into the MFB of mice found a decrease in cell proliferation at the V-SVZ (Baker *et al.,* 2004). They also reported a decrease in neurogenesis at the ventricular germinal niche. However, another group examined the effects of the partial progressive model by injecting 6-OHDA directly into the dorsal striatum of rats and also reported no increase in neurogenesis, but that there was a massive increase in cell proliferation at the V-SVZ (Aponso *et al,* 2008). This becomes further complicated by the data reported in the following studies. Winner *et al.* (2008) injected 6-OHDA into the MFB of adult rats and examined the effects on proliferation and neurogenesis. They found that unlike Baker *et al.* (2004) and Aponso *et al.* (2008), the injury resulted in a significant decrease in cell proliferation.

They also report that this decrease in V-SVZ cell proliferation is accompanied by an increase in the neurogenesis of extant proliferating cells (Winner *et al.,* 2008). Finally, a study examining the effects of unilateral lesioning of the striatum of adult rats using 6-OHDA reported findings that were the complete opposite of what was originally reported by Baker *et al.* (2004). This study reports that the striatal lesioning led to an increase in progenitor cell proliferation at the V-SVZ and that this was accompanied by a concomitant increase in neurogenesis (Liu *et al.,* 2006). Simultaneous increases in cell proliferation and neurogenesis in response to 6-OHDA mediated lesioning of the midbrain has also been reported in non-rodent animal models. Parish *et al.* (2006) found that salamanders responded to the striatal deafferentiation of dopaminergic projections by significantly increasing the levels of cell proliferation and neurogenesis—in their case they report a complete regeneration of the dopaminergic system. Thus, the question how the brain responds to a 6-OHDA-mediated lesioning of the striatum remains a controversial matter. We hypothesized that Penk-tdTomato^{+} cells may be involved in a proliferation and/or a neurogenic response resulting from a Parkinson's disease-like phenotype.

We chose to use the partial progressive model of Parkinson's disease generated by the unilateral injection of 6-OHDA. We reasoned that the slower degenerative profile more closely resembled the true disease phenotype. Injection into the MFB causes a more severe lesion, where complete deafferentiation has been reported to occur in as little as 24 hours (Baker *et al.,* 2004; Winner *et al.,* 2008; Smith and Heuer, 2012). This acute loss, especially in mice, can be accompanied by sudden death (Personal e-mail communication, Prof. Peter Magill; Smith and Heuer, 2012). For the above reasons, we chose to inject 6-OHDA directly into the dorsal striatum of IPNK mice to generate the partial progressive model of Parkinson's disease to test our hypothesis.

6.15 Experimental paradigm and revision.

We initially attempted to generate the lesion by a single direct injection of 6-OHDA (750nl, 6mg/ml) into the dorsal striatum (AP: -0.1mm from bregma; ML: 2.0mm from bregma; DV: -2.25mm from bregma). We reasoned that the proliferative and neurogenic responses could happen within a matter of days, as reported in the studies outlined above. Also, we were concerned about the health of the animals, as studies utilizing 6-OHDA injections in the MFB report frequently animal death. Thus, we decided to perform the analysis 7 days post-6-OHDA injection. However, when we stained using an antibody against the enzyme Tyrosine Hydroxylase (TH; a widely-accepted method for detecting dopaminergic projections) we could not detect the striatal deafferentiation that we were attempting to induce. We could not detect any changes in the intensity of TH staining in the sham or 6-OHDA injected animals, or between the ipsilateral or contralateral sides.

We believe that the lesion did not occur in this experimental paradigm either because the volume of 6-OHDA was not sufficient or because we did not allow a long enough time for the deafferentiation to occur. For this reason, we had to revise our methodology in order to generate an appropriate lesion. A recent study analysing the cellular basis for levodopa-induced dyskinesia in Parkinsonian mice (aged 3 to 6 months) performed two injections in the dorsal striatum (rostral and caudal) and found that the deafferentiation was maximized 21 days post-injection (Girasole *et al.,* 2018). Thus, we reasoned that a similar protocol (two injections: rostral site: 0.6 mm in front of Bregma; 2.3 mm lateral, 2.5 mm ventral; caudal site: -0.1 mm behind Bregma, 2.4 mm lateral, 2.4 mm ventral) in mice aged 2 to 3 months would generate a large enough lesion so that we could properly assess the response of Penk-tdTomato^{+} V-SVZ cells in a model of partial progressive Parkinson's disease. Indeed, we found that this led to a specific lesioning of the dorsal striatum in the injected ipsilateral hemisphere (Figure 5.1). Although we injected at two sites in the dorsal striatum, when we systematically analysed 9 sections, each spaced 90 μm apart, to span the area where we believed the lesioning would occur, we found that were not two separate

lesion sites—instead, the lesion occurred adjacent to the caudal site (*data not shown*). Thus, to perform the quantitative analysis we focused the analysis around the site of this large lesion. The quantitative method we used—using three sections separated by 30 μm and summing the cell numbers for each animal—was adopted from the initial report of neurogenesis in the MCAO model by the laboratory of Ole Lindvall (Arvidsson *et al.,* 2002).

6.16 Penk-tdTomato$^+$ cells proliferate in response to 6-OHDA injection into the dorsal striatum.

As outlined above, there have been many conflicting reports regarding the striatal NSC proliferation response in the partial progressive model of Parkinson's disease. In our study, we found an unequivocal increase in proliferation within the striatum in response to the 6-OHDA injections 21 days post-injection (Figures 5.2 and 5.5a). This finding is in line with the reports of Liu *et al.* (2006), Parish *et al.* (2007), Aponso *et al.* (2008) who also observe an increase in the proliferation in response to the striatal deafferentiation of dopaminergic projections. Although the magnitude was considerably smaller, we also found a significant increase in the number of proliferating Penk-tdTomato$^+$ cells (Figures 5.3 and 5.5b). This indicates that Penk-tdTomato$^+$ cells are one among several subgroups of proliferating cells.

This was a significant finding, but our experimental design did not allow us to determine the exact dynamics of the proliferation response of the Penk-tdTomato$^+$ cells—in that we are not able to tell whether the increased number of proliferating Penk-tdTomato$^+$ cells was truly due to the migration of cells from the V-SVZ. Many groups have found that there is a migration of cells from the V-SVZ in response to a striatal injury. Using cultured brain slices, Kojima *et al.* (2010) showed direct evidence that cells from the V-SVZ migrate at the neuroblast stage along blood vessels towards the site of injury—in a model of MCAO-induced ischemia. This directly supported the seminal findings of Arvidsson *et al.* (2002), who reported the presence migrating chains of

proliferating cells in response to their MCAO-induced stroke model. Zhang *et al.* (2007) arrived at the same conclusion using a similar model. Gordon *et al.* (2007) also found that the increase in proliferation due to brain injury (mediated by quinolinic acid injection) in rats was also due to the migration of neuroblasts. Interestingly, a recent study that generated a stroke-mediated lesion in the cortex using endothelin-1 injection in mice reported the same increase in proliferation but found that this was due to the migration of cells from the V-SVZ at the NSC stage (Faiz *et al.,* 2015). They posit that the NSCs migrate to the site of the cortical lesion and then proceed to differentiate and proliferate locally.

Complicating the issue even further is a recent study from the group of Jonas Frisen. Here, they used MCAO to generate a stroke and the *Connexin30*-CRE mouse model to follow the astrocyte response (Magnusson *et al.,* 2014). Using this model, they report increases in proliferation and neurogenesis but they find that one-thirds to two-thirds of this is due to the *in loco* conversion of astrocytes (labeled by their CRE line) into the neurogenic lineage. They report that this is due to the inhibition of Notch1 signaling in the astrocytes, which causes them to behave as neural stem cells. Interestingly, in our sequencing and immunostaining we are able to show that Connexin30 labels Group2 cells (which are a subgroup of putative quiescent NSCs). Thus, there may be a chance that some of this was actually due to Connexin30^{+} quiescent NSCs. Regardless, this study does introduce the possibility that the increase in proliferation we report in the striatal parenchyma in response to the 6-OHDA injection was not due to the migration of NSCs or neuroblasts from the V-SVZ, but to the conversion of Penk-tdTomato^{+} astrocytes. However, since the development of the line in the Minichiello group we have not seen Penk-tdTomato labelling astrocytes in the striatal parenchyma. Thus, the more likely possibility is that Penk-tdTomato^{+} NSCs and/or neuroblasts migrate from the V-SVZs and proliferate at the site of injury. In the future, we will need to design an experiment to directly address this.

In a pilot experiment, we stereotactically injected a virus expressing GFP into the lateral ventricle in the hemisphere where we also injected 6-OHDA, with the aim of double-labeling only the Penk-tdTomato cells lining the ventricular wall. We were able to see a few GFP$^+$/Penk-tdTomato$^+$ cells away from the lateral ventricle, indicating that they may be migrating in response to the injury. However, the experiment became unanalyzable due to the leakage of the GFP virus during the injection—we found cortical cell types, among possibly other cell types, being labeled. Thus, we could not tell with any certainty whether migrating cells had actually originated from the V-SVZ. We will need to refine the injection method in the future to only label ventricular NSCs. Although extremely difficult, we could also attempt to directly visualize the migrating Penk-tdTomato$^+$ cells in cultured slices as per the study of Kojima *et al.* (2010). This will allow us to unequivocally show whether Penk-tdTomato$^+$ cells migrate from the V-SVZ and at what stage they are migrating (by the expression of Group1, Group2, or Group3 markers). Future work will also be required to determine the mechanism by which this proliferation is induced—whether it is due to the loss of direct dopaminergic innervation caused by the 6-OHDA-mediated deafferentiation or to the release of chemokines or other released factors from the site of the lesion.

6.17 Penk-tdTomato$^+$ cells do not become neurogenic in response to 6-OHDA injection into the dorsal striatum.

As described above, there have been conflicting findings regarding the neurogenic response of the rodent brain when striatal deafferentiation of dopaminergic projections occurs. Winner *et al.* (2008) report a significant increase in neurogenesis without a concomitant increase in proliferation, while Liu *et al.* (2006) and Parish *et al.* (2007) report a significant increase in both proliferation and neurogenesis. Baker *et al.* (2004) and Aponso *et al.* (2008) find no evidence of neurogenesis in response to 6-OHDA injection. We tested the neurogenic potential of Penk-

tdTomato$^+$/BrdU$^+$ cells by determining whether they colocalised with the neurogenic markers NeuN and βIII-tubulin. We first used NeuN and found that there was no significant increase in NeuN$^+$ cells that were also Penk-tdTomato$^+$/BrdU$^+$ (Figures 5.4). We were concerned that NeuN might only label mature neurons and potentially miss newly-generated neurons, thus we sought to use an alternative marker of neuroblasts—βIII-tubulin. This analysis supported our initial findings—that there was, indeed, no significant increase in neurogenesis due to the unilateral injection of 6-OHDA 21 days post-injection in the ipsilateral striatal parenchyma (Figure 5.5 C and D). Thus, our current experiment supports the findings of Baker *et al.* (2004) and Aponso *et al.* (2008) with regards to a lack of neurogenesis.

However, another interesting possibility is that neurogenesis did indeed occur but that it occurred before the 21 days. We decided to wait for 21 days in order to generate as complete a lesion as possible, as per the Girasole *et al.* (2018) experimental protocol, after we found that a 7-day post-6-OHDA injection wait did not lead to a striatal lesion. In the initial study by Olle Lindvall's group, they waited 5 weeks after MCAO-induced stroke to perform the analysis and were able to detect a significant increase in the number of BrdU$^+$/Dcx$^+$ or βIII-tubulin$^+$ cells (Arvidsson *et al.,* 2002). The injury paradigm is of course quite different from the one we utilized in our study. In the two studies where neurogenesis was detected in rodent brains in response to 6-OHDA injection, the groups waited 14 days post-injection (Liu *et al.,* 2006; Winner *et al.,* 2008). In the two studies where neurogenesis was not detected, the groups waited 28 days post-injection to perform the immunochemical analyses (Baker *et al.,* 2004; Aponso *et al.,* 2008). Thus, in our attempt to obtain as complete a lesion as possible by waiting 21 days post-6-OHDA injection, we may have missed the neurogenic window. In the future, waiting 14 days may lead to the desired lesioning while also allowing for the detection of neurogenesis.

However, a study by Gordon *et al.* (2007) of the temporal profile of the neurogenic response in the injured striatal parenchyma suggests that the picture might be more complicated

still. They studied the effect of quinolinic acid-induced striatal cell loss in rats and found that neurogenesis peaked as soon as 2 days post-injection, after which neurogenesis was virtually undetectable after as little as 7 days. In our experiment, we waited 4 days after 6-OHDA injection to first inject BrdU. Gordon *et al.* hypothesize that the short duration of the neurogenic window is due to the loss or inhibition of migratory signals generated from the damaged area. It must be considered that the effect of their injury is much more acute than 6-OHDA injection, but it does indicate that the striatal neurogenic response is acute and not chronic. Thus, in the future we will need to design our experiment such that we can obtain a large enough lesion while being able to capture the acute neurogenic response. While our current findings suggest that Penk-tdTomato^{+} cells proliferate but are not neurogenic in response to 6-OHDA 21 days after injection, future work will be required to more closely and accurately assess the neurogenic and migratory responses of the cells of interest by a more systematic time course analysis and 6-OHDA lesion modulation.

Conclusions and Outlook

1. The BAC-Penk-Cre$^{tg/+}$;Rosa26-tdTomato$^{tg/+}$ (IPNK) line labels cells within the V-SVZ germinal niche.
2. Two different deep RNA sequencing experiments identified three different subgroups from this niche, representing quiescent NSCs and neuroblasts.
3. One of these groups, “Group15/Group1”, represents a novel subgroup of quiescent NSCs when compared to published reports.
4. This subgroup is enriched for gene markers (Prlr, 5HT$_{2C}$-receptor) that have been hypothesised to be expressed by B-cells, but which have not been shown in previously published reports of RNA sequencing of B-cells.
5. The novel subgroup is also enriched for the novel markers KiR7.1 (*KCNJ13*) and the folate receptor.
6. The gene markers for each of the three clusters were validated *in vivo* by immunochemistry.
7. KiR7.1 can be used to isolate cells, by FACS, that exhibit the gene expression profile of “Group1/Group15” cells, thereby acting as a type of functional validation.
8. In a 6-OHDA-mediated model of Parkinson’s disease, Penk-tdTomato^{+} cells increase their proliferation and can be found in the striatal parenchyma, near the injury site.
9. In the same model, 21 days post-6-OHDA injection, Penk-tdTomato^{+} cells do not exhibit neurogenesis.

The data presented in this thesis points to a new subgroup of B-cells within the V-SVZ niche. The gene expression and pathway upregulation that we found in this subgroup are line with hypotheses put forth by various groups regarding B-cells. Prlr, 5HT$_{2C}$-receptor, KiR7.1, and the folate receptor can be used to specifically study the functions and dynamics of this subgroup of B-cells. The novel markers identified herein may become useful tools to understand how fate-

specification of B-cells occurs during embryogenesis. Future work will be required to understand the functional importance and temporal dynamics of Penk expression in these B-cells during embryogenesis and beyond. Fate mapping experiments will also be required to understand whether Penk-tdTomato^{+} B-cells are predisposed, or fate-specified, to generate particular cell types. Finally, the newly identified markers will allow us to study this subgroup of B-cells in healthy and diseased animal brains and will help identify new ways to reactivate dormant NSCs in these contexts.

Bibliography

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

Altman, J. Are new neurons formed in the brains of adult mammals? *Science (80-.).* **135**, 1127–8 (1962).

Andrew, S. E. *et al.* The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. **4**, 398–403 (1993).

Antony, P. M. A. & Diederich, N. J. Parkinson ' s disease mouse models in translational research. 401–419 (2011). doi:10.1007/s00335-011-9330-x

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

Arvidsson, A., Collin, T., Kirk, D., Kokaia, Z. & O, L. Neuronal replacement from endogenous precursors in the adult. *Nat Med* **9**, 548–553 (2003).

Bagga, V., Dunnett, S. B. & Fricker, R. A. The 6-OHDA mouse model of Parkinson's disease - Terminal striatal lesions provide a superior measure of neuronal loss and replacement than median forebrain bundle lesions. *Behav. Brain Res.* **288**, 107–117 (2015).

Baker, S. A., Baker, K. A. & Hagg, T. Dopaminergic nigrostriatal projections regulate neural precursor proliferation in the adult mouse subventricular zone. *Eur. J. Neurosci.* **20**, 575–579 (2004).

Baumann, N. & Pham-Dinh, D. Biology of oligodendrocyte and myelin in the mammalian central nervous system. *Physiol. Rev.* **81**, 871–927 (2001).

Benraiss, A. *et al.* Sustained mobilization of endogenous neural progenitors delays disease progression in a transgenic model of huntington's disease. *Cell Stem Cell* **12**, 787–799 (2013).

Bentley, D. R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**, 53–9 (2008).

Bergsland, M., Werme, M., Malewicz, M., Perlmann, T. & Muhr, J. The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes Dev.* **20**, 3475–3486 (2006).

Besusso, D. *et al.* controls inhibition of locomotor behavior. *Nat. Commun.* **4**, 1–12 (2013).

Blume, E. Street drugs yield primate Parkinson's model. *J. Am. Med. Assoc.* **250**, 13–14 (1983).

Bolam, J. P., Hanley, J. J., Booth, A. C. & Bevan, M. D. Synaptic organisation of the basal ganglia. *J. Anat.* **196**, 527–542 (2000).

Boldrini, M. *et al.* Short Article Human Hippocampal Neurogenesis Persists throughout Aging Short Article Human Hippocampal Neurogenesis Persists throughout Aging. 589–599 (2018). doi:10.1016/j.stem.2018.03.015

Boutin, C., Schmitz, B., Cremer, H. & Diestel, S. NCAM expression induces neurogenesis in vivo. *Eur. J. Neurosci.* **30**, 1209–1218 (2009).

Brown, G. C. & Neher, J. J. Microglial phagocytosis of live neurons. *Nat. Publ. Gr.* **15**, 209–216 (2014).

Brown, R. S. E. *et al.* Conditional Deletion of the Prolactin Receptor Reveals Functional Subpopulations of Dopamine Neurons in the Arcuate Nucleus of the Hypothalamus. *J. Neurosci.* **36**, 9173–9185 (2016).

Brown, R. S. E., Kokay, I. C., Herbison, A. E. & Grattan, D. R. Distribution of prolactin-responsive neurons in the mouse forebrain. *J. Comp. Neurol.* **518**, 92–102 (2010).

Burns, R. S. *et al.* A primate model of parkinsonism: selective destruction of dopaminergic neurons in the pars compacta of the substantia nigra by N-methyl-4-phenyl-1,2,3,6-tetrahydropyridine. *Proc. Natl. Acad. Sci.* **80**, 4546–4550 (2006).

Cahoy, J. D. *et al.* A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J. Neurosci.* **28**, 264–78 (2008).

Calaora, V., Chazal, G., Nielsen, P. J., Rougon, G. & Moreau, H. mCD24 expression in the developing mouse brain and in zones of secondary neurogenesis in the adult. *Neuroscience* **73**, 581–594 (1996).

Capilla-Gonzalez, V., Cebrian-Silla, A., Guerrero-Cazares, H., Garcia-Verdugo, J. M. & Quiñones-Hinojosa, A. Age-related changes in astrocytic and ependymal cells of the subventricular zone. *Glia* **62**, 790–803 (2014).

Cassidy, R. M. *et al.* Neurogenesis in adult mammalian Substancia Nigra.pdf. **100**, 7925–7930 (2003).

Chaker, Z., Codega, P. & Doetsch, F. A mosaic world: puzzles revealed by adult neural stem cell heterogeneity. *Wiley Interdiscip. Rev. Dev. Biol.* **5**, 640–658 (2016).

Chevalier, G. & Deniau, J. M. Disinhibition as a basic process in the expression of striatal functions. II. The striato-nigral influence on thalamocortical cells of the ventromedial thalamic nucleus. *Brain Res.* **334**, 227–233 (1985).

Codega, P. *et al.* Article Prospective Identification and Purification of Quiescent Adult Neural Stem Cells from Their In Vivo Niche. *Neuron* **82**, 545–559 (2014).

Conover, J. C. & Shook, B. A. Aging of the Subventricular Zone Neural Stem Cell Niche. *Aging Dis.* **2**, 149–163 (2011).

Crankshaw, D. J., Crosby, D. A. & Morrison, J. J. Effects of the KIR7.1 Blocker VU590 on Spontaneous and Agonist-Induced Contractions of Human Pregnant Myometrium. *Reprod. Sci.* **24**, 1402–1409 (2017).

Crowther, R. A., Hasegawa, M., Spillantini, M. G., Jakes, R. & Goedert, M. α-Synuclein in filamentous inclusions of Lewy bodies from Parkinson's disease and dementia with Lewy bodies. *Proc. Natl. Acad. Sci.* **95**, 6469–6473 (2002).

Curtis, M. A. *et al.* Increased cell proliferation and neurogenesis in the adult human Huntington's disease brain. *Proc. Natl. Acad. Sci.* **100**, 9023–9027 (2003).

Curtis, M. A. *et al.* The distribution of progenitor cells in the subependymal layer of the lateral ventricle in the normal and Huntington's disease human brain. *Neuroscience* **132**, 777–788 (2005).

Dauer, W. & Przedborski, S. Parkinson's Disease: Mechanisms and Models. *Neuron* **39**, 889–909 (2003).

Delgado, A. C. *et al.* Endothelial NT-3 Delivered by Vasculature and CSF Promotes Quiescence of Subependymal Neural Stem Cells through Nitric Oxide Induction. *Neuron* **83**, 572–585 (2014).

Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

Doetsch, F., Caille, I., Lim, D. A., Garcı, J. M. & Alvarez-buylla, A. Subventricular Zone Astrocytes Are Neural Stem Cells in the Adult Mammalian Brain. **97**, 703–716 (1999).

Drissen, R. *et al.* Distinct myeloid progenitor-differentiation pathways identified through single-cell RNA sequencing. *Nat. Immunol.* **17**, 666–676 (2016).

Dulken, B. & Brunet, A. Stem Cell Aging and Sex: Are We Missing Something? *Cell Stem Cell* **16**, 588–590 (2015).

Fox, S. Animal Models of Movement Disorders: Volume II. *Int. J. Neurosci.* **122**, 480–481 (2012).

Freundlieb, N. Dopaminergic Substantia Nigra Neurons Project Topographically Organized to the Subventricular Zone and Stimulate Precursor Cell Proliferation in Aged Primates. *J. Neurosci.* **26**, 2321–2325 (2006).

Fuentealba, L. C. *et al.* Embryonic Origin of Postnatal Neural Stem Cells. *Cell* **161**, 1644–1655 (2015).

Geladopoulos, T., Sakellaridis, N. & Vernadakis, A. Differential maturation of μ and δ opioid receptors in the chick embryonic brain. *Neurochem. Res.* **12**, 279–288 (1987).

Gennet, N., Tamburini, C., Nan, X. & Li, M. FolR1: A novel cell surface marker for isolating midbrain dopamine neural progenitors and nascent dopamine neurons. *Sci. Rep.* **6**, 1–10 (2016).

Givan, A. L. Flow Cytometry: An Introduction. *Methods Mol. Biol. Flow Cytom. Protoc. 2nd ed.* (2004).

Gokce, O. *et al.* Cellular Taxonomy of the Mouse Striatum as Revealed by Single-Cell RNA-Seq. *Cell Rep.* 1–12 (2015).

Goldman, S. A. & Nottebohm, F. Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. *Proc. Natl. Acad. Sci.* **80**, 2390–2394 (2006).

Graybiel, A. M. The Basal Ganglia. *Curr. Biol.* **10**, 509–511 (2008).

Guez-Barber, D. *et al.* FACS Identifies Unique Cocaine-Induced Gene Regulation in Selectively Activated Adult Striatal Neurons. *J. Neurosci.* **31**, 4251–4259 (2011).

Guez-Barber, D., Fanous, S., Harvey, B. K., Zhang, Y. & Hope, B. T. FACS purification of immunolabeled cell types from adult rat brain. **203**, 416 (2013).

Han, I., You, Y., Kordower, J. H., Brady, S. T. & Morfini, G. A. Differential vulnerability of neurons in Huntington's disease: The role of cell type-specific features. *J. Neurochem.* **113**, 1073–1091 (2010).

He, X. J. & Nakayama, H. Neurogenesis in Neurotoxin-induced Animal Models for Parkinson's Disease-A Review of the Current Status. *J. Toxicol. Pathol.* **22**, 101–108 (2009).

Heid, C. A., Stevens, J., Livak, K. J. & Williams, P. M. Real Time Quantitative PCR. 986–994 (1994).

Herzenberg, L. A. *et al.* The History and Future of the Fluorescence Activated Cell Sorter and Flow Cytometry. **1827**, 1819–1827 (2002).

Höglinger, G. U. *et al.* Dopamine depletion impairs precursor cell proliferation in Parkinson disease. *Nat. Neurosci.* **7**, 726–735 (2004).

Holland, P. M., Abramson, R. D., Watson, R. & Gelfand, D. H. Detection of specific polymerase chain reaction product by utilizing the 5 ' -> 3 ' exonuclease activity of Thermus aquaticus DNA polymerase. **88**, 7276–7280 (1991).

Holtman, I. R. *et al.* Induction of a common microglia gene expression signature by aging and neurodegenerative conditions: a co-expression meta-analysis. *Acta Neuropathol. Commun.* **3**, 31 (2015).

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

Johansson *et al.* Identification of a Neural Stem Cell in the Adult Mammalian Central Nervous System. *Cell* **96**, 25–34 (1999).

Kawaguchi, D., Furutachi, S., Kawai, H., Hozumi, K. & Gotoh, Y. Dll1 maintains quiescence of adult neural stem cells and segregates asymmetrically during mitosis. *Nat. Commun.* **4**, 1880 (2013).

Keays, D. A. *et al.* The role of Tuba1a in adult hippocampal neurogenesis and the formation of the dentate gyrus. *Dev. Neurosci.* **32**, 268–277 (2010).

Kemp, J. & Powell, T. The structure of the caudate nucleus of the cat: light and electron microscopy. **262**, 383–401 (1971).

Kim, Y. *et al.* Dopamine stimulation of postnatal murine subventricular zone neurogenesis via the D3 receptor. *J. Neurochem.* **114**, 750–760 (2010).

Kita, H., Kosaka, T. & Heizmann, C. W. Parvalbumin-immunoreactive neurons in the rat neostriatum: a light and electron microscopic study. *Brain Res.* **536**, 1–15 (1990).

Kojima, T. *et al.* Subventricular zone-derived neural progenitor cells migrate along a blood vessel scaffold toward the post-stroke striatum. *Stem Cells* **28**, 545–554 (2010).

Kokay, I. C. *et al.* Analysis of prolactin receptor expression in the murine brain using a novel prolactin receptor reporter mouse. *J. Neuroendocrinol.* **30**, 1–20 (2018).

Kokovay, E. *et al.* VCAM1 is essential to maintain the structure of the SVZ niche and acts as an environmental sensor to regulate SVZ lineage progression. *Cell Stem Cell* **11**, 220–230 (2012).

Kreitzer, A. C. Physiology and pharmacology of striatal neurons. *Annu. Rev. Neurosci.* **32**, 127–47 (2009).

Kreitzer, A. C. & Malenka, R. C. Striatal Plasticity and Basal Ganglia Circuit Function. *Neuron* **60**, 543–554 (2008).

Kreitzer, A. C. & Malenka, R. C. Endocannabinoid-mediated rescue of striatal LTD and motor deficits in Parkinson's disease models. *Nature* **445**, 643–647 (2007).

Kreitzer, A. C. & Malenka, R. C. Striatal Plasticity and Basal Ganglia Circuit Function. *Neuron* **60**, 543–554 (2008).

Kriegstein, A. The Glial Nature of Embryonic and Adult Neural Stem Cells. *Annu. Rev. Neurosci.* 149–184 (2011).

Kuhlbrodt, K., Herbarth, B., Sock, E., Hermans-Borgmeyer, I. & Wegner, M. Sox10, a Novel Transcriptional Modulator in Glial Cells. *J. Neurosci.* **18**, 237–250 (2018).

Kujala, T. *et al.* Calabresi et al. ACh mediated modulation of striatal function. **2236**, 120–126 (2000).

Lan Zhang, R. *et al.* Neuroblast Division during Migration toward the Ischemic Striatum: A Study of Dynamic Migratory and Proliferative Characteristics of Neuroblasts from the Subventricular Zone. *J. Neurosci.* **27**, 3157–3162 (2007).

Lau, L. M. L. de & Breteler, M. M. B. Epidemiology of Parkinson's disease. *Lancet* **5**, 525–535 (2006).

Lee, C. *et al.* The Molecular Profiles of Neural Stem Cell Niche in the Adult Subventricular Zone. *PLoS One* **7**, (2012).

Leeman, D. S. *et al.* Lysosome activation clears aggregates and enhances quiescent neural stem cell activation during aging. **1283**, 1277–1283 (2018).

Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

Liebmann, M., Stahr, A., Guenther, M., Witte, O. W. & Frahm, C. Astrocytic Cx43 and Cx30 differentially modulate adult neurogenesis in mice. *Neurosci. Lett.* **545**, 40–45 (2013).

Lim, D. A. & Alvarez-Buylla, A. Adult neural stem cells stake their ground. *Trends Neurosci.* **37**, 563–571 (2014).

Lin, J. M. *et al.* The antitumor activity of the human FOLR1-specific monoclonal antibody, farletuzumab, in an ovarian cancer mouse model is mediated by antibody-dependent cellular cytotoxicity. *Cancer Biol. Ther.* **14**, 1032–1038 (2013).

Liu, B. fang *et al.* Proliferation of neural precursors in the subventricular zone after chemical lesions of the nigrostriatal pathway in rat brain. *Brain Res.* **1106**, 30–39 (2006).

Liu, X., Wang, Q., Haydar, T. F. & Bordey, A. Nonsynaptic GABA signaling in postnatal subventricular zone controls proliferation of GFAP-expressing progenitors. *Nat. Neurosci.* **8**, 1179–1187 (2005).

Lois, C. & Alvarez-buylla, A. Chain Migration of Neuronal Precursors. *Science (80-.).* **271**, 3–6 (1996).

Lun, M. P., Monuki, E. S. & Lehtinen, M. K. Development and functions of the choroid plexus-cerebrospinal fluid system. *Nat. Rev. Neurosci.* **16**, 445–457 (2015).

MacDonald, M. E. *et al.* A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. *Cell* **72**, 971–83 (1993).

MacMillan, J. C. *et al.* Relationship between trinucleotide repeat expansion and phenotypic variation in Huntington's disease. *Nat. Genet.* **4**, 393–397 (1993).

Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).

Madisen, L. *et al.* A robust and high-throughput Cre Repooting and characterization. *Nat Neurosci* **13**, 133–140 (2010).

Magnusson, J. P. *et al.* A latent neurogenic program in astrocytes regulated by Notch signaling in the mouse. *Science (80-.).* **346**, 237–241 (2014).

Mak, G. K. *et al.* Male pheromone – stimulated neurogenesis in the adult female brain : possible role in mating behavior. **10**, 1003–1012 (2008).

Mamber, C., Kozareva, D. A., Kamphuis, W. & Hol, E. M. Shades of gray: The delineation of marker expression within the adult rodent subventricular zone. *Prog. Neurobiol.* **111**, 1–16 (2013).

Marín, O., Anderson, S. A. & Rubenstein, J. L. R. Origin and Molecular Specification of Striatal Interneurons. *J. Neurosci.* **20**, 6063–6076 (2018).

Martin, J. B. & Gusella, J. F. Huntington's Disease: Pathogenesis and Management. *N. Engl. J. Med.* **315**, 1267–1276 (1986).

Marxreiter, F., Regensburger, M. & Winkler, J. Adult neurogenesis in Parkinson's disease. *Cell. Mol. Life Sci.* **70**, 459–473 (2013).

Merkle, F. T. *et al.* Adult neural stem cells in distinct microdomains generate previously unknown interneuron types. *Nat. Neurosci.* **17**, 207–214 (2014).

Merkle, F. T., Mirzadeh, Z. & Alvarez-Buylla, A. Mosaic organization of neural stem cells in the adult brain (Supporting Online Material). *Science (80-.).* **317**, 381–4 (2007).

Mirzadeh, Z., Doetsch, F., Sawamoto, K., Wichterle, H. & Alvarez-Buylla, A. The Subventricular Zone En-face: Wholemount Staining and Ependymal Flow. *J. Vis. Exp.* 1–8 (2010). doi:10.3791/1938

Mirzadeh, Z., Merkle, F. T., Soriano-Navarro, M., Garcia-Verdugo, J. M. & Alvarez-Buylla, A. Neural stem cells confer unique pinwheel architecture to the ventricular surface in neurogenic regions of the adult brain. *Cell Stem Cell* **3**, 265–78 (2008).

Mizrak, D. *et al.* Single-Cell Analysis of Regional Differences in Adult V-SVZ Neural Stem Cell Lineages Article Single-Cell Analysis of Regional Differences in Adult V-SVZ Neural Stem Cell Lineages. *CellReports* **26**, 394-406.e5 (2019).

Mohanty, V., Siddiqui, M. R., Tomita, T. & Mayanil, C. S. Folate receptor alpha is more than just a folate transporter. *Neurogenesis* **4**, e1263717 (2017).

Morizur, L. *et al.* Distinct Molecular Signatures of Quiescent and Activated Adult Neural Stem Cells Reveal Specific Interactions with Their Microenvironment. *Stem Cell Reports* **11**, 565–577 (2018).

Mulder, J. *et al.* Endocannabinoid signaling controls pyramidal cell specification and long-range axon patterning. *Proc. Natl. Acad. Sci.* **105**, 8760–8765 (2008).

Nath, A. R., Drissen, R., Guo, F., Nerlov, C. & Minichiello, L. Studying BDNF/TrkB Signaling: High-Throughput Microfluidic Gene Expression Analysis from Rare or Limited Samples of Adult and Aged Central Neurons. *Neuromethods* (2017). doi:10.1007/7657

Norstedt, G. *et al.* Development of Pure Prolactin Receptor Antagonists. *J. Biol. Chem.* **278**, 35988–35999 (2003).

Novak, A., Guo, C., Yang, W., Nagy, A. & Lobe, C. G. Z/EG, a Double Reporter Mouse Line That Expresses Enhanced Green Fluorescent Protein Upon Cre-Mediated Excision. *genesis* **155**, 147–155 (2000).

O'Keeffe, G. C. *et al.* Dopamine-induced proliferation of adult neural precursor cells in the mammalian subventricular zone is mediated through EGF. *Proc. Natl. Acad. Sci.* **106**, 8754–8759 (2009).

Obernier, K. & Alvarez-Buylla, A. Neural stem cells: origin, heterogeneity and regulation in the adult mammalian brain. *Development* **146**, 1–15 (2019).

Osman, F., Leutenegger, C., Golino, D. & Rowhani, A. Real-time RT-PCR (TaqMan®) assays for the detection of Grapevine Leafroll associated viruses 1 – 5 and 9. **141**, 22–29 (2007).

Oude Ophuis, R. J. A., Boender, A. J., van Rozen, A. J. & Adan, R. A. H. Cannabinoid, melanocortin and opioid receptor expression on DRD1 and DRD2 subpopulations in rat striatum. *Front. Neuroanat.* **8**, 1–12 (2014).

Paez-Gonzalez, P., Asrican, B., Rodriguez, E. & Kuo, C. T. Identification of distinct ChAT+ neurons and activity-dependent control of postnatal SVZ neurogenesis. *Nat. Neurosci.* **17**, 934–942 (2014).

Papanikolaou, M. Expression and Function of KiR7.1 in the Murine Central Nervous System. *Dissertation* (2014).

Parish, C. L., Beljajeva, A., Arenas, E. & Simon, A. Midbrain dopaminergic neurogenesis and behavioural recovery in a salamander lesion-induced regeneration model. *Development* **134**, 2881–2887 (2007).

Patrick, T. A., Kranz, D. M., Van Dyke, T. A. & Roy, E. J. Folate receptors as potential therapeutic targets in choroid plexus tumors of SV40 transgenic mice. *J. Neurooncol.* **32**, 111–123 (1997).

Paul, A., Chaker, Z. & Doetsch, F. Hypothalamic regulation of regionally distinct adult neural stem cells and neurogenesis. *Science (80-.).* **356**, 1383–1386 (2017).

Pawluski, J. L., Brummelte, S., Barha, C. K., Crozier, T. M. & Galea, L. A. M. Effects of steroid hormones on neurogenesis in the hippocampus of the adult female rodent during the estrous cycle, pregnancy, lactation and aging. *Front. Neuroendocrinol.* **30**, 343–357 (2009).

Penney, J. B., Vonsattel, J. P., MacDonald, M. E., Gusella, J. F. & Myers, R. H. CAG repeat number governs the development rate of pathology in huntington's disease. *Ann. Neurol.* **41**, 689–692 (1997).

Perera, M. *et al.* Defective neuronogenesis in the absence of Dlx5. *Mol. Cell. Neurosci.* **25**, 153–161 (2004).

Perese, D. A., Ulman, J., Viola, J., Ewing, S. E. & Bankiewicz, K. S. A 6-hydroxydopamine-induced selective parkinsonian rat model. **494**, 285–293 (1989).

Phillips, W. Abnormalities of Neurogenesis in the R6/2 Mouse Model of Huntington's Disease Are Attributable to the In Vivo Microenvironment. *J. Neurosci.* **25**, 11564–11576 (2005).

Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–8 (2013).

Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–81 (2014).

Platel, J. *et al.* Neuronal integration in the adult mouse olfactory bulb is a non-selective addition process. 1–21 (2019).

Plotkin, J. L. *et al.* Impaired TrkB receptor signaling underlies corticostriatal dysfunction in Huntington's disease. *Neuron* **83**, 178–188 (2014).

Ponti, G. *et al.* Cell cycle and lineage progression of neural progenitors in the ventricular-subventricular zones of adult mice. *Proc. Natl. Acad. Sci.* **110**, E1045–E1054 (2013).

Prüss, H. *et al.* Potassium channel expression in adult murine neural progenitor cells. *Neuroscience* **180**, 19–29 (2011).

Scholzen, T. & Gerdes, J. The Ki-67 Protein: From the Known and the Unknown. *J. Cell. Physiol.* **182**, 311–322 (2000).

Scott, C. E. *et al.* SOX9 induces and maintains neural stem cells. *Nat. Neurosci.* **13**, 1181–1189 (2010).

Shah, P. T. *et al.* Single-Cell Transcriptomics and Fate Mapping of Ependymal Cells Reveals an Absence of Neural Stem Cell Function. *Cell* **173**, 1045-1057.e9 (2018).

Shan, X. *et al.* Enhanced De Novo Neurogenesis and Dopaminergic Neurogenesis in the Substantia Nigra of 1-Methyl-4-phyenyl-1,2,3,6-Tetrahydropyridine-Induced Parkinson's Disease-Like Mice. *Stem Cells* **24**, 1280–1287 (2006).

Shen, Q. *et al.* Adult SVZ Stem Cells Lie in a Vascular Niche: A Quantitative Analysis of Niche Cell-Cell Interactions. *Cell Stem Cell* **3**, 289–300 (2008).

Shingo, T. *et al.* Pregnancy-Stimulated Neurogenesis in the Adult Female Forebrain Mediated by Prolactin. **299**, 117–121 (2003).

Simola, N. & Morelli, M. Mechanisms of Neurotoxicity of 6-OHDA. *Neurotox. Res.* **11**, 151–167 (2007).

Song, M. *et al.* Optogenetic stimulation of glutamatergic neuronal activity in the striatum enhances neurogenesis in the subventricular zone of normal and stroke mice. *Neurobiol. Dis.* **98**, 9–24 (2017).

Sorrells, S. F. *et al.* Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. *Nature* **555**, 377–381 (2018).

Stoker, T. B., Torsney, K. M., Barker, R. A. & Nolan, Y. Emerging Treatment Approaches for Parkinson ' s Disease. **12**, 1–10 (2018).

Surmeier, D. J., Ding, J., Day, M., Wang, Z. & Shen, W. D1 and D2 dopamine-receptor modulation of striatal glutamatergic signaling in striatal medium spiny neurons. *Trends Neurosci.* **30**, 228–235 (2007).

Tavazoie, M. *et al.* A Specialized Vascular Niche for Adult Neural Stem Cells. *Cell Stem Cell* **3**, 279–288 (2008).

Tecuapetla, F., Jin, X., Lima, S. Q. & Costa, R. M. Complementary Contributions of Striatal Projection Pathways to Action Initiation and Execution. *Cell* **166**, 703–715 (2016).

Tepper, J. M., Tecuapetla, F., Koós, T. & Ibáñez-Sandoval, O. Heterogeneity and Diversity of Striatal GABAergic Interneurons. *Front. Neuroanat.* **4**, 1–18 (2011).

Uitterlinden, G., Robak, L. A., Jansen, I. E., Rooij, J. Van & Kraaij, R. Excessive burden of lysosomal storage disorder gene variants in Parkinson's disease. 3191–3203 (2017). doi:10.1093/brain/awx285

Van Den Berge, S. A. *et al.* The proliferative capacity of the subventricular zone is maintained in the parkinsonian brain. *Brain* **134**, 3249–3263 (2011).

Van Den Berge, S. A. *et al.* The proliferative capacity of the subventricular zone is maintained in the parkinsonian brain. *Brain* **134**, 3249–3263 (2011).

Vatanashevanopakorn, C. *et al.* Studying BDNF/TrkB Signaling: Transcriptome Analysis from a Limited Number of Purified Adult or Aged Murine Brain Neurons. *Neuromethods* (2017). doi:10.1007/7657

Venter, J. C., Smith, H. O. & Hood, L. A new strategy for genome sequencing. *Nature* **381**, 364–366 (1996).

Walker, F. O. Huntington's Disease. *Semin Neurol* **27**, 143–150 (2007).

Westmoreland, J. J., Hancock, C. R. & Condie, B. G. Neuronal development of embryonic stem cells: A model of GABAergic neuron differentiation. *Biochem. Biophys. Res. Commun.* **284**, 674–680 (2001).

Winner, B. *et al.* Striatal deafferentation increases dopaminergic neurogenesis in the adult olfactory bulb. *Exp. Neurol.* **197**, 113–121 (2006).

Winner, B. & Winkler, J. Adult Neurogenesis in Neurodegenerative. *Cold Spring Harb. Perspect. Biol.* (2015).

Wolf, H. K. *et al.* NeuN: A useful neuronal marker for diagnostic histopathology. *J. Histochem. Cytochem.* **44**, 1167–1171 (1996).

Xi, J. H. *et al.* Evidence of apoptosis in the subventricular zone and rostral migratory stream in the MPTP mouse model of Parkinson disease. *J. Neuropathol. Exp. Neurol.* **65**, 873–882 (2006).

Yin, W. *et al.* The potassium channel KCNJ13 is essential for smooth muscle cytoskeletal organization during mouse tracheal tubulogenesis. *Nat. Commun.* **9**, 1–13 (2018).

Yuzwa, S. A. *et al.* Developmental Emergence of Adult Neural Stem Cells as Revealed by Single-Cell Transcriptional Profiling. *Cell Rep.* **21**, 3970–3986 (2017).

Zeng, X. S., Geng, W. S., Jia, J. J., Chen, L. & Zhang, P. P. Cellular and molecular basis of neurodegeneration in Parkinson disease. *Front. Aging Neurosci.* **10**, 1–16 (2018).

Zhang, J., Moats-Staats, B. M., Ye, P. & D'Ercole, A. J. Expression of Insulin-Like Growth Factor System Genes During the Early Postnatal Neurogenesis in the Mouse Hippocampus. *J. Neurosci. Res.* **85**, 1618–1627 (2007).

Zhang, X., Goncalves, R. & Mosser, D. M. The Isolation and Characterization of Murine Macrophages. *Curr. Protoc. Immunol.* **November**, 1–18 (2008).

Appendix

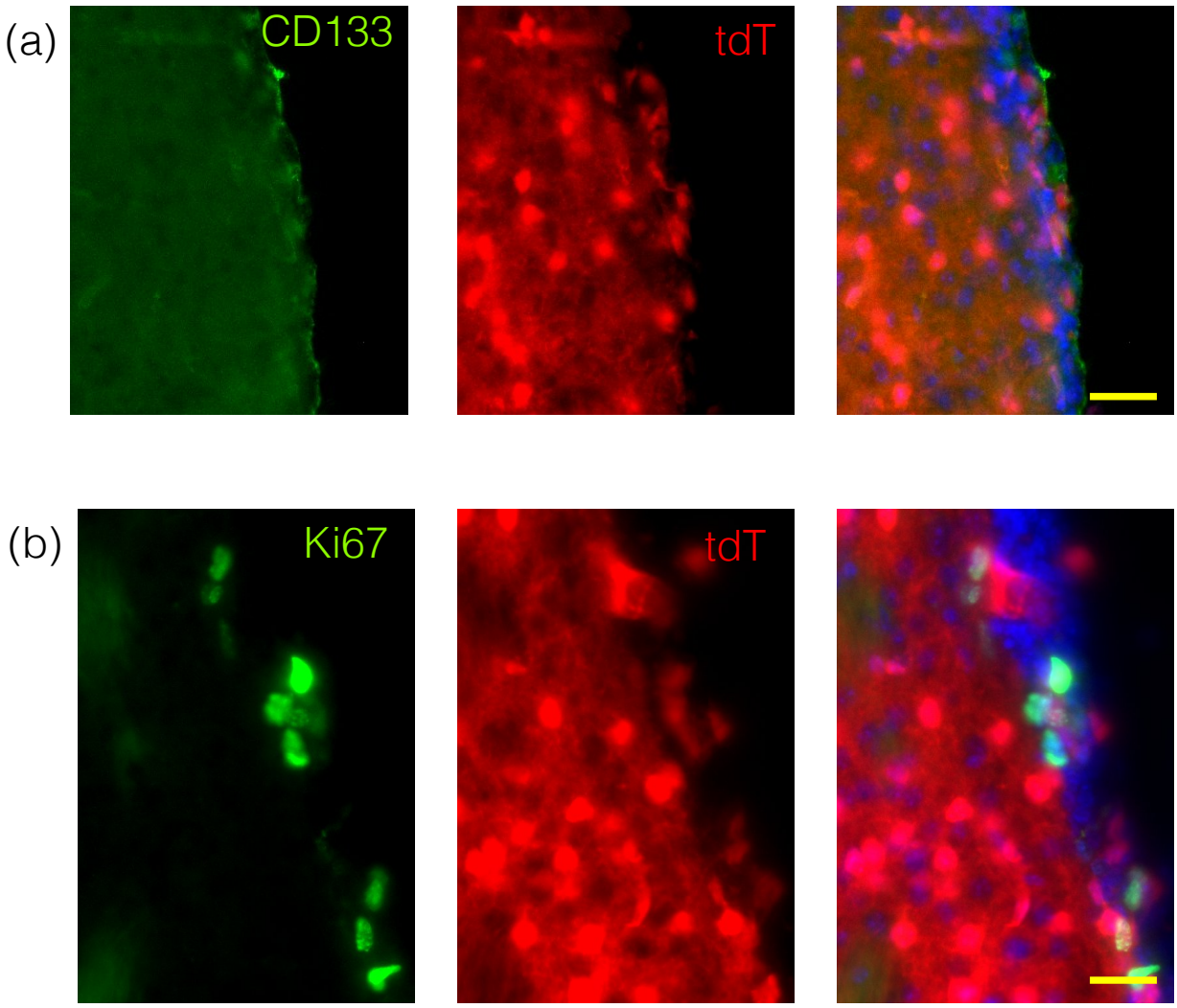


Figure A1. Analysing the colocalisation of ependymal and proliferation markers in the cells of the V-SVZ. A 2-month old IPNK brain was cryosectioned at 30µm and immunostained with antibodies against CD133 and Ki67. Presented are zoomed images of the V-SVZ of representative sections used for the quantification presented in Figure 4.1. Scale bars apply each panel, scale bar = 20µm in all panels.

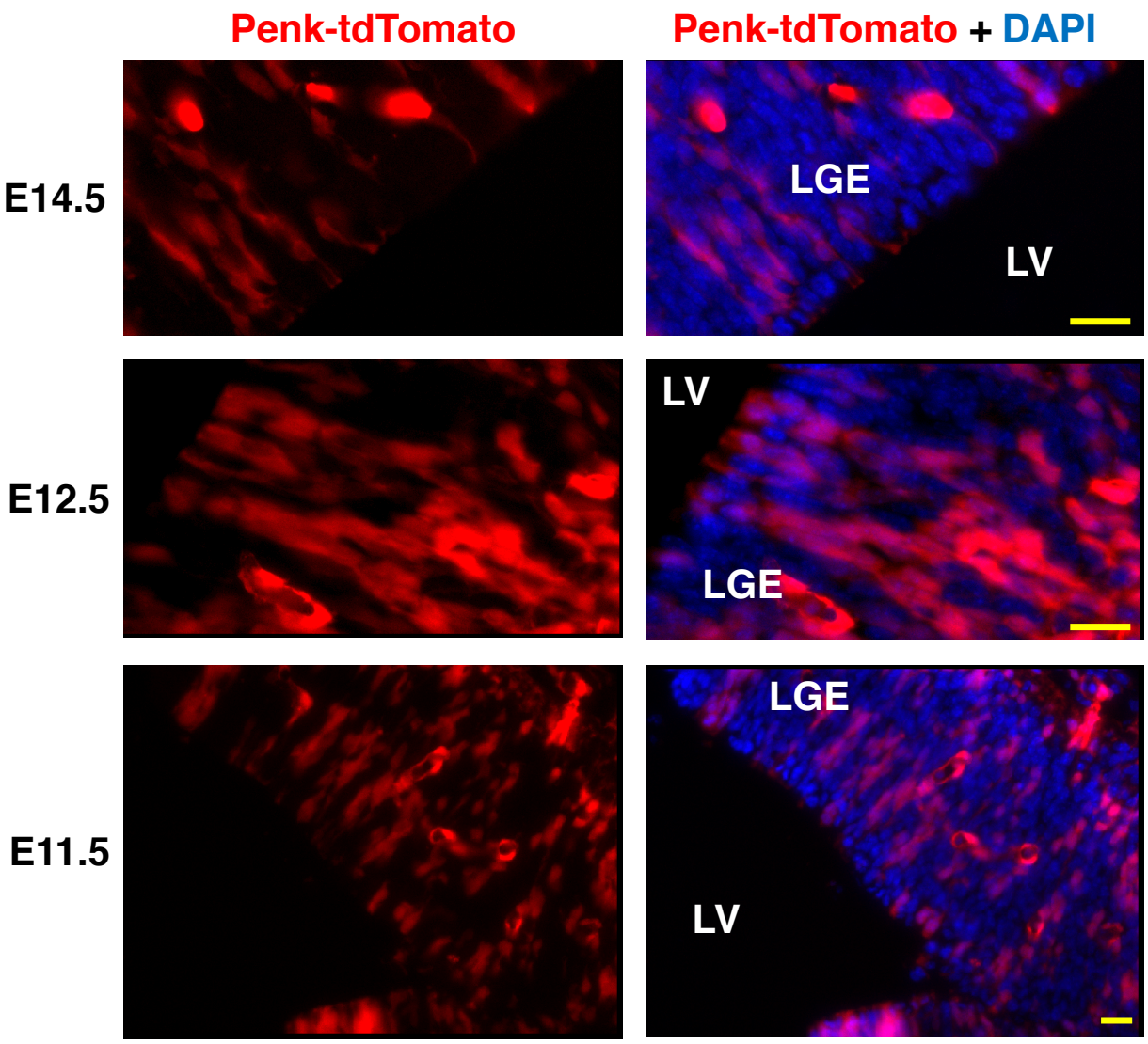


Figure A2. Analysing the expression of Penk-tdTomato in embryonic brains. Embryonic IPNK brains (E14.5, E12.5, and E11.5) were cryosectioned at 10µm and treated with only DAPI. Presented are zoomed images of the lateral ventricle (**LV**) and the lateral ganglionic eminence (**LGE**; where B-cells are generated). Strong Penk-tdTomato staining can be seen in the LGE at all three embryonic stages. Scale bars apply each panel, scale bar = 20µm in all panels.