Understanding the Mechanism Underlying the Neural-to-Glial Transition in the Developing Spinal Cord

Thomas Watson

University College London
and
The Francis Crick Institute
PhD Supervisor: James Briscoe

A thesis submitted for the degree of
Doctor of Philosophy
University College London
October 2018
Declaration

I, Thomas Watson, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.
Abstract

In the developing spinal cord, neurons and glia are generated sequentially. In vivo, this process is under tight spatial and temporal control. SOX9, NFIA, and NFIB form the basis of a transcriptional network which coordinates the activation of the glial lineage. However, molecular detail of the regulatory network controlling the timing of their expression is limited. We have developed an in vitro model system using embryonic stem cells that reproduces the dynamics and regulatory relationships observed in vivo. We leverage the in vitro system to investigate the temporal control of the neural-to-glial transition.

Previous work has implicated TGF-β and NOTCH signalling in controlling developmental transitions throughout central nervous system development. We used the in vitro system to investigate the role of these pathways in the neural-to-glial transition. Manipulating TGF-β signalling shifted the timing of glial lineage activation whilst activating NOTCH signalling was sufficient to induce the expression of glial genes. We suggest that these pathways might form a mechanism by which neuron differentiation instructs developmental progression.

Using single-cell transcriptomics, we reconstructed the developmental trajectory of progenitors transitioning to a glial fate. By plotting the transcriptional changes that accompany the transition we identified novel candidate genes involved. Using CRISPR-Cas9 and lentiviral approaches, we test these candidates and investigate their effect on the transition. Through this approach we reveal ZFP536, a transcription factor required for the correct timing of the neural-to-glial transition. ZFP536 is expressed during the transition and regulates SOX9 expression both in vitro and in vivo. Our work provides new understanding of how the timing of the neural-to-glial transition is controlled in the developing spinal cord.
Impact Statement

Understanding the mechanisms underlying the temporal control of developmental transitions is of broad interest to developmental biologists. Accurately encoding time is important for generating the correct numbers of differentiated cell types that will form a fully functional multicellular organism. Additionally, changes in the timing of developmental transitions has led to major evolutionary advances. For example, the expansion of the cortex has led to the higher cognitive ability of humanoids. This can be attributed in part to a delay in a developmental transition from neuron differentiation to gliogenesis. Understanding how cells encode time will allow us to understand how these evolutionary advances were made.

This work also contributes towards the advance of regenerative medicine and tissue engineering. The invention of techniques to generate induced pluripotent stem (iPS) cells has allowed the potential for cell replacement therapies. This requires the directed differentiation of pluripotent cells. A detailed understanding of the mechanisms underlying developmental decision making will enable the design of differentiation protocols that generate functional cell types with high purity. Additionally, generating these differentiated cell types from patient-derived iPS cells allows the investigation of the cellular biology underlying their disease.
Acknowledgement

Firstly, I would like to thank my supervisor, James, whose pushed and supported me throughout my Ph.D. He has given me every opportunity to explore my project and encouraged me to develop new and existing skills. A group of people have contributed directly to the work and are referred to, where appropriate, in the main text. Nevertheless, I acknowledge their contributions here. Robert Blassberg for teaching me everything I now know about molecular biology. Designing and manufacturing the reagents to quickly generate tools that allow the manipulation of any gene of interest was great fun. Julien Delile for answering every question I had on R programming and for setting time aside for the analysis of my single cell sequencing data. The chick electroporation work would not have been possible without Despina Stamataki, who suffered my inept attempts at injecting DNA into Chick neural tubes. Despina and Vicki Metzis both lent their experience to the Zfp536 in situ hybridisation experiments. Katherine Exelby and Manuela Melchionda helped me to dissect, embed and section neural tubes. Manuela did the SOX9 time course in mouse embryos. Andreas Sagner provided the Day 5 single cell sequencing data set but was also fantastic for providing feedback of the thesis. I must thank all members of the Briscoe Lab, past and present, for their contribution to scientific discussion and for generating a fantastic working environment - it has been a lot of fun.

At the Francis Crick Institute, we are fortunate to have a number of core facilities that aid us in pushing our science forward. I’ve used a number of them heavily. The flow cytometry facility performed the FACS experiments of the lentiviral work. The genomics equipment park performed a lot of the RNA extractions and DNA purifications. The high-throughput sequencing facility performed the single cell sequencing experiments. BRF handled the mice, whilst the media preparation facility produced reagents such as TAE and LB media.

Finally, I would like to acknowledge the contribution of my friends and family whose tireless support has pushed me forwards over these four years. I am no poet and it’s difficult to find the words to describe what you mean to me but I love you all. My
parents, however, deserve a special mention. You have given me everything and thank you just isn’t sufficient. I love you both so much - this body of work is dedicated to you.
Table of Contents

Abstract .................................................................................................................................................. 3
Acknowledgement.................................................................................................................................. 5
Table of Contents .................................................................................................................................. 7
Table of figures ....................................................................................................................................... 10
List of tables .......................................................................................................................................... 12
Abbreviations ....................................................................................................................................... 13
Chapter 1.Introduction .......................................................................................................................... 16
  1.1 The developmental origin of the central nervous system ................................................................. 16
  1.2 Neurons and glia are the building blocks of the CNS .................................................................. 17
  1.3 Timing mechanisms in development ............................................................................................ 19
    1.3.1 Systemic Control ...................................................................................................................... 19
    1.3.2 Tissue Level Control .............................................................................................................. 20
    1.3.3 Cell Intrinsic Control ............................................................................................................. 22
  1.4 The development of the spinal cord .............................................................................................. 24
    1.4.1 Patterning and differentiation of neural progenitors ............................................................... 24
    1.4.2 NFIA and NFIB are the molecular correlates of gliogenesis ............................................... 27
    1.4.3 The function of NFIA and NFIB in gliogenesis .................................................................... 27
    1.4.4 The role of NFIA in neuron differentiation ........................................................................... 28
    1.4.5 Terminal glial differentiation .................................................................................................. 29
    1.4.6 Oligodendrogeneration is controlled by a distinct regulatory network .................. 31
  1.5 A gene regulatory network controls the timing of NFIA and NFIB induction and terminal gliogenesis ........................................................................................................................................ 32
    1.5.1 SOX9 plays a central role in gliogenesis and the timing of NFIA and NFIB induction .................................................................................................................................................. 32
    1.5.2 BRN2 activates NFIA with SOX9 but via distinct mechanisms ............................................ 34
    1.5.3 Multiple roles for micro RNAs in the neural-to-glial transition ............................................ 35
  1.6 Tissue level control gliogenesis .................................................................................................. 36
  1.7 Systemic Cues .............................................................................................................................. 39
  1.8 Elucidating the mechanisms behind the neural-to-glial transition ......................................... 39
    1.8.1 In vitro approaches .................................................................................................................... 39
    1.8.2 Single cell sequencing and pseudotemporal reconstruction ................................................. 40
  1.9 Aims ............................................................................................................................................... 41
Chapter 2.Materials and methods ....................................................................................................... 43
  2.1 Tissue Culture Techniques ......................................................................................................... 43
    2.1.1 Culture of MEFs ...................................................................................................................... 43
    2.1.2 Preparation of feeders .......................................................................................................... 43
    2.1.3 Thawing feeders for ES cell culture ....................................................................................... 44
    2.1.4 Culture of ES cells ............................................................................................................... 44
    2.1.5 Differentiation of ES cells to spinal cord neural progenitors ............................................. 45
    2.1.6 Transforming Growth Factor (TGF)-β manipulations ...................................................... 46
    2.1.7 Differentiation to GFAP positive glia .................................................................................... 46
    2.1.8 Analysing samples by quantitative PCR (qPCR) ................................................................ 46
    2.1.9 Analysing samples by immunofluorescence ...................................................................... 47
    2.1.10 Analysing samples by flow cytometry ................................................................................. 47
  2.2 General Molecular Biology Techniques .................................................................................... 48
Reconstruction of the neural signalling pathways affecting the neural transition

ZFP536 affects the neur

ZFP536 is required for the correct timing of the neural transition

ZFP536 is sufficient to activate glial gene expression

Activation of NOTCH signalling induces glial genes

A lentiviral reporter system to test candidates

Loss of gliogenesis upon deletion of SOX9

Loss of gliogenesis upon combinatorial deletion of NFIA and NFIB

Single cell sequencing and analysis

In situ hybridisation on sectioned embryos

In situ hybridisation probe design and manufacture

Sample collection and sequencing of single cells

CRISPR/Cas9

Lentivirus production

Transformation and growth of chemically competent cells

Restriction digestion

PCR

Discussion

Manipulating TGF-β signalling affects the timing of the neural-to-glia

transition

ZFP536 affects the neural-to-glia transition via SOX9

ZFP536 is required for the correct timing of the neural-to-glia transition

ZFP536 affects the neural-to-glia transition via SOX9

An in vitro model of the neural-to-glial transition

The in vitro system recapitulates the in vivo phenotypes of glial genes

Loss of gliogenesis upon combinatorial deletion of NFIA and NFIB

Loss of gliogenesis upon deletion of SOX9

Reconstructing the neural-to-glia transition by single cell sequencing

The effect of NOTCH signalling on the neural-to-glia transition

A lentiviral reporter system to test candidates

Activation of NOTCH signalling induces glial genes

Chapter 3. An in vitro model of the neural-to-glial transition

Chapter 4. The in vitro system recapitulates the in vivo phenotypes of glial genes

Chapter 5. Manipulating TGF-β signalling affects the timing of the neural-to-glial transition

Chapter 6. Reconstructing the neural-to-glial transition by single cell sequencing

Chapter 7. The effect of NOTCH signalling on the neural-to-glial transition

Chapter 8. ZFP536 affects the neural-to-glial transition via SOX9

Chapter 9. Discussion

An in vitro model of the neural-to-glial transition

iNPs recapitulate the phenotypes of glial genes and gives additional insight into the molecular mechanism

Signalling pathways affecting the neural-to-glial transition

TGF-β signalling links neuron differentiation to the neural-to-glial transition

NOTCH signalling uncouples neuron differentiation and the neural-to-glial transition

Reconstruction of the neural-to-glial transition by single cell sequencing identifies candidate genes

Lin28a, Lin28b, Trim71

Trim8

Cux2

Npas3
9.5 ZFP536 affects SOX9 expression and hence the timing of the neural-to-glial transition ................................................................. 149
9.5.1 What regulates Zfp536? .................................................................................................................. 151

9.6 The timing of the neural-to-glial transition .................................................................................. 151
9.6.1 Cell autonomous control of the transition .............................................................................. 152
9.6.2 Tissue level control of the transition ....................................................................................... 153

9.7 Conclusions ................................................................................................................................. 154

Appendix A ........................................................................................................................................ 157
Appendix B ........................................................................................................................................ 168
Reference List .................................................................................................................................... 170
## Table of figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The ‘activation-transformation’ hypothesis</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>Cortex development in the mouse</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>The patterning of the neural tube</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>The canonical JAK-STAT signalling pathway</td>
<td>31</td>
</tr>
<tr>
<td>5</td>
<td>An overview of Transforming Growth Factor-β signalling</td>
<td>37</td>
</tr>
<tr>
<td>6</td>
<td>Differentiation of mouse ES cells to iNPs and glia</td>
<td>63</td>
</tr>
<tr>
<td>7</td>
<td>Neuron differentiation begins at D6 in vitro</td>
<td>63</td>
</tr>
<tr>
<td>8</td>
<td>The dynamics of SOX9, NFIA and NFIB are recapitulated in vitro</td>
<td>64</td>
</tr>
<tr>
<td>9</td>
<td>The neural-to-glial transition is recapitulated in vitro</td>
<td>65</td>
</tr>
<tr>
<td>10</td>
<td>The in vitro model recapitulates in vivo gene expression dynamics</td>
<td>66</td>
</tr>
<tr>
<td>11</td>
<td>Terminal gliogenesis is recapitated in vitro</td>
<td>67</td>
</tr>
<tr>
<td>12</td>
<td>Confirmation of NFIA deletion in NFIA and NFIA/B knockouts</td>
<td>76</td>
</tr>
<tr>
<td>13</td>
<td>Confirmation of NFIB deletion in NFIA and NFIA/B knockouts</td>
<td>77</td>
</tr>
<tr>
<td>14</td>
<td>Neuron differentiation is affected in NFIB(^{-}) but not NFIA(^{-}), NFIA/B(^{-})</td>
<td>78</td>
</tr>
<tr>
<td>15</td>
<td>Glial lineage activation is impaired in NFIA(^{-}), NFIA/B(^{-})</td>
<td>78</td>
</tr>
<tr>
<td>16</td>
<td>Abrogation of terminal gliogenesis in NFIA/B(^{-}) cells</td>
<td>79</td>
</tr>
<tr>
<td>17</td>
<td>Confirmation of SOX9 deletion in SOX9(^{-}) cells</td>
<td>80</td>
</tr>
<tr>
<td>18</td>
<td>Neuron differentiation is unaffected by deletion of SOX9</td>
<td>81</td>
</tr>
<tr>
<td>19</td>
<td>The expression of glial lineage genes is reduced in SOX9(^{-}) cells</td>
<td>82</td>
</tr>
<tr>
<td>20</td>
<td>TGF-β1-3 are expressed at E12.5 with NFIA</td>
<td>89</td>
</tr>
<tr>
<td>21</td>
<td>TGF-B signalling is active in vitro</td>
<td>90</td>
</tr>
<tr>
<td>22</td>
<td>Manipulating TGF-β signalling affects neuron differentiation</td>
<td>90</td>
</tr>
<tr>
<td>23</td>
<td>Manipulating TGF-β signalling affects Sox9 expression</td>
<td>91</td>
</tr>
<tr>
<td>24</td>
<td>Manipulating TGF-β signalling affects the onset of gliogenesis</td>
<td>92</td>
</tr>
<tr>
<td>25</td>
<td>Manipulating TGF-β signalling affects glial lineage activation</td>
<td>93</td>
</tr>
<tr>
<td>26</td>
<td>The timing of the neural-to-glial transition varies with the length of TGF-β inhibition</td>
<td>94</td>
</tr>
<tr>
<td>27</td>
<td>Sox9 is not required for the effect of TGF-β signalling on the neural-to-glial transition</td>
<td>95</td>
</tr>
<tr>
<td>28</td>
<td>Validation of the samples used for single cell sequencing</td>
<td>102</td>
</tr>
<tr>
<td>29</td>
<td>Removal of neurons and mesoderm cells from the data set</td>
<td>103</td>
</tr>
</tbody>
</table>
Figure 30 Gene modules correlated with sample time ................................. 104
Figure 31 Principle Component Analysis (PCA) using time gene modules captures developmental time................................................................................................................. 105
Figure 32 Smoothing single cell expression data using LOESS........................ 106
Figure 33 Identifying dynamically expressed genes........................................ 107
Figure 34 Genes with decreasing expression across pseudotime.................. 108
Figure 35 Genes with increasing expression across pseudotime.................... 109
Figure 36 Pseudotemporal reconstruction reproduces in vitro dynamics ...... 110
Figure 37 Development of a lentiviral reporter system.................................. 116
Figure 38 Constitutive expression of HES1 blocks neuron differentiation....... 117
Figure 39 Activation of NOTCH signalling induces glial genes.................... 118
Figure 40 Zfp536 and SOX9 expression in vivo ............................................ 129
Figure 41 Constitutive expression of ZFP536 in vitro and in vivo .................. 130
Figure 42 Neuron differentiation is unaffected by deletion of ZFP536......... 131
Figure 43 SOX9 expression is delayed in ZFP536(1) ........................................ 132
Figure 44 Deletion of ZFP536 reduces glial gene expression.......................... 133
Figure 45 Glial phenotypes are recapitulated in an alternate deletion of ZFP536 134
Figure 46 Zfp536 induction is unaffected by deletion of SOX9 ..................... 135
Figure 47 Combined deletion of SOX9/ZFP536 does not affect the onset of neuron differentiation.................................................................................................................. 135
Figure 48 No additional reduction of glial genes upon combined deletion of ZFP536 and SOX9.................................................................................................................. 136
List of tables

Table 1 List of antibodies........................................................................................................ 57
Table 2 Primers for qPCR...................................................................................................... 57
Table 3 List of plasmids......................................................................................................... 58
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APCDD1</td>
<td>APC Down-Regulated 1</td>
</tr>
<tr>
<td>ASCL1</td>
<td>Achaete-Scute Family BHLH Transcription Factor</td>
</tr>
<tr>
<td>BMP</td>
<td>Bone Morphogenic Protein</td>
</tr>
<tr>
<td>BrdU</td>
<td>Bromodeoxyuridine</td>
</tr>
<tr>
<td>BHLH</td>
<td>Basic Helix-Loop-Helix</td>
</tr>
<tr>
<td>BRN2</td>
<td>POU Class 3 Homeobox 2</td>
</tr>
<tr>
<td>CDX</td>
<td>Caudal Type Homeobox</td>
</tr>
<tr>
<td>CHIR</td>
<td>CHIR99021</td>
</tr>
<tr>
<td>CLE</td>
<td>Caudal Lateral Epiblast</td>
</tr>
<tr>
<td>CNS</td>
<td>Central Nervous System</td>
</tr>
<tr>
<td>CNTF</td>
<td>Neurotrophic Factor</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered Regulatory Interspaced Short Palindromic Repeats</td>
</tr>
<tr>
<td>CUX2</td>
<td>Cut Like Homeobox 2</td>
</tr>
<tr>
<td>DBX</td>
<td>Developing Brain Homeobox 2</td>
</tr>
<tr>
<td>DICER</td>
<td>Dicer 1, Ribonuclease 3</td>
</tr>
<tr>
<td>DROSHA</td>
<td>Drosha Ribonuclease 3</td>
</tr>
<tr>
<td>e</td>
<td>Enhancer</td>
</tr>
<tr>
<td>E</td>
<td>Embryonic day (or days post coitum)</td>
</tr>
<tr>
<td>EGF</td>
<td>Epidermal Growth Factor</td>
</tr>
<tr>
<td>ES</td>
<td>Embryonic Stem</td>
</tr>
<tr>
<td>FABP7</td>
<td>Fatty Acid Binding Protein</td>
</tr>
<tr>
<td>FGF</td>
<td>Fibroblast Growth Factor</td>
</tr>
<tr>
<td>FOXC</td>
<td>Forkhead Box C</td>
</tr>
<tr>
<td>FOXG1</td>
<td>Forkhead Box G1</td>
</tr>
<tr>
<td>GDF</td>
<td>Growth Differentiation Factor</td>
</tr>
<tr>
<td>GFAP</td>
<td>Glial Fibrillary Acidic Protein</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescent Protein</td>
</tr>
<tr>
<td>GLAST</td>
<td>Glutamate/Aspartate Transporter (also SLC1A3)</td>
</tr>
<tr>
<td>HES</td>
<td>Hairy Enhancer of Split</td>
</tr>
<tr>
<td>HEY1</td>
<td>HES Related BHLHL Transcription Factor with YRPW Motif 1</td>
</tr>
<tr>
<td>HOPX</td>
<td>HOP Homeobox</td>
</tr>
</tbody>
</table>
ID  Inhibitor of DNA Binding
IRES  Internal Ribosomal Entry Site
JAK  Janus Kinase
kb  kilobases
LBX  Ladybird Homeobox
LHX  LIM Homeobox
LIF  Leukaemia Inhibitory Factor
LIN28  LIN-28 Homolog
LIMX  LIM Homeobox Transcription Factor
Meox1  Mesenchyme Homeobox
MEFs  Mouse Embryonic Fibroblasts
miR  micro RNA
MMD2  Monocyte to Macrophage Differentiation Associated
MSX  Msh Homeobox
NEUN  RBFOX3 (RNA Binding Fox-1 Homolog 3)
NFI  Nuclear Factor I
NGN  Neurogenin
NICD  Notch Intracellular Domain
NKX  NK Homeobox
NMP  Neuromesodermal Progenitor
NPAS3  Neuronal PAS Domain Protein 3
OLIG  Oligodendrocyte Transcription Factor
OPC  Oligodendrocyte Progenitor Cell
PAX  Paired Box
RELN  REELIN
RBPJ  Recombination Signal Binding Protein for Immunoglobulin Kappa J Region
SCL  Stem Cell Leukaemia (also known as TAL1)
SHH  Sonic Hedgehog
SLC1A3  see GLAST
SLIT1  Slit Guidance Ligand 1
SMAD  SMAD Family Member
SOX  SRY-box
SP8  Sp8 Transcription Factor
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAT</td>
<td>Signal Transducer and Activator of Transcription</td>
</tr>
<tr>
<td>T/BRA</td>
<td>BRACHYURY</td>
</tr>
<tr>
<td>TGF-β</td>
<td>Transforming Growth Factor-beta</td>
</tr>
<tr>
<td>TUBB3</td>
<td>Tubulin Beta 3 Class III</td>
</tr>
<tr>
<td>TWIST1</td>
<td>Twist Family BHLH Transcription Factor 1</td>
</tr>
<tr>
<td>VEGF</td>
<td>Vascular Endothelial Growth Factor</td>
</tr>
<tr>
<td>WNT</td>
<td>Wnt Family Member</td>
</tr>
<tr>
<td>ZCCHC24</td>
<td>Zinc Finger CCHC-Type Containing 24</td>
</tr>
<tr>
<td>ZFP</td>
<td>Zinc Finger Protein (also called ZNF)</td>
</tr>
</tbody>
</table>
Chapter 1. Introduction

Embryonic development is a remarkable process. A single cell, the zygote, contains the requisite information to generate a fully-functional multicellular organism. Before development has finished it will go through a number of biological processes including cell division, growth, patterning and differentiation. The zygote’s descendants will integrate both intrinsic information and extrinsic cues to orient themselves in time and space in order to generate organised tissues and organs. Understanding the mechanistic detail underlying these processes has been the ambition of developmental biology.

1.1 The developmental origin of the central nervous system

Gastrulation is the process by which the unpatterned blastula is converted into an organized multi-layered structure. During mouse gastrulation ingress of cells through the primitive streak results in the formation of the three primary germ layers: endoderm, mesoderm and ectoderm (Stern 2005, Kojima et al. 2014). The dorsal-most layer, the ectoderm, will give rise to the epidermis, neural plate, and neural crest. Adjacent to the ectoderm, at the dorsal tip of the primitive streak, is a structure known as ‘the organizer’ or ‘the node’ (Spemann 1921, Spemann and Mangold 1924, Waddington 1936, Stern 2005). The organizer instructs the formation of the neural plate (Levine and Brivanlou 2007, Anderson and Stern 2016). The remaining ectoderm will generate the epidermis. The neural plate grows and folds inwards, the outermost regions coming together to form the neural tube. This process is called neurulation (Schoenwolf and Smith 1990).

The neural tube is the embryonic structure which will generate the entire adult central nervous system (CNS) with the most anterior regions forming the brain and the more posterior segment giving rise to the spinal cord. Until recently, it was thought that posterior regions arose via patterning of the neural plate. This hypothesis was first formulated by Nieuwkoop and termed ‘activation-transformation’ (Nieuwkoop and Nigtevecht 1954). In this model, the neural lineage is first activated and then patterned into region-specific domains (Figure 1).
However, recent evidence has shown that anterior and posterior segments have distinct developmental origins. Lineage tracing experiments have shown that the spinal cord shares a common ancestry with paraxial mesoderm lineages (Brown and Storey 2000, Tzouanacou et al. 2009). These neuromesodermal progenitors (NMPs), characterised by the expression of T/BRA and SOX2, are located in the caudal lateral epiblast (CLE) and the node-streak border (Wilson et al. 2009). NMPs retained in this region during body axis elongation undergo progressive posteriorisation. FGF and WNT signals induce the expression of CDX transcription factors and the sequential activation of \( Hox \) genes. As cells exit the CLE they acquire neural or paraxial mesoderm fates but retain their specific \( Hox \) code. This results in the patterning of the anterior-posterior axis. These data challenged the ‘activation-transformation’ hypothesis and the prevailing view of neural induction (Metzis et al. 2017).

![Figure 1](image)

**Figure 1** The ‘activation-transformation’ hypothesis

A schematic of the ‘activation-transformation’ hypothesis. The neural lineage is activated in ectoderm tissue, becoming the prospective forebrain. Subsequently, a second signal progressively patterns cells into mibrain, hindbrain and spinal cord. This figure is adapted from (Stern 2001).

### 1.2 Neurons and glia are the building blocks of the CNS

The CNS is comprised of two major differentiated cell types: neurons and glia. Neurons form the circuitry of the CNS, conveying information between the brain and the periphery. Sensory neurons conduct signals from the periphery to the CNS, whilst motor neurons conduct information from the CNS to the periphery and primarily control muscle contraction. Interneurons form the connections between sensory and motor neurons. Neurons can be identified by their morphology but also by the expression of pan-neuronal markers, for example, TUBB3, DCX, ELAVL3. Neurons can be further partitioned based on their expression of neurotransmitters,
cell morphology and biological position. During development neuronal subtype specification is under both spatial and temporal control (Jessell 2000).

Glia, meaning “glue” in Ancient Greek, support neurons and perform homeostatic functions. Although they were discovered in 1856 by Rudolf Virchow, only recently are we beginning to appreciate the diversity of subtypes and their functions (Kettenmann and Verkhratsky 2008). In general, there are 3 major classes in the CNS: astrocytes, oligodendrocytes and microglia. Microglia are specialised macrophages that perform immunological functions. These originate in the yolk sac and enter the nervous system from the periphery at mid-gestation (Ginhoux et al. 2010). Oligodendrocytes are the cells which myelinate neuronal axons, forming an insulating sheath which allows electrochemical signals to propagate more efficiently. Markers include proteins specific for their function such as MBP, PLP, O4 (Bradl and Lassmann 2010). Astrocytes regulate CNS homeostasis (Chen and Swanson 2003). They perform several functions essential for neuronal activity, such as water and glutamate transport, and ion buffering (Allen and Lyons 2018). Astrocytes also regulate the blood-brain barrier (Cabezas et al. 2014). Unlike neurons and oligodendrocytes, astrocytes are less well defined. Classically, they were distinguished histologically into either protoplasmic (located in grey matter) or fibrous (in white matter) populations (Andriezen 1893, Kölliker 1896, Ramón y Cajal 1909). More recently, GFAP, S100B, AQP4, have all been used to identify astrocytic populations (Roots 1981, Nielsen et al. 1997). However, further work is required to delineate astrocytic subpopulations.

Similar to neurons, the production of glial subtypes during development is also under spatial control (Briscoe and Small 2015). However, in all regions of the neural tube glial production follows neuron differentiation. In the cortex, neuronal production begins at E11.5 and in the spinal cord at E9, whilst gliogenesis begins at E16.5 in the cortex and E12.5 in the spinal cord (Rowitch 2004, Miller and Gauthier 2007, Guérout et al. 2014). How the timing of the neural-to-glial transition is controlled is not well understood, although mechanisms have been proposed (Toma et al. 2016).
1.3 Timing mechanisms in development

How progenitors generate specific cell types with temporal precision is a fundamental question in developmental biology. Neurons and glia across all regions of the CNS are generated sequentially and with precise timing. A temporal delay in the neural-to-glial transition in the developing cortex relative to other regions allows the increase in total neuron production necessary for cognitive function. The expansion of the cortex is one of the evolutionary mechanisms thought to have increased the cognitive abilities of humanoids (Rakic 2009). Therefore, temporal control has important functional and evolutionary consequences.

In biological systems, time can be encoded at 3 levels: systemically, at the tissue level, or cell intrinsically. In this section we discuss timing mechanisms, focussing on examples present in neural development.

1.3.1 Systemic Control

Systemic cues have been proposed to control developmental transitions both in vertebrates and invertebrates. In *Drosophila*, the hormone ecdysone produced by the prothoracic gland, is the master regulator of molting and metamorphosis (Yamanaka et al. 2013). Ecdysone release is triggered by prothoracicotrophic hormone, a neurotransmitter secreted by the brain in response to various environmental and developmental cues. This suggests a mechanism where developmental transitions can be systemically controlled. Recent work describes how ecdysone modulates a developmental transition in mid-larval neural stem cells which is essential for generating proper neuronal and glial cell diversity (Syed et al. 2017). Whether a similar mechanism occurs in mammals is not clear. In the hindbrain perfused vascular networks are formed by E10.5, although whether either the presence of vasculature or hormones circulating in the blood instruct neural progenitor development is not known (Fantin et al. 2013, Tata et al. 2015).

Temperature and metabolism have been implicated in the timing of decision making. For mammals, where gestation is kept at ~37°C, metabolism is a more
relevant parameter. Several metabolic switches occur during early embryogenesis (Folmes and Terzic 2014). Mitochondria are inherited from the mother and at early stages metabolism is oxidative. By the blastocyst stage metabolism has been remodelled due to the increased availability of glucose and is largely glycolytic. Following implantation there is a gradual transition back to oxidative metabolism due to the onset of mitochondrial biogenesis (Folmes and Terzic 2014). Mutations to several metabolic enzymes have been described to have a developmental delay (Johnson et al. 2003). Nevertheless, the requirement for metabolic changes at later stages of embryogenesis is not well understood. In heart formation the development of mitochondrial networks is essential for cardiomyocyte differentiation whilst in the retina the transition from progenitor proliferation to differentiation is associated with a conversion from glycolytic to oxidative metabolism (Agathocleous et al. 2012, Folmes et al. 2012). However, it is unclear to what extent metabolism instructs developmental decisions.

Another potential mechanism of systemic control is the cell intrinsic circadian clock. Circadian rhythms are ~24 hour oscillations in gene expression, physiology and behaviour. In adults they are corrected by photoreceptors in the eye, responding to the light-dark cycle. Oscillating circadian clock genes emerge during development (Dekens and Whitmore 2008, Yagita et al. 2010, Wreschnig et al. 2014). Although no mechanistic link has been established, their appearance coincides with the termination of neuron differentiation in the hypothalamus (Wreschnig et al. 2014). Reprogramming neural progenitors to embryonic stem cells abolishes circadian oscillation, suggesting the underlying mechanism is cell intrinsic.

### 1.3.2 Tissue Level Control

One argument against systemic control of developmental timing is the development of *in vitro* methods for differentiating embryonic stem cells. These pluripotent cells can be directed to different cell types with high efficiency. In these experiments, the temporal sequence of developing progenitors is comparable to their *in vivo* counterparts (Gaspard et al. 2008). Three-dimensional aggregation of pluripotent cells in combination with differentiation cues can generate organoids. In this
system, not only is the temporal generation of cell types preserved, but also their spatial organisation (Eiraku et al. 2008, Kadoshima et al. 2013, Clevers 2016). This argues for timing being controlled either cell autonomously or at the tissue level.

Evidence for tissue level control of developmental timing initially came from heterochronic transplantation experiments (McConnell and Kaznowski 1991, Desai and McConnell 2000). Developmentally young neural progenitors transplanted into an older host cortex adapt to the ‘aged’ environment generating later-born neurons. However older progenitors were not capable of generating early-born neuronal subtypes when transplanted into younger hosts. Therefore, although the environment can instruct progenitor maturation it cannot overcome the gradual restriction of progenitor competency. Evidence that the local environment is sufficient to overcome systemic signals comes from transplantation of chimeric zebra finch cortices into quail (Chen et al. 2012). Birds with vocal learning abilities such as zebra finch have larger brains relative to their non-learning counterparts. Donor cortices were still enlarged in their quail host. This argues that environmental influence of developmental timing occurs on a local and not systemic length scale.

Recent work has begun to deconvolve the mechanism by which environmental cues affect developmental decisions. In the cortex, neuronal subtypes are generated sequentially: subplate neurons are produced first, deep-layer subcortical neurons follow and upper-layer intracortical projection neurons last (Figure 2) (Molyneaux et al. 2007). Specific ablation of deep layer neurons delayed a progenitor transition from deep layer to upper layer neuron production (Toma et al. 2014). Therefore, differentiated cell types control the timing of temporal progression. The precise mechanism remains unclear, but several signalling pathways have been implicated. Activation of TGF-β signalling in progenitors has been shown to control the progression between two sequential waves of neuronal generation in the developing hindbrain (Dias et al. 2014). The onset of neuron differentiation is characterised by a transition from symmetric to asymmetric cell divisions. NOTCH signalling is activated in neighbouring cells upon neuronal differentiation. The downstream effect of NOTCH activation is to maintain the progenitor status. However, it has also been shown to induce the expression of
genes controlling developmental progression (Namihira et al. 2009). This suggests a link between neuron differentiation and developmental transitions.

![Figure 2 Cortex development in the mouse](image)

In the ventricular zone (VZ), neuroepithelial cells (NEC) transition to radial glia cells (RGC) before undergoing neuron differentiation at embryonic day 12. RGCs leave the apical surface and enter the subventricular zone before terminally differentiating. Neuron subtypes are generated sequentially and migrate into the cortical plate (CP), taking up residence in an inside-out manner. At E17.5, progenitors start producing astrocytes. This figure is taken from (Bjornsson et al. 2015).

### 1.3.3 Cell Intrinsic Control

Single cell culture of neural progenitors isolated from different stages of the developing mammalian cortex reproduced the temporal generation of cell types observed in vivo (Shen et al. 2006). This highlights the importance of cell intrinsic mechanisms. In the mammalian cortex each neurogenic progenitor goes through a consistent number of cell divisions and produces similar proportions of neuronal subtypes (Takahashi et al. 1999, Gao et al. 2014). This suggested that cell divisions might control temporal progression. However, transient cell cycle arrest of neural progenitors is not sufficient to change developmental timing (Okamoto et al. 2016). At least in neural development, cell divisions appear not to play a role.
Conventional timekeeping devices can either count up or count down. Examples of both have been found in biological systems. In early *Xenopus laevis* development the dilution of four replication factors through rapid cell divisions has been proposed to regulate the timing of the midblastula transition (Collart *et al.* 2013). An example of a count up timer is the proliferation of oligodendrocyte progenitor cells in the rat optic nerve, which can divide a maximum of 8 times before terminal differentiation (Raff 2011). This is in part due to the accumulation of cell cycle inhibitors through progressive cell divisions (Gao *et al.* 1997, Dugas *et al.* 2007). The “accumulation to threshold” hypothesis has received support from other areas. In neural progenitors, HES1 and ASCL1 oscillate out of phase to maintain progenitor status. Gradual accumulation of ASCL1 eventually abrogates HES1 expression and results in neuron differentiation (Imayoshi *et al.* 2013).

Additional cell autonomous mechanisms have been proposed. A linear cascade of transcription factors can shuttle progenitors through sequential developmental transitions. Changing the network topology, for example, by adding repressor loops, can produce variable transition rates. Additionally, modifying the edges, the strength each gene interacts with the next, can alter the pace of temporal progression (Alon 2007). *Drosophila* neuroblasts sequentially express Hunchback, Krüppel, Pdm1, Castor and Grainyhead (Pearson and Doe 2003). Seven-up repression of Hunchback acts as a negative feedback loop (Kanai *et al.* 2005). The downregulation of Hunchback sets in motion a transcriptional cascade that results in the temporal production of differentiated progeny. What triggers Seven-up expression is not known.

There is emerging evidence that transcriptional cascades function in the temporal progression of progenitors in mammals. FOXG1 promotes a transition in progenitor potential in the cortex from generating early to late neuronal cell types (Kumamoto *et al.* 2013, Toma *et al.* 2014). In the spinal cord, the expression of NFI transcription factors induces a transition from neuron differentiation to gliogenesis (Deneen *et al.* 2006). Epigenetic mechanisms may feature as repressor loops. Loss of RING1B, a component of the Polycomb repressor complex, results in an extension of early neuron differentiation (Morimoto-Suzuki *et al.* 2014, Corley and Kroll 2015). What confers the specificity to epigenetic proteins is not clear. Further
work must be done in order to obtain a molecular description of the transcriptional changes in maturing progenitors before the regulatory networks can begin to be defined.

1.4 The development of the spinal cord

In all regions of the neural tube progenitors sequentially produce neurons and glia. In the developing cortex, however, the phase of neuron differentiation is longer. Neurons are first produced at embryonic day 12 (E12) and glial differentiation begins after E16.5 (Miller and Gauthier 2007). In the spinal cord, the transition is earlier, both in developmental time, and in relation to the onset of neuron generation. The start of neuron differentiation is at E9 whilst gliogenesis commences at E11.5 (Deneen et al. 2006). The phase of neuron production is extended in the developing cortex since progenitors sequentially generate subtypes of neurons which migrate away from the progenitor zone to take residence in specific layers of the cortex (Figure 2) (Molyneaux et al. 2007). Both the relative temporal and spatial simplicity, make the developing spinal cord an attractive model system to investigate the transition. In this section we describe the molecular events which define the development of the spinal cord.

1.4.1 Patterning and differentiation of neural progenitors

Following commitment to the neural lineage, progenitors are patterned into regional domains, a process which is well studied. Sonic Hedgehog (SHH) produced by the notochord and floor plate, as well as BMPs and WNTs emanating from the roof plate, are the primary signals responsible for progenitor patterning (Briscoe and Small 2015). In ventral regions, SHH is initially secreted from the notochord, a structure underlying the neural tube. The resulting gradient activates gene expression in a concentration-dependent manner, whilst cross-repressive interactions between transcription factors such as NKX2.2, OLIG2, and PAX6 result in the formation of distinct domains. At later stages the notochord retracts and by ~E10.5 has lost contact with the neural tube. The maintenance of the progenitor pattern is then dependent on continued SHH production from the floor plate (Yu et al. 2013). BMPs and WNTs act to pattern the dorsal domains
of the neural tube (Wilson and Maden 2005). Ectopic activation of the BMP signalling pathway was able to induce the expression of MSX1 and MSX2 in more ventral regions and to repress DBX1 and DBX2, markers of intermediate progenitors (Timmer et al. 2002). However, the underlying transcriptional network interpreting dorsal signalling is not well defined. Nevertheless, antiparallel morphogen gradients result in the formation of 11 distinct progenitor domains (Figure 3).

**Figure 3 The patterning of the neural tube**

SHH secreted from the notochord (NC) and floor plate (FP) and BMP/WNT emanating from the roof plate pattern the neural tubes into 11 distinct progenitor domains (pd1 to p3). These antiparallel morphogen gradients are interpreted by transcriptional networks that restrict progenitor identity. The 11 progenitor domains produce distinct classes of neurons: the pMN domain produces motor neurons (MN) whilst the remaining regions differentiate to different subclasses of interneurons (dI1 to V3). SOX2 expression demarcates the progenitor zone. This figure is taken from (Briscoe and Small 2015).

At E9, neural progenitors begin to leave the ventricular zone and terminally differentiate into neurons. Differentiation begins ventrally before spreading to dorsal domains (Kicheva et al. 2014). Progenitor patterning results in the production of molecularly distinct subclasses of neurons (Figure 3). For example, progenitors expressing OLIG2 produce motor neurons whilst the remaining domains give rise to different classes of interneurons (Novitch et al. 2001). Not only does the OLIG2
domain generate different progeny but the differentiation rate is also higher (Kicheva et al. 2014). This is in part due to OLIG2 repression of Hes5, a transcriptional repressor which inhibits the expression of proneural genes (Sagner et al. 2018).

Hes genes are the canonical targets of NOTCH signalling and form the basis of a molecular network governing the decision between differentiation and self-renewal. The transcriptional output of NOTCH signalling is mediated by a complex of proteins, which includes RBP-J. Deletion of RBP-J in mice results in the loss of the progenitor pool as they prematurely differentiate into neurons (Imayoshi et al. 2010). The level of HES expression is an important parameter in neuronal commitment. HES oscillates out of phase with ASCL1 in progenitors to maintain progenitor status (Imayoshi and Kageyama 2014). Proneural transcription factors, such as NGN1/2 and ASCL1, play an instructive role in directing neuron differentiation and neuronal subtype identity (Guillemot and Hassan 2017). Progressive accumulation of ASCL1 results in complete inhibition of HES and the induction of neuron differentiation (Imayoshi et al. 2013).

Proneural genes not only function to repress NOTCH target genes, but also repress SOX-B1 transcription factors. The B1-subfamily of Sox genes, comprising Sox1-3, are expressed in neural progenitors and negatively regulate neuron differentiation in part by repressing proneural genes (Bylund et al. 2003, Graham et al. 2003, Holmberg et al. 2008). Although NOTCH signalling and SOXB1 proteins appear to have similar functions, their mechanisms of action are distinct. Constitutive expression of SOX3 was able prevent neuron differentiation in the absence of NOTCH signalling. In contrast, activation of NOTCH signalling was not sufficient to maintain progenitor status when a dominant-negative form of SOX3 was constitutively expressed (Holmberg et al. 2008). This suggests a functional hierarchy between the SOXB1 family and NOTCH. Nevertheless, these two mechanisms act in concert to control the balance between neuron differentiation and progenitor maintenance.
1.4.2 NFIA and NFIB are the molecular correlates of gliogenesis

In vivo, neuron differentiation abates around E12.5 when progenitors switch towards a glial fate (Kessaris et al. 2001, Rowitch 2004). Evidence for a transition in progenitor competency comes from transplant studies. OLIG2-positive progenitors isolated from E9.5 neural tubes were capable of generating both neurons and glia when transplanted into the chick neural tube. However, cells isolated from E13.5 embryos only generated glia (Mukouyama et al. 2006). Historically, the onset of glial differentiation was defined based on the temporal expression pattern of the astrocyte-specific glutamate/aspartate transporter, GLAST. Glast expression is first detected in progenitors located ventrally at E11.5 before being expressed in progenitors across the dorsal-ventral axis (Shibata et al. 1997, Ogawa et al. 2005, Deneen et al. 2006).

The transcription factors NFIA and NFIB were identified by comparing the transcriptomes of OLIG2-positive cells isolated from mice at neurogenic (E9.5-E10.5) and gliogenic stages (E11.5-E12.5). The expression pattern of NFIA and NFIB is comparable to Glast, also beginning ventrally at E11.5. Further work went on to suggest that NFIA and NFIB are necessary and sufficient for glial differentiation. Deletion of either NFIA or NFIB severely reduced the number of GFAP-positive glia. It has been suggested that the residual gliogenesis is due to redundancy between the two genes. Constitutive expression of NFIA and NFIB in chick neural tubes was sufficient to induce Glast and early astrocytogenesis (Deneen et al. 2006). In addition, overexpression of NFIA or NFIB are sufficient to reprogram fibroblasts into GFAP-positive glia (Caiazzo et al. 2015). Curiously, the reprogramming efficiency was improved if both NFIA and NFIB were used. This argues against functional redundancy or implicates NFI dosage as being important for glial commitment.

1.4.3 The function of NFIA and NFIB in gliogenesis

NFIA and NFIB are expressed throughout the glial lineage. What is their function during this progression? NFIA functions to activate genes important for terminal gliogenesis. Analysis of transcriptional changes between E11.5 and E12.5 in
OLIG2-positive progenitors identified a cohort of candidate genes induced following NFIA induction. The expression of Apcdd1, Mmd2, Zcchc24, and Hopx were all reduced in NFIA-deficient mice. Interestingly constitutive expression of Mmd2, Zcchc24 or Apcdd1 in chick was sufficient to restore Glast expression when Nfia was inhibited. At later stages Glast expressing cells were observed migrating out into the mantle zone suggesting gliogenesis had been restored (Kang et al. 2012). Whether NFIB plays an analogous role is unclear.

NFIC and NFIX are the remaining two proteins in the NFI family of transcription factors. Of the two, NFIX has been shown to play a role in spinal cord gliogenesis (Matuzelski et al. 2017). NFIX is induced at E13.5 following GLAST, NFIA and NFIB induction. NFIX deficient mice have reduced gliogenesis but the expression of Glast, NFIA and NFIB is unaffected. Nfix is regulated by NFIA and NFIB as deletion of either reduced NFIX expression. Curiously only NFIB was sufficient to activate the Nfix promotor, suggesting an indirect role for NFIA. Therefore, although functional redundancy has been postulated, the mechanism by which they activate downstream targets is molecularly distinct.

1.4.4 The role of NFIA in neuron differentiation

Neuron generation in the developing mouse spinal cord ends at approximately E12.5. NFIA and NFIB expression is induced at E11.5 (Deneen et al. 2006, Matuzelski et al. 2017). The temporal pattern implicates NFIA and NFIB in the termination of neuron differentiation, but the evidence is contradictory. Constitutive expression of NFIA did not affect the expression of neuronal markers. However, inhibition of Nfia resulted in continued neuron differentiation. Reciprocally, the expression of progenitor markers was decreased (Deneen et al. 2006). In contrast, in the developing hippocampus, NFIA deficient mice had an expanded ventricular zone, delayed neuron differentiation, and reduced gliogenesis (Piper et al. 2010). The different phenotypes could be attributed to the function of NFIA being location and/or stage dependent. How NFIA affects neuron differentiation is not clear, but NOTCH signalling could be involved. Activation of NOTCH signalling results in the inhibition of neuron differentiation. Combined repression of Nfia, Nfib, Nfic, and Nfix
was sufficient to activate the NOTCH target Hey1 in glioblastoma (Brun et al. 2018). Additionally, Hey1 along with Hes1 and Hes5 were all upregulated in the hippocampus of NFIA-deficient mice (Piper et al. 2010). Whether NFIA is merely modulating the NOTCH response or abrogating NOTCH signalling is not clear. Astrocytogenesis appears to require NOTCH signalling. Abrogation of NOTCH signalling either before or after NFIA induction returned GFAP phenotypes, suggesting NOTCH is required at multiple stages of progenitor development. Only after E13 was gliogenesis independent of NOTCH (Kang et al. 2013). This coincides with the migration of progenitors away from the ventricular zone. Activation of the NOTCH pathway also caused demethylation of the Gfap promoter (Namihira et al. 2009). The latter was not due to a change in expression of DNA methyltransferases, the proteins that control DNA methylation status. Instead, NFI was found to bind the Gfap promoter and removal of NFIA inhibited its demethylation by NOTCH. This was also the case for other astrocytic genes such as S100β and Aqp4 (Namihira et al. 2009). Therefore, not only does NFIA activate downstream genes but it plays a central role in activating astrocytic markers.

1.4.5 Terminal glial differentiation

At E13.5, following NFI induction, progenitors begin to migrate away from the ventricular zone and terminally differentiate. The OLIG2-positive domain produces the myelinating glia oligodendrocytes, whilst the remaining regions generate astrocytes. There are at least three spatially distinct regions of astrocytes, defined by their expression of RELN and SLIT1 (Hochstim et al. 2008). These subtypes arise from progenitor domains expressing PAX6 and NKX6.1, respectively, and these transcription factors direct astrocyte specification. It was further shown that NKK2.2 was necessary for the expression of SLIT1 (Genethliou et al. 2009). Terminal differentiation is characterised by two waves of cell division. The second wave, occurring after the progenitors have left the ventricular zone, occurs in a progressive ventral to dorsal pattern (Tien et al. 2012, Molofsky et al. 2013). Fate mapping experiments using conditional reporters induced by various dorsal-ventral markers showed that astrocytes take residence in spatial domains that reflected
their embryonic region of origin (Tsai et al. 2012). Therefore, patterning plays an instructive role in glial subtype specification and spatial organisation.

Terminal astrocytogenesis requires JAK-STAT signalling. Upon stimulation by cytokines, such as LIF, CNTF and Cardiotrophin-1, the STAT receptor activates JAK which in turn phosphorylates STAT, allowing it to translocate to the nucleus and affect transcription (Figure 4) (Levy and Darnell Jr 2002). Although the source of the ligand is not certain, evidence from the cortex points towards it being secreted from neurons. Conditioned medium from wild-type, but not Cardiotrophin-1 knockout, neuronal cultures was sufficient to induce gliogenesis in vitro (Barnabé-Heider et al. 2005). In addition, activation of JAK-STAT signalling or constitutive expression of STAT3 is sufficient to induce early gliogenesis (Barnabé-Heider et al. 2005, Hong and Song 2014). Deletion of the receptor cofactor gp130 or the LIF receptor results in a loss of GFAP expression in E18.5 spinal cords (Koblar et al. 1998, Nakashima et al. 1999). At this time, phosphorylated-STAT3 is detected in the mantle zone and appears coincident with the glia markers GFAP and S100β. Conditional deletion of STAT3 (but not knockout of STAT1) results in a loss of terminal gliogenesis. The presence of STAT binding sites in close proximity to the Gfap promotor suggest the effect is direct (Namihira et al. 2009). STAT3 activation was also sufficient to increase the proportion of NFIA-positive cells (Hong and Song 2014). Whether this is due to direct regulation of NFIA by STAT3 or an indirect effect of inducing the second wave of cell division is not clear. Nevertheless, at later stages, activation of JAK-STAT signalling is required for the terminal differentiation of glia.
Ligand binding to the STAT receptor results in the activation of JAK and the phosphorylation of STAT. This allows STAT to translocate to the nucleus where it affects transcription. This figure is taken from (Levy and Darnell Jr 2002).

**Figure 4 The canonical JAK-STAT signalling pathway**

1.4.6 **Oligodendrogenesis is controlled by a distinct regulatory network**

Following motor neuron production, oligodendrocytes are produced from the OLIG2-positive motor neuron domain. OLIG1/2 are essential for the generation of oligodendrocytes as their deletion results in astrocyte production (Lu et al. 2002, Takebayashi et al. 2002, Zhou and Anderson 2002). Deletion of SCL, expressed in the p2 progenitor domain, expands OLIG2 expression dorsally and concomitantly increases oligodendrocyte production (Muroyama et al. 2005). This is due to OLIG2’s ability to neutralise SOX9/NFIA activation of the astrocytic lineage (Glasgow et al. 2014). Therefore, OLIG2 has both neurogenic and gliogenic functions. The phosphorylation status of OLIG2 has been proposed to mediate the transition between motor neuron and oligodendrocyte production. At later stages, OLIG2 is dephosphorylated, destabilising homodimerization and promoting heterodimerization with proneural transcription factors such as NGN2 (Li et al. 2011). Whether this is a general mechanism underlying the transition at other dorsal-ventral positions is not clear. SOX10 is expressed from E11.5 in the motor
neuron domain, and its deletion results in a fate change from oligodendrogenesis to astrocytogenesis (Stolt et al. 2005). SOX10 directly interacts with NFIA, antagonizing its pro-astrocytic function (Glasgow et al. 2014). Inhibition of Nfia abrogated OLIG2 expression (Deneen et al. 2006). Therefore, a distinct regulatory network diverts progenitors from astrocytogenesis to generating myelinating oligodendrocytes.

1.5 A gene regulatory network controls the timing of NFIA and NFIB induction and terminal gliogenesis

In the spinal cord, similar to the rest of the embryonic neural tube gliogenesis follows neuron differentiation. Our understanding of the underlying network controlling the timing of this transition is limited. NFIA and NFIB are the molecular correlates of the glial lineage. Therefore, understanding the timing of gliogenesis revolves around defining the mechanism of NFIA and NFIB induction. As discussed earlier the timing of developmental decisions can be under systemic, tissue level, or cell autonomous control. In the upcoming sections we review what is known about NFIA and NFIB regulation in the spinal cord. We begin with cell autonomous mechanisms before discussing other possibilities.

1.5.1 SOX9 plays a central role in gliogenesis and the timing of NFIA and NFIB induction

How the switch from neuron differentiation to gliogenesis is initiated is not well understood but SOX9 has been implicated (Kang et al. 2012). SOX9 expression begins in progenitors located ventrally at E9.5 before spreading throughout the dorsal-ventral axis (Stolt et al. 2003). This coincides with the onset of neuron differentiation and precedes NFIA and NFIB expression. Like NFIA and NFIB, SOX9 is expressed throughout the glial lineage and has multiple roles in gliogenesis. SOX9 is essential for the formation of neural stem (NS) cells. NS cells are defined by their ability to self-renew whilst maintaining their potential to generate neurons and glia. They can only be isolated from the neural tube from E10.5 after SOX9 expression (Scott et al. 2010). Consistent with this, SOX9
deficient mice have reduced gliogenesis (Stolt et al. 2003). Whether this is a consequence of aberrant glial lineage activation or SOX9 having multiple roles throughout gliogenesis is not clear. Although NFIA and NFIB are sufficient to reprogram fibroblasts into GFAP-positive glia, the addition of SOX9 to the reprogramming cocktail increases the efficiency (Caiazzo et al. 2015). Together these data point towards SOX9 being critical for gliogenesis.

SOX9 has been shown to be important for the correct timing of NFIA expression. Conditional deletion of SOX9 results in an approximately 24 hour delay in NFIA and Glast (Kang et al. 2012). In these experiments, SOX9 deletion was under the control of a Nestin-Cre which is not expressed until E10.5. Therefore, residual SOX9 expression is observed at E10.5 (Stolt et al. 2003). It is difficult to determine whether the 24 hour delay is a result of the technical limitations of the experiment or due to the existence of alternative mechanisms responsible for NFIA induction. Constitutive expression of SOX9 in chick was sufficient to induce NFIA. Moreover, a dominant repressor form of SOX9, Sox9-En, inhibited the expression of NFIA and Glast. Glast was restored upon co-expression with NFIA suggesting SOX9 is upstream and functions via NFIA (Kang et al. 2012). SOX9 has been suggested to regulate Nfia via a conserved noncoding element, e123, 5' of the transcriptional start site (TSS). Whether this region is required for the correct timing of Nfia expression is unclear. Nevertheless, these data support a model where SOX9 directly regulates Nfia and is required for the correct timing of its expression. It is not clear if a similar relationship exists between SOX9 and NFIB.

Following NFIA induction, the glial lineage is activated by the induction of target genes such as Apced1, Mmd2 and Zcchc24. NFIA is required for target gene activation and their overexpression is sufficient to restore gliogenesis in the absence of Nfia (Stolt et al. 2003). SOX9 deficient mice had reduced Apced1, Mmd2, Zcchc24 expression at E12.5. The phenotype does not seem to be a result of delayed NFIA expression but instead SOX9 is directly involved. SOX9 and NFIA coimmunoprecipitate and only their combined expression was sufficient to activate the promoter regions of target genes (Kang et al. 2012). Consistent with this, constitutive expression of NFIA was not sufficient to activate downstream target
genes when co-expressed with Sox9-En. These data demonstrate that SOX9 and NFIA function collaboratively to activate the glial lineage.

In the developing cerebellum, a different role for SOX9 has been proposed. Specific deletion of SOX9 was found not to affect glial cell specification but led to prolonged neuron differentiation (Vong et al. 2015). This suggests SOX9 is important for shutting down neuron differentiation and argues for an uncoupling of the neurogenic and gliogenic programmes. In the spinal cord, conditional deletion of SOX9 resulted in a transient increase in motor neuron differentiation. V2 interneurons are generated from the progenitor domain immediately dorsal to the motor neuron domain. V2 interneurons were also increased in SOX9 deficient spinal cords. The proportion continued to increase to E16.5 suggesting continued neuron differentiation. This was not a consequence of progenitor identity switching, as domain sizes were comparable between genotypes (Stolt et al. 2003). Therefore, SOX9 expression is required for the timely termination of neuron differentiation. Whether this is direct or a consequence of delayed glial gene induction is not clear. Nevertheless, SOX9 appears to couple the termination of neuron differentiation to the onset of gliogenesis. What controls SOX9 dynamics is not well understood.

1.5.2 BRN2 activates NFIA with SOX9 but via distinct mechanisms

In addition to SOX9, the transcription factor BRN2 (also known as POU3F2) has been implicated in NFIA induction. Similar to SOX9, BRN2 expression begins at approximately E9.5 shortly after neuron differentiation and is maintained to at least E13.5 (Lupu et al. 2008, Jin et al. 2009, Glasgow et al. 2017). BRN2 was first shown to regulate Nestin and Fabp7, two genes with a similar timing of expression to SOX9 (Josephson et al. 1998). BRN2 deficient mice have a 24 hour delay in NFIA expression, although Glast expression is unaffected. The delay appears comparable to SOX9 deficient embryos. Combined deletion of SOX9 and BRN2 resulted in a further reduction in NFIA expression at E12.5 (Glasgow et al. 2017). The neural tubes of these mice appeared smaller, suggesting neuron differentiation and/or progenitor proliferation could also be affected.
BRN2 has been reported to directly regulate NFIA expression via a conserved noncoding element, e161, located in the first intron of Nfia (Glasgow et al. 2017). Curiously SOX9 was unable to regulate e161 and conversely BRN2 was unable to activate e123, the element which SOX9 regulates. Therefore, SOX9 and BRN2 were proposed to regulate Nfia via distinct mechanisms although they converge to affect its expression. At E12.5, SOX9 and BRN2 coimmunoprecipitate and chromatin conformation capture experiments showed that e161, e123 and the Nfia TSS co-interact (Glasgow et al. 2017). This association was detectable at E10.5 prior to NFIA induction and before SOX9 and BRN2 had engaged with e123 and e161. In fact, the chromatin organisation around Nfia is preserved in motor neuron populations where NFIA but not SOX9 or BRN2 are expressed (Glasgow et al. 2017). This suggests that SOX9 and BRN2 are not required for the organisation of the Nfia locus and the decision to transition to gliogenesis is made in advance of Nfia transcription. How chromatin organisation and engagement is controlled is not clear.

1.5.3 Multiple roles for micro RNAs in the neural-to-glial transition

Several lines of evidence highlight a role for micro RNAs (miRNA) in the neural-to-glial transition, although the data appears contradictory. Following their transcription, miRNAs are processed by DROSHA in the nucleus and transformed into mature miRNAs in the cytoplasm by DICER (Jiang and Yan 2016). Conditional deletion of DICER in ventral regions of the developing neural tube at E11.5 resulted in a loss of gliogenesis ventrally. Intriguingly DICER knockout mice did not have prolonged or ectopic neuron differentiation arguing that neuron differentiation and gliogenesis can be mechanistically uncoupled (Zheng et al. 2010). This is in agreement with the role of SOX9 in the cerebellum where specific deletion of SOX9 led to prolonged neuron differentiation without affecting gliogenesis (Vong et al. 2015).

Mature miRNAs largely function by binding the 3' UTR of messenger transcripts, inhibiting their translation (Cai et al. 2009). A global screen identified 10 mis-
regulated miRNA’s in neurospheres where NFIA is overexpressed. Constitutive expression of miR-124, miR-153, miR-219 all significantly decreased the number of differentiated glia (Tsuyama et al. 2015). Curiously, no miRNAs were found to promote gliogenesis, apparently contradicting the DICER knockout phenotype. miR-153 was found to directly regulate Nfia and Nfib and its downregulation is required for NFIA and NFIB expression (Tsuyama et al. 2015). This points towards a model where the neural-to-gliial transition is actively repressed. miR-124 is expressed in differentiating neurons in the adult brain and regulates the balance between progenitor maintenance and neuron differentiation (Cheng et al. 2009, Tsuyama et al. 2015). It functions by targeting Sox9. Constitutive expression of Sox9 without its 3’ UTR was sufficient to overcome the effects of miR-124. It is unclear whether this network is directly affecting SOX9 mediated NFIA induction and the neural-to-gliial transition.

1.6 Tissue level control gliogenesis

The requirement for signalling cues in neural development is emphasised by the finding that a cocktail of nine signalling effectors was sufficient to reprogram MEFs into functional neural stem cells. Removal of A83-01, an inhibitor of TGF-β signalling, severely affected the reprogramming efficiency of the induction cocktail (Zhang et al. 2016). The TGF-β signalling pathway effects transcription through SMAD proteins (Gaarenstroom and Hill 2014). Ligand binding to the receptor results in the phosphorylation of receptor specific SMAD proteins. Phosphorylation triggers binding of SMAD4, translocation to the nucleus and transcriptional regulation (Figure 5). ID transcription factors are canonical TGF-β targets and play an important role in progenitor maintenance. Electroporation of Smad3 in the chick embryonic spinal cord led to inhibition of ID, activation of proneural bHLH factors, and early neuron differentiation (Garcia-Campmany and Marti 2007).
Figure 5 An overview of Transforming Growth Factor-β signalling

An overview of the TGF-β signalling pathway from ligand binding to transcripting. Ligand binding to either to the BMP or TGF-β receptors result in the phosphorylation of receptor specific SMAD proteins. Upon phosphorylation SMADs complex with SMAD4 before being transported into the nucleus to affect transcription. This figure is taken from (Gaarenstroom and Hill 2014).

GDF11 is a TGF-β ligand which signals through Activin receptors. In the developing spinal cord, it is transcribed in newly born neurons adjacent to the progenitor zone at E10-E11.5 (Shi and Liu 2011). Exposure to GDF11 resulted in a decrease in neuron differentiation and a concomitant increase in oligodendrocyte generation in neurospheres derived from E9.5 spinal cords. In GDF11 deficient mice the rate of neuronal differentiation was reduced. This delayed the differentiation of later born neuronal subtypes (Shi and Liu 2011). This is corroborated by data from the ventral hindbrain where TGF-β was identified as a signal that mediated a switch between two phases of neuron differentiation (Dias et al. 2014). In this system motor neurons are produced first and serotonergic neurons second, after which neuron differentiation abates and oligodendrocyte-like precursors (OLPs) are specified. Overexpression of a constitutively active form of the TGF-β receptor resulted in an
early transition to the second phase. It is not clear whether this an indirect effect of increasing the rate of neuron production. Nevertheless, this discovery was used to develop an in vitro differentiation protocol where a short pulse of TGF-β pathway activation was used to turn on the second wave of neuron differentiation (Dias et al. 2014). This supports a model where TGF-β ligands are secreted by neurons to instruct progenitors. Whether this is simply controlling the rate of neuron production or a developmental transition is unclear. The effect of TGF-β signalling on the neural-to-glial transition has yet to be investigated.

MIB1 is an E3 ubiquitin ligase that is required for the endocytosis of NOTCH ligands. MIB1 deficient mice have defects similar to mice lacking NOTCH signalling. Specifically, embryos have increased neuronal differentiation at the expense of progenitor maintenance (Yoon et al. 2008). Conditional knockout of MIB1 using a Nestin-Cre resulted in a loss of NFIA expression at E12.5 (Kang et al. 2013). Consistent with this, activated NOTCH is sufficient to induce Nfia transcription (Namihira et al. 2009). These data suggest that NOTCH signalling acts upstream of NFIA. Activation of NOTCH1 signalling was sufficient to induce Sox9 expression and astrogliogenesis in vitro. The effect was attributed directly to SOX9 as knockdown of Sox9 reversed the NOTCH1 phenotype (Martini et al. 2013). This might suggest NOTCH activation of NFIA is via SOX9. However, activated NOTCH was found to directly bind the Nfia promotor region suggesting the effect might be direct (Namihira et al. 2009). Whether NOTCH requires SOX9 to activate Nfia expression is not clear.

Multipotent neurospheres can only be isolated from the spinal cord after E10.5, following SOX9 expression in neural progenitors. Activation of SHH signalling promoted neurosphere formation at E9.5 via SOX9 induction (Scott et al. 2010). This is consistent with the expression pattern of SOX9. SHH is produced from the notochord and floor plate, structures located ventrally. SOX9, NFIA and NFIB are all first expressed proximal to the notochord before spreading dorsally (Stolt et al. 2003, Deneen et al. 2006). The primary function of SHH signalling is to pattern the neural tube into discrete domains (Briscoe and Small 2015). This implicates progenitor patterning in the correct timing of the neural-to-glial transition. At later stages floor plate derived SHH is responsible for maintaining the correct neural
progenitor identity and floor plate specific deletion of SHH causes a loss of ventral patterning. Whilst this resulted in the misspecification of neurons and glia it did not affect the timing of SOX9, NFIA and NFIB expression (Yu et al. 2013). This uncouples progenitor patterning and the timing of glial lineage activation. Nevertheless, initial SHH production promotes SOX9 expression and multipotentiality.

1.7 Systemic Cues

As progenitors are generating neurons and transitioning towards the glial lineage, the vascularization of the neural tube is taking place (Tata et al. 2015). This process is best characterised in the developing mouse hindbrain. Vascularization begins at around E9.75, with sprouts appearing from the perineural space and growing towards the ventricular zone. By E10 vessels extend sprouts orthogonally, eventually connecting with neighbouring vessels. Perfused vascular networks appear from E10.5 onwards (Fantin et al. 2013). VEGF secretion from neural progenitors is required to direct vascularization (Raab et al. 2004). It is unclear whether this instructive relationship between neural progenitors and the vasculature is reciprocated. Nevertheless, the vasculature may provide a mechanism, either directly or through exposure to circulating hormones, for the systemic control of the neural-to-glial transition.

1.8 Elucidating the mechanisms behind the neural-to-glial transition

1.8.1 In vitro approaches

In vitro methods used to study gliogenesis include using Neural Stem (NS) cells derived from the neural tubes of E14.5 mice. NS cells are defined by their ability to self-renew long-term using a combination of FGF and EGF, whilst retaining their potential to differentiate into both neurons and glia. NS cells have been generated from all axial levels including the spinal cord and retain their anterior-posterior identity in vitro (Kelly et al. 2009). Neuronal differentiation is initiated upon growth
factor withdrawal and in some cases, with the addition of neurotrophic factors. Glia differentiation is initiated upon the addition of serum or BMP (Conti et al. 2005, Pollard et al. 2006). At early stages in vivo BMP2/4/7 are secreted from the roof plate (Bond et al. 2012, Le Dreau and Marti 2013). Whether they affect the most ventral domains and operate at later stages is not clear.

Embryonic stem cells can be utilised to provide insight into how developmental processes such as the neural-to-glial switch are controlled. Protocols to derive neural progenitors are well established and a commonly used protocol produces cells with an anterior-dorsal identity (Ying et al. 2003). The addition of specific signals can program cells towards different regional subtypes. Retinoic acid can be used to induce more posterior gene expression programs giving neural progenitors with hindbrain identity. Activation of SHH signalling biases cells towards a ventral phenotype (Jessell 2000, Gaspard et al. 2008, Gouti et al. 2014). The timing of sequential neuron differentiation in these cultures was comparable to their in vivo counterparts (Gaspard et al. 2008). This led to the suggestion that developmental timing was controlled cell intrinsically.

Recently developed methods have allowed the generation of neural progenitors with a spinal cord identity from embryonic stem cells (Gouti et al. 2014, Turner et al. 2014, Lippmann et al. 2015, Tsakiridis et al. 2015). During this process cells transition through a Neuro-Mesodermal Progenitor that has the potential to give rise to both posterior neural tissue and paraxial mesoderm lineages. This process recapitulates in vivo development where clonal analysis has shown that posterior neural progenitors and paraxial mesoderm share a common precursor (Tzouanacou et al. 2009). Recent work has focussed on the ability of these neural progenitors to generate neurons but their capacity to make glia has yet to be investigated.

1.8.2 Single cell sequencing and pseudotemporal reconstruction

Single-cell sequencing is a recently developed technique that captures the transcriptome of individual cells (Tang et al. 2009). Although most studies have
aimed at generating an atlas of cell types, by harvesting samples through time it is possible to reconstruct lineage specific transcriptional dynamics. This has led to insight into the mechanism underlying lineage decision making and the identification of novel cell states (Papatsenko et al. 2015, Sagner et al. 2018). A number of methods for obtaining the transcriptomes of single cells now exist. In general, they vary in their output and accuracy. For example, Drop-seq can profile large numbers of cells by encapsulating them in aqueous droplets (Macosko et al. 2015). Smart-seq on the Fluidigm C1 uses microfluidics to isolate single cells (Picelli et al. 2013). This approach captures fewer cells but with greater transcriptional accuracy (Ziegenhain et al. 2017). The development of single cell transcriptomics has precipitated the development of analysis pipelines capable of deconvolving large datasets. In general, these pipelines begin by cleansing and normalising the data and use a selection of genes along with dimensionality reduction techniques to order cells in ‘pseudotime’ (Wu et al. 2017). This allows the identification of a list of candidate genes involved in developmental decisions. The next hurdle is finding high-throughput molecular tools to manipulate these genes and characterise their functions.

1.9 Aims

Our mechanistic understanding of the neural-to-glial transition is limited. Whilst SOX9, NFIA and NFIB have been identified as critical nodes in the network governing glial competence, the pathways controlling their expression are unclear. In this study we focus on providing new mechanistic insight into the acquisition of glial fate. We began by building on a differentiation protocol that differentiates embryonic stem cells into neural progenitors with a spinal cord identity (Gouti et al. 2014). As progenitors mature, they undergo transcriptional changes indicative of a cell transitioning through the neural-to-glial transition. At later stages, cells express glial specific genes including the differentiated marker GFAP. The in vitro protocol was then harnessed to investigate the transition. Specifically, our aims were to assess the functional redundancy between NFIA and NFIB, to characterise the role of TGF-β signalling in the transition, and to identify transcription factors which effect the SOX9, NFIA and NFIB gene regulatory network. To investigate the latter, we
leveraged single cell sequencing and pseudotemporal reconstruction to identify a list of candidate genes. To gain mechanistic insight we used existing molecular tools such as CRISPR-Cas9 and developed an existing lentiviral system. Taken together, our data provides insight into the mechanism underlying the neural-to-glial fate transition.
Chapter 2. Materials and Methods

2.1 Tissue Culture Techniques

All tissue culture was performed in a Class II Laminar flow hood using aseptic technique. Unless otherwise stated all centrifugation steps are at 1,000 rpm (~0.3 xg), for 4 minutes on a bench top centrifuge.

2.1.1 Culture of MEFs

Mouse Embryonic Fibroblasts (MEFs) were isolated from E12.5 mouse embryos. Extraembryonic tissue, head and liver were removed. Embryos were washed in PBS before being passed through a syringe with an 18 Gauge needle 3x to dissociate the tissue. Cells from 3 embryos were plated into one 10 cm plate in MEF medium (Dulbecco’s Modified Eagle Medium (DMEM) containing 15% FBS, 1X GlutaMAX, 1X Pen/Strep (all Gibco)). Media was replenished every 48 hours. Once confluency was reached cells were dissociated with 0.25% Trypsin/EDTA (Gibco) and split 1:3 (Passage 1 (P1)). After P1 MEFs reached confluency the cells were dissociated and frozen x1 10 cm plate per vial in FBS (Gibco) + 10% DMSO (Sigma) until further use. To prepare feeders one vial of frozen MEFs was thawed into 4 10 cm plates. Cells were split 1:4 upon reaching confluency. This was repeated for the next passage whilst the two subsequent passages were carried out at 1:3. MEFs were mitotically inactivated with mitomycin C (Sigma) (see below).

2.1.2 Preparation of feeders

To prepare mitomycin C, 2 mg was dissolved in 2 mL PBS (Gibco) using a needle and syringe. The 2 mL mitomycin C solution was diluted in 18 mL MEF medium. This 10X stock was frozen down until required. To treat MEFs with mitomycin C the stock solution was diluted 1:10 in MEF medium and 5 mL added to a 10 cm dish of confluent MEFs. Cells were incubated at 37°C for 2.5-3 hours, after which the mitomycin C containing medium was aspirated and the cells washed with PBS. Cells were harvested using 0.25% Trypsin/EDTA and resuspended after centrifugation in MEF medium to a concentration of 5 x 10^6 cells per mL. 0.5 mL of
the cell suspension was transferred to a cryovial and mixed with 0.5 mL FBS + 20% DMSO to obtain 2.5 x10^6 cells per vial. Cells were frozen until further use (Feeders). Each batch of feeders was tested for their ability to support embryonic stem (ES) cells before use.

2.1.3 Thawing feeders for ES cell culture

A single vial of 2.5x10^6 cells was used to cover 80 cm^2 of tissue culture dishes although this varies between batches. Feeders were thawed at 37°C and immediately transferred to a centrifuge tube containing 5 mL MEF medium. The cryovial was washed with an additional 1 mL MEF media. Following centrifugation cells were aliquoted into tissue culture dishes coated for at least 30 minutes at 37°C with 0.1% gelatine (Sigma). Feeders should evenly cover the culture surface (note: it takes 48 hours for feeders to properly spread). Cells were fed 24 hours after thawing with MEF media. 48 hours after thawing feeders are ready to support embryonic stem (ES) cell growth. One batch of thawed feeders were used for up to two weeks although they were fed with new medium regularly and checked for signs of necrosis.

2.1.4 Culture of ES cells

The cell line HM1 (Magin et al. 1992) was cultured in pluripotent ES cell conditions (Evans and Kaufman 1981). Prior to thawing, the feeder layer was washed with PBS and ES medium (KO-DMEM + 10% FBS + 1X GlutaMAX + 1X Pen/Strep + 2-mercaptoethanol (all Gibco)) containing 1000 U/mL ESGRO Leukaemia Inhibitory Factor (Merck Millipore) (ES+LIF) added. ES cells were thawed using the same methodology as for MEFs but using ES media. ES cells were frozen at twice the density required for use (see below) meaning 1 vial of ES cells was thawed per dish. Every 24 hours cells were washed with PBS and fed with new ES+LIF. 48 hours after thawing cells were dissociated using Accutase, centrifuged and split onto new feeders at 25,000 cells per cm^2 to split in 2 days or 10,000 cells per cm^2 to split in 3 days. ES cells were frozen comparably to feeders but in a medium
containing 50% ES media, 40% FBS, 10% DMSO at a density of 50,000 cells per cm².

### 2.1.5 Differentiation of ES cells to spinal cord neural progenitors

24 hours before starting the differentiation, CellBIND dishes (Corning) were incubated at 37°C with 0.1% gelatine (Sigma). ES cells were cultured for at least 2 passages after thawing before commencing with the differentiation. A protocol for generating spinal cord neural progenitors was previously developed (Gouti et al. 2014). Prior to differentiation feeders were removed by sequential panning steps. ES cells were harvested using Accutase (ThermoFisher Scientific) into centrifuge tubes containing ES media. Following centrifugation cells were resuspended in 0.2 mL ES medium per cm². The ES cell suspension was then transferred to a new tissue culture dish twice the size of the ancestor dish. For example, for ES cells cultured in a 10 cm² dish, cells were resuspended in 2 mL ES media and panned in a 20 cm² dish. The cell mixture containing ES cells and feeders was incubated at 37°C for 20 minutes. This time is sufficient for feeders but not ES cells to attach to the tissue culture surface. After 20 minutes cells were transferred to a new panning dish and the process repeated. Following the second incubation ES cells were harvested into a centrifuge tube, pelleted by centrifugation, resuspended in PBS, re-centrifuged and resuspended in N2B27 (Advanced DMEM/F12 and Neurobasal medium (1:1), supplemented with 1X N2, 1X B27, 1X GlutaMAX, 1X Pen/Strep, 0.1 mM 2-mercaptoethanol (all Gibco), 40 µg/mL BSA (Sigma)). 100,000 cells were aliquoted into a 3 cm² CellBIND dish (Corning) containing 1.5 mL N2B27 + 10 ng/mL FGF-2 (day (D) 0). The seeding density is a critical step in the differentiation protocol and may need to be adjusted between users, batches of tissue culture plastic and batches of reagents. On D2 the media was replaced with N2B27 + 10 ng/mL FGF-2 + 5 µM CHIR99021 (Axon). On D3 and D4 cells were fed with N2B27 + 100 nM RA (Sigma). We extended the spinal cord by culturing the cells in N2B27 from D5 onwards. After D6 a washing step was included in the feeding process to remove cell debris. Differentiations were quality controlled based on the efficiency of neural differentiation. A successful differentiation yielded >70% neural
progenitors at D5 (assayed by SOX2 flow cytometry or SOX1 immunostaining or Sox2 expression by qPCR).

2.1.6 Transforming Growth Factor (TGF)-β manipulations

TGF-β2 (R&D Systems) was added at D5 of the differentiation at 2 ng/mL in N2B27 (Dias et al. 2014). Media was replenished at D6. From D7 onwards progenitors were kept in N2B27. We used the chemical inhibitor SB431542 (SB, 10 µM in N2B27, Tocris) from D5 onwards to inhibit TGF-β signalling. For experiments where SB was removed at D7, cells were first washed with N2B27 before being kept in N2B27 for the remainder of the experiment.

2.1.7 Differentiation to GFAP positive glia

To generate GFAP-positive glia progenitors were cultured until D15. From D10, cells were maintained either in N2B27 or exposed to LIF (500 U/mL in N2B27, Merck Millipore). Cells were washed with N2B27 daily.

2.1.8 Analysing samples by quantitative PCR (qPCR)

Plates containing differentiated cells were washed twice with 0.5 mL N2B27 and twice with PBS before processing. Samples were either dissociated and centrifuged first or lysed directly. 350 µL of RLT buffer was added per sample and stored at -80°C until processing. Samples were not left more than 3 weeks. RNA was extracted using the RNeasy Mini Kit (Qiagen) according to the manufacturer’s instructions. 1-1.5 µg of RNA was added to the reverse transcription reaction (SuperScript III, Invitrogen). The resulting 20 µL of cDNA was diluted 1:12 in RNase-free water. qPCR was performed in 384-well plates in a total reaction volume of 10 µL containing 4.5 µL diluted cDNA, 0.5 µL 10 µM primers (forward and reverse) and 5 µL of Platinum SYBR Green qPCR SuperMix-UDG with ROX (Invitrogen). Plates were run using the standard program on the QuantStudio 5 Real-Time PCR system (ThermoFisher Scientific).
2.1.9 Analysing samples by immunofluorescence

Samples were washed twice with 0.5 mL N2B27 and twice with PBS before incubating in 0.7 mL 4% paraformaldehyde at 4°C for 20 minutes. The paraformaldehyde was removed, cells washed with PBS and stored in 2 mL PBS at 4°C. Samples were not left for more than 3 weeks. Prior to antibody staining cells were blocked and permeabilised in 1% BSA and 0.1% Triton (block/perm) for 1 hour at room temperature. Primary antibodies were diluted in block/perm and incubated at 4°C overnight. The following day samples were washed in PBS 3x 5 minutes. Secondary antibodies were diluted in block/perm and incubated at room temperature for 45 minutes. Samples were again washed in PBS 3x 5 minutes. The last wash was removed and samples mounted in Prolong Gold Antifade reagent (Invitrogen) with a glass coverslip. Images were collected on a Zeiss Imager.Z2 microscope fitted with an Apotome.2 structured illumination module using a 20x magnification lens (NA = 0.75). For structured illumination 5 phase images were acquired and a z-projection was made using 10 sections. Images were processed using FIJI and analysed using CellProfiler.

2.1.10 Analysing samples by flow cytometry

Cells were dissociated using Accutase, washed in PBS and fixed in 4% paraformaldehyde for 20 minutes at 4°C. Fixative was removed by centrifugation and samples stored in PBS for analysis (<3 weeks). Before staining comparable numbers of cells were aliquoted into 1.5 mL Eppendorf tubes. Cells were centrifuged at 2,000 rpm for 4 minutes and resuspended in 0.1 mL block/perm. After an incubation period of 45 minutes at room temperature antibodies were pre-diluted to a 10X mix and 10 µL added per sample. Cells were incubated with antibody for 45 minutes at room temperature. The antibody mix was removed after centrifugation and cells washed twice with 0.5 mL PBS. Cells were transferred to a flow cytometry tube (Falcon) and acquired on an LSR Fortessa (BD Biosciences). Analysis was performed using FlowJo.
Chapter 2 Materials and Methods

2.2 General Molecular Biology Techniques

2.2.1 PCR

Polymerase Chain Reaction (PCR) was performed using the Q5 Polymerase Kit (New England BioLabs (NEB)) on a T100 Thermocycler (Bio-Rad Laboratories) according to the manufacturer’s instructions. Primers were designed and checked for mismatches using Primer BLAST (NCBI (Ye et al. 2012)). Primer annealing temperatures were calculated using the Tm Calculator (NEB). PCR products were checked for purity by agarose gel (1-2% in 1x TAE (20 mM TRIS acetate, 1 mM Na₂EDTA.2H₂O, pH 8.5)) electrophoresis (95V, 0.5A, 45 minutes). PCR reactions with single products were purified on a QIAquick PCR Purification column (Qiagen). PCR reactions yielding multiple products were purified by QIAquick gel extraction (Qiagen). The manufacturer’s protocol was followed in both cases.

2.2.2 Restriction digestion

DNA was digested using specific restriction enzymes under the conditions specified by the manufacturer (NEB). 1-2 µg of DNA was digested at 37°C for 2-24 hours. Following digestion, linear DNA was dephosphorylated with Antarctic Phosphatase (NEB) to prevent re-ligation. The reaction was then purified on a QIAquick PCR Purification column (Qiagen) according to the manufacturer’s instructions.

2.2.3 Transformation and growth of chemically competent cells

Following cloning (see below) plasmids were transformed into DH5-alpha Competent E.coli (NEB). For DNA generated from cloning reactions a 1:10 mix of DNA to competent cells was used. To transform purified vectors a 1:1000 dilution was made. The mixture was incubated on ice for 20 minutes, heat-shocked at 42°C for 45 seconds, recovered on ice for 2 minutes before adding 120 µL of SOC medium (Invitrogen). The mixture was incubated at 37°C for > 1 hour under shaking to allow the bacteria to enter the growth phase. 100 µL was aliquoted and spread onto Luria-Bertani (LB) agar plates containing the appropriate antibiotic. Often serial dilutions were made for more efficient cloning reactions. The plates
were incubated overnight at 37°C to allow colonies to form. Colonies were picked and used to inoculate 5 mL of LB media containing the appropriate antibiotic and grown overnight at 37°C with shaking. A 50 µL aliquot of each culture was taken and stored at 4°C to inoculate larger preps. Cultures were centrifuged at 4,500 rpm for 20 mins and the DNA isolated from the bacterial pellet using the Plasmid DNA Miniprep Kit (Qiagen). DNA samples were screened by sequencing (Mix2Seq kit, Eurofins Genomics). Aliquots from the initial cultures of correct clones were used to inoculate 100 mL of LB media containing antibiotic. This larger preparation was grown overnight at 37°C with shaking. The culture was divided equally between two centrifuge tubes and pelleted at 4,500 rpm for 30 minutes. DNA was extracted using the Plasmid DNA Maxiprep Kit (Qiagen) and quantified by measuring the absorbance at 260 nM on a NanoDrop ND-1000 spectrometer (LabTech). The DNA sequence was checked by sequencing.

2.2.4 Lentivirus production

cDNA was cloned into the pENTR 11 Dual Selection vector (Invitrogen). pENTR was digested using EcoRI, EcoRV. Primers were designed to the start and end of the target cDNA with a target annealing temperature of 69°C. The oligos were synthesised with a 30 base pair sequence homologous to the cut pENTR. The amplified cDNA was ligated into pENTR using the InFusion cloning kit (Clontech). 1/5 was transformed into competent cells. The insertion was checked by sequencing. The cDNA in pENTR was transferred into pLEX_307, modified to contain IRES-NLS-mKate2, using BP Clonase II (Invitrogen) by following the manufacturers protocol. 1/5 of the pLEX vector was transformed as described above except all incubation steps were at 30°C.

All virus work was performed in accordance with the recommended health and safety guidelines for category level 2 material. Virus was produced in HEK293T cells in a modified version of a protocol developed previously (Tiscornia et al. 2006). Briefly, pLEX-cDNA-mKate2 (0.9 µg) along with the packaging vectors pMDL (0.6 µg), pVSVG (0.3 µg), pREV (0.2 µg), and 6 µL Xtremegene HP reagent were diluted in 200 µL OptiMEM (Gibco). The mixture was incubated at room
temperature for 20 minutes and then added dropwise to 1 10 cm plates. Virus-containing supernatant was harvested at 48 and 72 hours after transfection and the virus pelleted by ultracentrifugation (18,000 xg, 4°C, 90 minutes, SW32Ti rotor). The supernatant was removed and the pellet resuspended in 0.2 mL Hank’s Balanced salt solution. The resuspended virus was transferred to a 1.5 mL Eppendorf tube and shaken at room temperature for 15 minutes. Coagulated protein was removed by a brief centrifuge at 13,000 rpm on a bench-top centrifuge. The supernatant was aliquoted into new 1.5 mL Eppendorf tubes and stored at -80°C until required.

Before infection, D4 samples were washed twice in N2B27 and 10 µL of purified virus to cells in 1 mL of N2B27. At D5, D6, D7 cells were washed with 0.5 mL N2B27 before feeding. On D7 cells were dissociated with Accutase (as described previously, centrifuged, washed with PBS, centrifuged and resuspended in PBS + 0.1% BSA supplemented with the live cell dye, Calcein Violet (Life Technologies). Infected cells were sorted by Fluorescence Activated Cell Sorting (FACS) on either an FACS Aria III or an Influx (BD Biosciences) by the Crick Flow Cytometry STP. An equal number of negative cells were isolated as an internal control. Both fractions were pelleted and RNA extracted for analysis by qPCR.

2.2.5 CRISPR/Cas9

Guide RNAs (gRNA) were designed using the CRISPR Design Tool (crispr.mit.edu, Zhang lab). This software uses the total number of predicted off-targets, as well as the number and location of mismatches to compute scores for each gRNA. Guides with high scores were picked to minimise the chance of off-target affects. Guides were cloned into the px459 plasmid using a previously published protocol modified in-house by Sarah French (Ran et al. 2013). This makes use of two consecutive BbsI restriction sites. Oligos were ordered containing the BbsI sites and 5’ phosphorylated to prevent self-ligation. A G was added to the forward oligo if no G was present for the U6 promotor. Forward and reverse oligos were pre- annealed at 95°C for 10 minutes in T4 ligation buffer. A digestion/ligation reaction was performed by mixing px459, annealed oligo duplex cooled to room temperature,
BbsI, ATP, DTT, T4 Ligase, Tango buffer and incubating it on a thermocycler for 6 cycles of 37°C for 5 minutes, 21°C for 5 minutes. 1:10 of the reaction mixture was transformed into competent cells and DNA extracted as described above. Several densities of competent cells were plated in order to obtain isolated colonies.

ES cells were adapted to 2i+LIF conditions before electroporation (Ying et al. 2008). The feeders were panned from ES cell cultures as described above and plated on CellBIND dishes coated for 24 hours with 0.1% gelatine in N2B27 + CHIR99021 + PD184352 (Axon) + LIF (2i+LIF). After 3-4 days cells were split using Accutase. ES cells went through 3 passages before electroporation. 2 x10⁶ cells were electroporated with 2 µg of plasmid (this was doubled for the generation of the NFIA/B(-/-) cells) and mouse ES cell Nucleofector reagent (Lonza) using a Nucleofactor II electroporator (Amaxa), program A-023. The electroporated cells were divided in two 10cm CellBIND dishes coated overnight with 0.1% gelatine (25% in one and 75% in the second). 24 hours after seeding the cells were exposed for 48 hours to puromycin (1.5 ng/mL) in 2i+LIF before continued culture in 2i+LIF. Colonies were picked using a p10 pipette and dissociated in a round-bottomed 96-well plate in 20 µL 0.25% Trypsin/EDTA. Dissociated colonies were split equally between one well of a 96-well plate coated with feeders and a 48-well plate coated in 0.1% gelatine in ES+LIF. Cells on feeders were grown to confluency, dissociated with Accutase and resuspended in ES+LIF + 40% FBS + 10% DMSO and stored at -80°C. Cells plated in the 48-well were grown to confluency and DNA extracted using the PureLink Genomic DNA kit (Invitrogen). The target region was amplified by PCR using primers designed proximal to the gRNA target site. The PCR products were separated by agarose gel electrophoresis, extracted and sequenced as described above. When possible, homozygous clones were prioritised for further characterisation.

2.3 Single cell sequencing and analysis

2.3.1 Sample collection and sequencing of single cells

Single cell sequencing was performed as described previously (Sagner et al. 2018). Briefly, samples were washed twice with N2B27 and twice with PBS before
being dissociated with Accutase, Cells were collected into a tube containing N2B27 and centrifuged. The cell pellet was resuspended in PBS and half transferred to a new centrifuge tube for RNA extraction. The remaining cell suspension was centrifuged and the pellet resuspended in PBS. The cells were centrifuged again and the pellet resuspended in N2B27. The sample was filtered into a FACS tube (Falcon) and transferred to ice. The cell suspension was diluted to 250,000-400,000 cells/mL and individual cells captured on a 96 small or medium IFC chip using the Fluidigm C1 system. The capture efficiency was assessed manually by imaging each capture site on a NanoEntek JuLi cell imager. Capture sites containing single cells were processed for library preparation and sequencing. Full-length cDNA was prepared using the SMARTer Ultra Low RNA kit (Clontech) on the C1 chip according to the manufacturer’s protocol. ERCC RNA spikes were added and libraries prepared using the Illumina Nextera XT DNA Sample Preparation kit following a protocol supplied by Fluidigm. Sequencing of the DNA libraries was performed on the Hiseq 2500 or 4000 (Illumina) using 50- or 75-bp paired end runs.

2.3.2 Analysis

The initial analysis was performed by Julien Delile using a modified version of a previously published pipeline (Gouti et al. 2017, Sagner et al. 2018). Sequencing reads were aligned to the GRCm38 Ensembl mouse genome using Tophat2 (Kim et al. 2013) and counted using HTSeq-count (Anders et al. 2015). Cells were excluded based on their transcript abundance (< 5 x10^5 reads) and the number of expressed genes (< 1000). The reads were then normalized to read counts per million and log transformed.

Julien Delile then identified modules of genes with similar expression patterns. Genes with a Spearman correlation (r > 0.3) with at least three other genes were selected, grouped and further filtered using the following 3-step iterative process:

1. The remaining genes were assigned gene modules by performing a hierarchical clustering using the Spearman dissimilarity matrix of the UMI counts and Ward’s agglomeration criterion. The number of modules was selected heuristically by
determining the “elbow” position in the curve representing the total within-module gene level variation per number of modules.

2. A first gene module filtering criterion was applied to test whether enough cells were expressing the genes within the module. For each cell, an average expression level per module was obtained by averaging the z-scored log-transformed expression levels of all genes belonging to the module. These gene module averaged levels were then binarized independently by using a parameter-free adaptive thresholding method (R function binarize.array() from the ArrayBin package). A cell was considered expressing a gene module if the associated Boolean value was true. Modules with fewer than five cells expressing it were excluded.

3. A second gene module filtering criterion was applied to test whether the cells were expressing the gene module with consistently high levels in the majority of the genes within the module. We binarized the z-scored log-transformed expression levels of all the remaining genes independently. Then, for each module, we calculated the ratio of Boolean values in cells expressing the module (as defined in 2.). We excluded modules where less than 40% of these Boolean values were true.

The iterative loop was terminated when no gene module was removed in the last iteration (this section was written by Julien Delile).

In collaboration with Julien Delile, we performed hierarchical clustering on the cells using the identified gene modules. Cells not represented by any gene modules were excluded. Gene modules containing neuronal genes such as Tubb3, Dcx, Lhx1, Lhx5, Lbx1 and mesodermal markers such as Meox1, Foxc2, Twist1, were used to re-cluster the data in order to identify neurons and mesoderm. We used cutree (k = 3) to cut the dendrogram and isolate the neural progenitors. To order cells in pseudotime, 4 gene modules correlated with sample collection day (Spearman correlation (r > 0.6) were identified.

The following analysis is performed in R using the script provided in Appendix A (sections written by Julien Delile are annotated, the remainder has been written by the author). For each cell, the expression of all genes within each time-correlated module was averaged and Principle Component (PC) Analysis used to plot the
cells in 2 dimensions. PC1 represented pseudotime. Outliers along the first principle component were removed by plotting a histogram of cell counts in 30 equal bins along PC1. The expression of all genes along PC1 was smoothed using LOESS (span = 0.7, bins = 20). Genes that had low expression (mean < 1.4) and low dispersion (< 1.5 standard deviations away from the mean) were removed. Gene ontology was performed using the GO enrichment analysis tool (geneontology.org).

To compare the expression dynamic in silico to that in vitro the pseudotime axis was divided into 5 equal bins and the expression within each bin averaged. qPCR was performed on in vitro samples collected daily from D5 to D9. The expression from 3 differentiations was averaged and a Pearson correlation performed between the averaged in silico data and the in vitro data (Appendix B).

2.4 Embryo Manipulation

2.4.1 Embryo harvesting and embedding

Embryos were collected after the prescribed time period following the plug date (E0.5). All extraembryonic tissue was removed before embryos were fixed in 4% paraformaldehyde at 4°C. The duration of fixation was dependent on stage (E9.5-E10.5 2 hours, E11.5-E13.5 overnight). Embryos for in situ hybridisation were carried through a methanol series (25%-50%-75%-100%, each for 10 minutes at room temperature) and stored at -20°C. Before continued processing, embryos were rehydrated by reversing the methanol series. Embryos were equilibrated in sucrose overnight at 4°C and the brachial segment of the neural tube (anterior forelimb to anterior hindlimb) dissected away. All remaining embryonic tissue was discarded. Brachial neural tubes were embedded in gelatine and sectioned from forelimb to hindlimb on a cryostat. Sections were mounted on superfrost slides and stored at -80°C.
2.4.2 Immunofluorescence

The gelatin was melted by washing slides 3x in PBS at 42°C for 5 minutes. Antibody staining was performed using a comparable approach described for the staining of *in vitro* cells (see above). Briefly, sections were incubated in block/perm (PBS + 1% BSA + 0.1% Triton) for approximately 1 hour at room temperature before incubating overnight at 4°C in primary antibody diluted in block/perm. Primary antibody was washed away by incubating the slides for 5 minutes in PBS 3x. Secondary antibody was diluted in block/perm and incubated at room temperature for 1 hour. The washing step was repeated and coverslips mounted using ProLong Gold Antifade reagent (Invitrogen). Images were collected on a Zeiss Imager.Z2 microscope fitted with an Apotome.2 structured illumination module using a 20x magnification lens (NA = 0.75). For structured illumination 5 phase images were acquired and a z-projection was made using 8 sections. Images were processed using FIJI.

2.4.3 *In situ* hybridisation probe design and manufacture

The Zfp536 probe sequence used for *in situ* hybridisation was based on a previously validated probe (Qin et al. 2009). An approximately comparable region was cut from Zfp536 cDNA and cloned into pBluescript-KS(+/-). Zfp536 cDNA was digested using AfeI, NheI restriction enzymes and the cut fragment gel purified. pBluescript was digested using EcoRV, Xbal and dephosphorylated with Antarctic Phosphatase before being purified through a column (QIAquick PCR Purification Kit, Qiagen). The two fragments were ligated using the Rapid DNA Ligation kit (Roche) for 20 minutes at room temperature. The product was transformed into chemically competent cells, bacterial cultures expanded and DNA purified as described above. A positive plasmid (pBS-Zfp536) was identified by sequencing. The probe was transcribed by T7 polymerase using the Riboprobe Combination System (Promega) on HindIII digested pBS-Zfp536. To check the reaction produced a single product an aliquot was run on a 1% agarose gel.
2.4.4 \textit{In situ} hybridisation on sectioned embryos

All steps were carried out using RNase-free reagents. To accurately compare different time points each slide contained embryo sections from embryos collected daily from E9.5 to E13.5. Gelatine was removed by washing slides 3x in PBS at 42\textdegree{}C. Slides were incubated in prehybridisation buffer (50\% formamide, 5x SSC, 0.1\% Tween, 0.1 mg/mL Heparin Na Salt, H\textsubscript{2}O\textsubscript{2}) for 30 minutes at 69\textdegree{}C. Zfp536 antisense probe was diluted in hybridisation buffer (prehybridisation buffer + 0.1 mg/mL Torula yeast RNA, 0.1 mg/mL herring sperm DNA) (1:100) and incubated at 70\textdegree{}C for at least 10 minutes. 200 \mu L of diluted probe was added per slide, covered with parafilm and incubated overnight at 69\textdegree{}C in a sealed humidified chamber. Wash buffers 1 (5x SSC, 0.1\% Tween, H\textsubscript{2}O\textsubscript{2}) and 2 (0.2x SSC, 0.1\% Tween, H\textsubscript{2}O\textsubscript{2}) were made fresh and prewarmed to 69\textdegree{}C. Slides were washed in buffer 1 for 15 minutes at 69\textdegree{}C. Two 30 minutes washes were performed with buffer 2 at 69\textdegree{}C. The sections were then washed at room temperature with PBST (PBS + 0.1\% Tween) for 15 minutes. Slides were then incubated in blocking buffer (PBST + 2\% sheep serum) for 1 hour at room temperature in a humidified chamber. Anti-dioxigenin-AP antibody was diluted (1:3000) in blocking buffer and added to the slides. Antibody incubation was overnight at 4\textdegree{}C in a humidified chamber. Slides were washed in PBST (3x 10 minutes) and washed in NMNT (100 mM Tris pH 9, 5 mM MgCl\textsubscript{2}, 100mM NaCl, 0.1\% Tween) (2x 10 minutes). The sections were developed using BM Purple AP Substrate (Roche) at room temperature. After the staining appeared slides were washed twice in NTMT and once with PBST before fixation with 4\% paraformaldehyde (10 minutes, room temperature). Fixative was washed off with PBST and coverslips mounted with Prolong Gold. Images were taken on a Zeiss Imager.Z2 microscope using a 10x objective (Zeiss).

2.4.5 Chick \textit{in ovo} electroporation

Chick embryos were grown at 37\textdegree{}C in a humidified oven for 48 hours. Using a syringe and needle the egg shell was punctured and 5 mL of egg white removed. The needle hole was widened to allow access to the embryo. Ink was injected underneath the embryo to enable visualisation. pLEX-Zfp536-mKate2 was premixed with dye and PBS to a final concentration of 1 \mu g/\mu L. Using a mouth
pipette the mixture was injected into the lumen of the neural tube of HH stage 8-12 chick embryos (Hamburger and Hamilton 1992). DNA was electroporated into one half of the neural tube by placing electrodes unilaterally and passing 3 50 millisecond pulses of 30V using a BTX electroporator. 5 mL of egg white was replaced, the hole sealed with Sellotape and the eggs returned to the 37°C humidified oven. 24 hours later the embryos were screened for mKate2 fluorescence using a Leica MZFCIII fluorescence microscope. Positive embryos were removed from eggs and the brachial region dissected. Brachial neural tubes were embedded, sectioned and stained as described above for mouse embryos.

### Table 1 List of antibodies

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Dilution</th>
<th>Host</th>
<th>Clonality</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFIA</td>
<td>1:500</td>
<td>Rabbit</td>
<td>Polyclonal</td>
<td>Active Motif</td>
</tr>
<tr>
<td>NFIB</td>
<td>1:200</td>
<td>Rabbit</td>
<td>Monoclonal</td>
<td>Abcam</td>
</tr>
<tr>
<td>NFIB-AF488</td>
<td>1:200</td>
<td>Rabbit</td>
<td>Monoclonal</td>
<td>Abcam</td>
</tr>
<tr>
<td>SOX9</td>
<td>1:300</td>
<td>Goat</td>
<td>Polyclonal</td>
<td>R&amp;D Systems</td>
</tr>
<tr>
<td>SOX9</td>
<td>1:1000</td>
<td>Rabbit</td>
<td>Polyclonal</td>
<td>Millipore</td>
</tr>
<tr>
<td>SOX2</td>
<td>1:500</td>
<td>Goat</td>
<td>Polyclonal</td>
<td>R&amp;D Systems</td>
</tr>
<tr>
<td>SOX2-v450</td>
<td>1:50</td>
<td>Mouse</td>
<td>Monoclonal</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>GFAP</td>
<td>1:500</td>
<td>Rabbit</td>
<td>Polyclonal</td>
<td>DAKO</td>
</tr>
<tr>
<td>GFAP-AF488</td>
<td>1:50</td>
<td>Mouse</td>
<td>Monoclonal</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>TUBB3</td>
<td>1:2000</td>
<td>Mouse</td>
<td>Monoclonal</td>
<td>Covance</td>
</tr>
<tr>
<td>TUBB3-AF647</td>
<td>1:20</td>
<td>Mouse</td>
<td>Monoclonal</td>
<td>BD Biosciences</td>
</tr>
<tr>
<td>FLAG</td>
<td>1:2000</td>
<td>Rat</td>
<td>Monoclonal</td>
<td>Sigma</td>
</tr>
<tr>
<td>RFP</td>
<td>1:20000</td>
<td>Guinea Pig</td>
<td>Polyclonal</td>
<td>Kerafast</td>
</tr>
</tbody>
</table>

### Table 2 Primers for qPCR

<table>
<thead>
<tr>
<th>Gene</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nfia</td>
<td>AAGCCTCCAACCACATCAAC</td>
<td>TTTACAAGCTTGGAATCCCG</td>
</tr>
<tr>
<td>Sox9</td>
<td>AGGAAGCTGCCAGACCAGT</td>
<td>CGAAGGTCTCTTCTCTGCT</td>
</tr>
<tr>
<td>Gfap</td>
<td>(Abranches et al. 2006)</td>
<td>CTCTATGCCTCCTCCGAGACGAT</td>
</tr>
<tr>
<td>Nfib</td>
<td>GGTGAGTTGGGAGAATGTGTC</td>
<td>TTCTGCTTGTATTTCGGGCTTC</td>
</tr>
<tr>
<td>Acdd1</td>
<td>GCAGCTGCCCTCTCTCTCTCT</td>
<td>AGTGTGAAGGATTTGGAG</td>
</tr>
<tr>
<td>Nfix</td>
<td>AGGCCATGACAGTGGAGTC</td>
<td>CATCTCTTGCTGGTGTGAGA</td>
</tr>
<tr>
<td>Tubb3</td>
<td>CAACAGACGCCATCCAGG</td>
<td>CTTGGGGCCCTGGGCTCCGA</td>
</tr>
<tr>
<td>Sox2</td>
<td>ACCAGCTCCAGAACCTACAT</td>
<td>CCTGCTGCGAGTACGACT</td>
</tr>
<tr>
<td>Smad7</td>
<td>TTCGGACAAACAGAGTCAGC</td>
<td>GGTAACTGCTGCGGTTGTA</td>
</tr>
<tr>
<td>Zfp536 (Ning et al. 2015)</td>
<td>GCTCCTGCCTCAACAGAAAC</td>
<td>TTCGCTGCTTACTAGCTTGCT</td>
</tr>
<tr>
<td>Npas3  (Sha et al. 2011)</td>
<td>TCGGCATTTTCGTTTAGACC</td>
<td>AAGAAAGAGGCGGTGAATT</td>
</tr>
</tbody>
</table>
Lin28a (Peng et al. 2011) | CGGGCATCTGTAAAGTTGTC | CAGACCCTTGCTGACTTCT
Lin28b (Cotterman and Knoepfler 2009) | TGAAGAAGAACCCAAAGGAAGAC | TGATGATCAAGGCCACCACAGT
Hes1 | GGCAGAAGGCAAGAATAATTG | GTGCTTCAGACGTTTCAG
Cux2 | GGACTCAGCTATCCAACAAC | CTTGCACCACAGTTTCAAA
Hopx | TCCCAGCTGACTCTCAGAAG | AGTCGCTGACCTTACGTCT
Actin | TGGCCTCTAGCACCATGA | CCACCAGATCCACACAGG
Sp8 | CCCAACTTTCTCATTTCTCTTCTC | CACCGATGTGTGCATTACT

Table 3 List of plasmids

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>pLEX_307</td>
<td>Gift from David Root (Addgene # 41392)</td>
</tr>
<tr>
<td>Tet-O-FUW-Zfp536</td>
<td>Gift from Marius Wernig (Addgene # 45845)</td>
</tr>
<tr>
<td>pENTR-NICD1</td>
<td>Gift from Xin Chen (Addgene # 46048)</td>
</tr>
<tr>
<td>pENTR 11</td>
<td>Invitrogen</td>
</tr>
</tbody>
</table>
Chapter 3. **An in vitro model of the neural-to-glial transition**

As the spinal cord forms, neural progenitors become regionalised and between E9 and E12.5 in mouse progenitors differentiate into neurons (Rowitch 2004, Guérout et al. 2014). At E11.5 progenitors begin to express NFIA and NFIB and switch to generating glial subtypes (Deneen et al. 2006). The mechanism governing the timing of this transition remains unclear. *In vitro* models are valuable systems to study developmental questions as they are highly amenable to biochemical manipulations. A previously described method generates neural progenitors (iNPs) with a spinal cord identity (Gouti et al. 2014). This 5 day protocol posteriorizes ES cells using a short pulse of WNT signalling before inducing a neural fate with retinoic acid. iNPs have been shown to express regional markers and differentiate into neurons similar to their *in vivo* counterparts (Gouti et al. 2014, Sagner et al. 2018). Whether they generate glia had yet to be established. To develop an *in vitro* model of the neural-to-glial transition we chose to build upon the iNP method. iNPs were cultured in a basal media (N2B27) up to day (D) 15 and the schedule of temporally restricted markers assayed and compared to their known dynamics *in vivo*.

In the mouse, neuron differentiation begins at E9 (Kicheva et al. 2014). It has been shown previously that iNPs can generate neurons (Gouti et al. 2014, Sagner et al. 2018). We sought to resolve the timing of neuron generation *in vitro* and compare it to the *in vivo* schedule. To do this, D5 iNPs were cultured in N2B27 up to D11. Daily time points were collected and tested for the expression of the neuron marker TUBB3 by flow cytometry. Neurons appeared in culture at D6, and their proportion increased steadily through the differentiation. By D11 the percentage of neurons had reached 45% (Figure 7). Thus neuron differentiation *in vitro* begins at D6 suggesting at this time point iNPs are equivalent to *in vivo* progenitors at E9. This is consistent with neuron differentiation dynamics reported previously (Sagner et al. 2018).
The transcription factors NFIA & NFIB mark the onset of the glial lineage (Deneen et al. 2006). In the mouse spinal cord, this happens at E11.5, ~3 days after the onset of neuron differentiation. To test whether the temporal relationship between neuron differentiation and NFIA and NFIB induction is preserved in iNPs, we assayed their expression dynamics by qPCR, immunofluorescence and flow cytometry. At D6, at the onset of neuron differentiation, little to no Nfia/b expression was detectable by qPCR. Induction of Nfia/b was observed at D7/8 and continued to increase to D11 (Figure 8C). NFIA and NFIB are also expressed in paraxial mesoderm lineages (Chaudhry et al. 1997). SOX2 is a transcription factor that is expressed in neural progenitors throughout the period of neuron differentiation and gliogenesis. To exclude the possibility that the dynamic of Nfia, Nfib expression was solely from non-neural cell types, the intensity of NFIA and NFIB protein in SOX2-positive cells was measured by immunofluorescence and flow cytometry respectively. Expression of NFIA/B began at D8/9 and increased at D11 (Figure 8A,B, Figure 9). Thus, similar to in vivo, iNPs induce NFIA/B approximately 3 days after the onset of neuron differentiation, suggesting that the mechanism underlying this transition is conserved.

It has been proposed that NFIA induction in vivo depends on the activity of SOX9 (Kang et al. 2012). Deletion of SOX9 leads to a severe reduction in the number of glial cells in vivo (Stolt et al. 2003). However, the temporal expression profile of SOX9 resembles the timing of neuron differentiation, with onset at E9.5 (Stolt et al. 2003). To understand whether similar dynamics of Sox9 expression are captured in vitro, samples were collected daily from D5 to D11 and Sox9 expression measured by qPCR. Low levels of Sox9 expression were detected at D5. Expression continued to increase reaching its maximal expression at D8 when NFIA/B are induced. Sox9 expression then decreased from D8 to D11 (Figure 8C). Since SOX9 is also expressed in paraxial mesoderm lineages we confirmed this result by immunofluorescence at D5, D7, D9, D11. Very little SOX9 expression was observed in neural progenitors at D5. SOX9-positive cells at this stage were not SOX2-positive. SOX9 expression emerged at D7 and continued to be expressed at D7 and D9 (Figure 8A). Similar to in vivo, SOX9 is expressed prior to NFIA and NFIB induction in iNPs.
Glial cells depend on co-activation of target genes by NFIA and SOX9 (Kang et al. 2012). Overexpression of these target genes is able to rescue gliogenesis in the absence of SOX9 or NFIA. An example is *Apcdd1*, an antagonist of WNT signalling which is induced at E12.5, closely following NFIA expression (Kang et al. 2012). NFIB has been shown to activate the expression of another NFI family transcription factor, NFIX (Matuzelski et al. 2017). *In vivo*, NFIX expression follows NFIA/B and knockout mice show a delay in gliogenesis. To test whether the activation of these genes is recapitulated in iNPs, we analysed the induction of *Apcdd1* and *Nfix* by qPCR. Expression of *Apcdd1* was detectable at D3, before the cells had adopted a neural identity, but then decreased to barely detectable levels at D4. *Apcdd1* was re-expressed at D9 and increased until D11 (Figure 10). Induction of *Nfix* began at D8/9 and continued to increase to D11 (Figure 10). Thus, *Apcdd1* and *Nfix* expression follows NFIA and NFIB induction. These dynamics are consistent with the pattern of *Apcdd1*, *Nfix* expression *in vivo*. Therefore the *in vitro* model not only recapitulates the onset of the glial lineage but also the elaboration of the downstream transcriptional programme.

*In vivo*, glial progenitors migrate away from the progenitor zone and terminally differentiate. Terminal differentiation is characterised by the onset of GFAP expression, occurring *in vivo* at E17.5, 6 days after the induction of NFIA. As NFIA induction *in vitro* is at D8 we postulated that by D15 terminal differentiation should be observed. Progenitors were maintained in N2B27 and at D15, GFAP positive cells were quantified by immunofluorescence and flow cytometry. At this time point 2% of cells were positive for GFAP (Figure 11B). Immunofluorescence showed that GFAP-positive cells had a fibrous morphology reminiscent of differentiated astrocytes (Figure 11A). Differentiated glia localised to the edges of the colony suggesting that migration is a feature of terminal differentiation that could be directed *in vitro*. Taken together these data demonstrate that the *in vitro* model recapitulates all the basic features of the glial lineage, from onset to terminal differentiation.

Terminal differentiation of glia has been shown to require NFIA in combination with activation of the JAK/STAT pathway, and the small percentage of glia at D15 might therefore be the result of insufficient JAK/STAT signalling (Barnabé-Heider et al.)
2005, Deneen et al. 2006, Hong and Song 2014). To test this possibility, we exposed cells at D10 to LIF, an activator of the JAK/STAT signalling pathway (Nicola and Babon 2015), and measured the number of GFAP positive cells at D15 by flow cytometry and immunofluorescence. Exposure to LIF doubled the percentage of glia (Figure 11B). These data corroborate findings that JAK/STAT activation plays an important role in gliogenesis. It suggests that in vitro, signalling may be limiting the differentiation of glial progenitors and that further refinement of the protocol could increase the number of fully differentiated glia.

Together, these results define an in vitro differentiation protocol that allows iNPs to reproduce the basic features of the neural-to-glial transition. After the onset of neuron differentiation, progenitors sequentially switch on SOX9, NFIA/B, Apcdd1/Nfix, and finally GFAP. The schedule of expression is reminiscent of their in vivo counterparts. Intriguingly, the transition from neuron differentiation through to the activation of Apcdd1/Nfix occurs without the addition of exogenous signalling suggesting that the mechanism is either cell autonomous or that endogenous signalling is sufficient. Whether endogenous signals are produced from paraxial mesoderm, neurons, or iNPs themselves remains to be investigated. A feature of the model is that NFIA induction is coincident with maximal Sox9 expression. This relationship may be preserved in vivo. In the developing spinal cord, SOX9 expression begins at E9.5 and increases to E11.5 when NFIA and NFIB are induced (Stolt et al. 2003). As discussed previously the mechanism controlling the timing of the transition is poorly understood. It is tempting to speculate that the time taken for SOX9 to cross a certain threshold governs when NFIA is expressed. This raises questions over the molecular mechanism controlling SOX9 production. The in vitro model represents a tractable system to study the timing of the neural-to-glial transition.
Figure 6 Differentiation of mouse ES cells to iNPs and glia

A schematic describing the differentiation protocol. Mouse ES cells are seeded at day 0 (D0) in N2B27 + FGF. Between D2 and D3 cells CHIR is added to activate the WNT signalling pathway. At D3 the media is replaced with N2B27 + retinoic acid (RA). By day 5 cells, iNPs with a spinal cord identity have been generated (Gouti et al. 2014). From D5 onwards, cells are left to develop in N2B27.

Figure 7 Neuron differentiation begins at D6 in vitro

The percentage of SOX2 negative, TUBB3 positive cells (neurons) from D5 to D11 in iNP cultures measured by flow cytometry (n = 4-19). Error bars represent standard error of the mean of the mean. **** p < 0.0001, Mann-Whitney unpaired t test.
Figure 8 The dynamics of SOX9, NFIA and NFIB are recapitulated in vitro

SOX9, NFIA expression in iNPs cultured in N2B27 to D11 (n = 3-24). (A) Immunofluorescence images for DAPI, NFIA, SOX9, SOX2 at D5, D7, D9, D11. Images at 20x. Scale bars = 40μM. (B) Normalised mean NFIA and NFIB protein intensity in SOX2 positive nuclei measured by immunofluorescence and flow cytometry, respectively. NFIB measurements are at D7, D9, D11. Error bars represent standard error of the mean. (C) Sox9, Nfia, Nfib expression by qPCR. Data is normalised to the minimum and maximum values for each gene. Error bars represent standard error of the mean.
Figure 9 The neural-to-glial transition is recapitulated *in vitro*

Neuron differentiation and NFIB expression dynamics in iNPs cultured in N2B27 to D11 (n = 4-24). (A) Percentage of TUBB3+ neurons and NFIB+ progenitors across the differentiation measured by flow cytometry. For NFIB measurement cells were first gated on the progenitor marker SOX2. The percentage is normalised to the mean of D5 and mean of D11. (B) Expression of *Nfia* and *Nfib* from D3 to D11 measured by qPCR. (A-B) Error bars represent standard error of the mean. (C) Immunofluorescence for DAPI, SOX2 (Green), NFIB (Red), TUBB3 (Blue) at days 7, 9, 11. Images at 20x. Scale bars = 40μM.
Figure 10 The *in vitro* model recapitulates *in vivo* gene expression dynamics

Expression of Sox9, Nfia, Nfib, Apcd1, Nfix from D3 to D11 in iNPs measured by qPCR (n = 4-24). The data is normalised to the minimum and maximum value of each gene. Error bars represent standard error of the mean.
Figure 11 Terminal gliogenesis is recapitulated in vitro

Analysis of GFAP expression in iNPs cultured to D15. (A) Immunofluorescence images for DAPI, GFAP (Green), TUBB3 (Red), SOX2 (Blue) at D15. Images at 20x. Scale bars = 40μm. (B) Flow cytometry for GFAP at D15 in cells cultured from D10 to D15 +/- LIF. Error bars show standard error of the mean. ** p = 0.0079, Mann-Whitney unpaired t test.
Chapter 4. The \textit{in vitro} system recapitulates the \textit{in vivo} phenotypes of glial genes

\textit{In vivo} spinal cord neural progenitors sequentially generate neurons and glia. This transition in progenitor potency is characterised by specific changes in the expression of key genes. In the previous chapter we defined an \textit{in vitro} model that recapitulated these patterns of gene expression. Briefly, iNPs cultured in N2B27 first produced neurons before sequentially inducing SOX9, NFIA, NFIB and \textit{Apcdd1}, \textit{Nfix} and then generate differentiated glia (Chapter 3). SOX9, NFIA and NFIB, are implicated in controlling gliogenesis and are necessary for normal glial generation \textit{in vivo} (Stolt \textit{et al.} 2003, Deneen \textit{et al.} 2006). To functionally validate the \textit{in vitro} model, we knocked out these genes in embryonic stem (ES) cells and assayed their ability to transition and differentiate to glia.

4.1 Loss of gliogenesis upon combinatorial deletion of NFIA and NFIB

NFIA and NFIB appear concurrently and mark the onset of the glial lineage (Deneen \textit{et al.} 2006). \textit{In vivo}, loss of either causes a substantial reduction in differentiated glia. Furthermore, deletion of NFIA causes a reduction in the activation of downstream target genes such as \textit{Apcdd1} (Kang \textit{et al.} 2012). We wanted to test the faithfulness of the \textit{in vitro} system by removing NFIA in ES cells and measuring downstream target gene activation and terminal differentiation to test whether iNPs reproduce the phenotypes observed \textit{in vivo}. The residual glial differentiation in the absence of either NFIA or NFIB has been proposed to be due to redundancy between the two genes. Testing this redundancy \textit{in vivo} is challenging, since both NFIA and NFIB null mice have phenotypes in multiple tissues and die at birth (das Neves \textit{et al.} 1999, Gründer \textit{et al.} 2002). The \textit{in vitro} model represents a tractable system to test this redundancy.

The CRISPR/Cas9 technology can be used to introduce single point mutations in coding sequences. In combination with specific guide RNAs (gRNA), this allows editing of any gene of interest (Ran \textit{et al.} 2013). It is common to target the 5’
coding region to maximise the effect of frame shifts and/or chance of introducing a premature stop codon. \textit{Nfia} and \textit{Nfib} are composed of 9-12 exons and have several predicted isoforms. Exon 1 varies between isoforms and only encodes a few amino acids. gRNAs were designed to the start of exon 2 as it is larger and conserved between isoforms. Through this approach we generated stable ES cell lines with frame shift mutations in \textit{Nfia} and \textit{Nfib} individually and in combination (NFIA\textsuperscript{(-)}, NFIB\textsuperscript{(-)}, NFIA/B\textsuperscript{(-/-)}). NFIA\textsuperscript{(-)} had a 10 and 11 base pair deletion. NFIB\textsuperscript{(-)} had a 5 base pair deletion on both alleles. NFIA/B\textsuperscript{(-/-)} had a 1 and 4 base pair deletion in \textit{Nfia}, and a 11 and 25 base pair deletion in \textit{Nfib}. Using CRISPR/Cas9 technology, we produce frame shifts in multiple genes individually and simultaneously. In iNPs, induction of NFIA and NFIB is at D8/9. To test the consequence of the mutations on NFIA and NFIB protein expression we differentiated NFIA\textsuperscript{(-)}, NFIB\textsuperscript{(-)}, NFIA/B\textsuperscript{(-/-)} ES cells to D10 and assayed NFIA and NFIB in progenitors by immunofluorescence and western blotting. At D10, NFIA expression was observed in SOX2-positive progenitors in wild-type (WT) and NFIB\textsuperscript{(-)} cells. No NFIA expression was observed in either the NFIA\textsuperscript{(-)} or NFIA/B\textsuperscript{(-/-)} cells (Figure 12A, B). At D10, NFIB expression was observed in SOX2-positive progenitors in WT and NFIA\textsuperscript{(-)}. No NFIB expression was observed in either the NFIB\textsuperscript{(-)} or NFIA/B\textsuperscript{(-/-)} cells (Figure 13). This result was confirmed by western blotting (Figure 12C). Frameshifts produced using CRISPR/Cas9 are therefore sufficient to generate null NFIA and NFIB and NFIA/B cells.

\textit{In vivo}, the absence of NFIA and NFIB affects gliogenesis without any reported effect on neuron differentiation (Deneen et al. 2006). \textit{In vivo} neuron differentiation commences at E9 and \textit{in vitro} at D6. In both systems NFIA and NFIB expression begins \textasciitilde2 days later. We first tested whether removal of NFIA or NFIB altered the onset of neuron differentiation by assaying the neuron marker TUBB3 by flow cytometry at D7. There were approximately 12\% TUBB3 positive cells in the WT population. In NFIA\textsuperscript{(-)} and NFIA/B\textsuperscript{(-/-)} cells, the percentage of TUBB3 positive neurons was comparable to WT. In contrast, the percentage of TUBB3 positive cells in NFIB\textsuperscript{(-)} cells was approximately 5\% (Figure 14), indicating that NFIB\textsuperscript{(-)} cells have reduced neuron differentiation. Since this phenotype is prior to NFIB expression we reasoned that this cell line had a differentiation defect. This might
have been caused during the process of CRISPR targeting and clone selection. NFIB\(^{(c)}\) cells were removed from subsequent analysis.

\textit{In vivo}, SOX9 and NFIA form a complex and function to activate a cohort of target genes, including \textit{Apddd1}, important in glial differentiation (Kang \textit{et al.} 2012). It has been proposed that NFIA is required for \textit{Apddd1} expression. It remains to be demonstrated whether NFIB can also regulate \textit{Apddd1}. In iNPs, \textit{Apddd1} is induced at D9, following NFIA and NFIB. To test the requirement for NFIA and the role of NFIB, we assessed \textit{Apddd1} expression in NFIA\(^{(c)}\), NFIA/B\(^{(c/c)}\) at D5, D7, D9 by qPCR. Consistent with data presented in Chapter 3, at D5 and D7 \textit{Apddd1} expression was not detected in either WT, NFIA\(^{(c)}\) or NFIA/B\(^{(c/c)}\) cells. At D9 \textit{Apddd1} expression in the NFIA\(^{(c)}\) was reduced by 50% relative to the WT. In NFIA/B\(^{(c/c)}\) cells, \textit{Apddd1} was reduced by 72% (Figure 15A). The reduction of \textit{Apddd1} in NFIA-deficient iNPs is consistent with the known function of NFIA in the activation of glial genes with SOX9 (Kang \textit{et al.} 2012). The further reduction observed in the NFIA/B\(^{(c/c)}\) cultures suggests that NFIB can also activate SOX9/NFIA targets.

NFIB has been proposed to regulate another NFI transcription factor, \textit{Nfix} (Matuzelski \textit{et al.} 2017). \textit{Nfix} induction follows NFIA and NFIB, and its expression is reduced in NFIA and NFIB deficient mice. In iNPs, \textit{Nfix} is induced at D9 (Chapter 3). To test whether the regulatory relationship between NFIA/B and \textit{Nfix} is preserved \textit{in vitro}, we measured \textit{Nfix} transcription in WT, NFIA\(^{(c)}\), NFIA/B\(^{(c/c)}\) cells at D5, D7, D9 by qPCR. As reported previously, \textit{Nfix} expression is low at D5 and D7 and induced at D9 in WT cells (Chapter 3). In NFIA\(^{(c)}\) cells, \textit{Nfix} dynamics were comparable to WT. NFIA/B null cells had comparable \textit{Nfix} expression at D5 and D7, but its expression was reduced by 63% at D9 (Figure 15B). These data corroborate previous findings that \textit{Nfix} is regulated by NFIA/B. At the time points analysed we observed no change in \textit{Nfix} expression upon deletion of NFIA. Whether this is due to compensation by NFIB is not clear. Nevertheless, this is in contrast to the reported role of NFIA \textit{in vivo} (Matuzelski \textit{et al.} 2017).

Terminal glial differentiation is characterised by the expression of GFAP, which \textit{in vivo} is expressed at E18.5. NFIA or NFIB deficient mice had a severe reduction in gliogenesis (Deneen \textit{et al.} 2006). It has been suggested that the remaining glial
differentiation in each of the mutants is due to redundancy between the two genes. This predicts that NFIA\(^{-}\) cells should retain some glial differentiation, whereas the double knockout would not. iNPs generate GFAP positive cells at D15 (Chapter 3). To test the requirement for NFIA and NFIB in terminal gliogenesis, we cultured NFIA\(^{-}\) and NFIA/B\(^{(-/-)}\) cells to D15 and assessed the proportion of GFAP positive cells by qPCR, immunofluorescence and flow cytometry. GFAP positive cells were observed in the WT and NFIA\(^{-}\) cells by immunofluorescence but not in the NFIA/B\(^{(-/-)}\) cells (Figure 16A). Quantification of this result by flow cytometry revealed that NFIA\(^{-}\) had \(\sim 1\%\) GFAP positive cells, half that of the WT cells. No GFAP positive cells were seen in the NFIA/B\(^{(-/-)}\) (Figure 16B). This trend was corroborated by qPCR analysis, which showed Gfap expression was reduced by 80\% in NFIA\(^{-}\) compared to WT. The NFIA/B\(^{(-/-)}\) cultures had barely detectable levels of Gfap expression (Figure 16C). Double knockout of NFIA/B abrogates gliogenesis, supporting the hypothesis that there is redundancy between NFIA and NFIB. These data extend previous studies by demonstrating that NFIA and NFIB are required for gliogenesis.

In this section we delete NFIA and NFIB individually and in combination using CRISPR/Cas9 technology in combination with gene specific gRNAs. Unfortunately, NFIB null cells had defective neuron differentiation and could not be analysed further. It has been described previously that NFIA\(^{-}\) and NFIB\(^{-}\) mice have severe defects in gliogenesis (Deneen et al. 2006). The \textit{in vitro} model recapitulates the key features of the glial phenotypes of the NFIA knockout. Moreover, the loss of NFIA and NFIB results in the absence of GFAP positive glia. These data agree with the proposal that redundancy exists between NFIA and NFIB as well as emphasising their requirement for proper gliogenesis. SOX9 and NFIA act in combination to activate glial genes such as Apcdd1 (Kang et al. 2012). Consistent with this, deletion of NFIA results in reduced Apcdd1 expression. A more severe reduction was observed in NFIA/B\(^{(-/-)}\) cells suggesting a role for NFIB in Apcdd1 induction and glial lineage activation. Taken together these data provide functional validation of the \textit{in vitro} model as well as confirming the importance of NFIA/B in the glial lineage.
4.2 Loss of gliogenesis upon deletion of SOX9

Spinal cord neural progenitors first generate neurons before transitioning to gliogenesis. SOX9 is expressed in advance of NFIA and NFIB appearing coincident with neuron differentiation. SOX9 has been proposed to regulate gliogenesis in combination with NFIA and NFIB. For example, addition of SOX9 to NFIA and NFIB improved the efficiency of reprogramming mouse embryonic fibroblasts to GFAP positive glia (Caiazzo et al. 2015). Additionally, deletion of SOX9 in vivo causes a delay in NFIA induction and a subsequent delay in the activation of SOX9/NFIA target genes such as Apedd1 (Kang et al. 2012). At later time points SOX9 knockout mice also have a severe reduction in differentiated glia (Stolt et al. 2003). To further validate the in vitro system, we deleted SOX9 in ES cells and measured their ability to transition to glial progenitors and undergo terminal differentiation.

We used CRISPR/Cas9 to target Sox9. Guides were designed to the first of the three coding exons. Sequencing of the resulting clones identified a clone with a 1 base pair deletion in both alleles. In iNPs SOX9 is expressed at D7. To test whether the 1 base pair deletion was sufficient to abrogate SOX9 protein we differentiated SOX9(-) ES cells to D7 and assessed SOX9 expression by immunofluorescence. As described in Chapter 3, at D7 SOX9 was expressed in SOX2-positive progenitors in WT ES cells. No expression was detected in SOX9(-) cells (Figure 17A). We confirmed this result by western blot (Figure 17B). We conclude that the frame shift resulting from a 1 base pair deletion results in a loss of SOX9 protein.

Neuron differentiation in the mouse developing spinal cord begins at approximately E9 and in vitro at D6 (Kicheva et al. 2014). SOX9 expression appears coincident with the onset of neuron differentiation. Although deletion of SOX9 has been reported to prolong the phase of neuron differentiation, it has not been proposed to effect neurogenic onset (Stolt et al. 2003). However, previous work has deleted SOX9 at E10.5 after the start of neuron production. To test whether removal of SOX9 affects neuron differentiation we measured the neuron marker, TUBB3, in SOX9(-) and WT cells by flow cytometry at D7. In WT cells ~15% are TUBB3 positive. The percentage of TUBB3 positive cells in SOX9(-) was comparable to WT
cultures (Figure 18A). We confirmed this result by performing a time course of Tubb3 expression from D5 to D11 by qPCR. In WT cells Tubb3 expression is low at D5 and increases to D11. The SOX9(-) mimicked the WT dynamics, with low levels of expression at D5, steadily increasing to D11 (Figure 18B). This demonstrates that neuron differentiation is unaffected upon removal of SOX9. Although not statistically significant, there appeared a small increase in Tubb3 expression in SOX9(-) cells compared to WT at D11, perhaps suggesting the phase of neuron differentiation is also extended in vitro. Late neuron differentiation may be difficult to assay in a system which lacks the presence of neurotrophic factors required for neuronal survival. Nevertheless, SOX9(-) iNPs generate neurons with a comparable dynamic to WT cells.

In vivo the deletion of SOX9 results in an approximately 24 hour delay in NFIA (Kang et al. 2012). Whether NFIB is also delayed in SOX9 mutant mice is not clear. To test the function of SOX9 in vitro, we differentiated SOX9(-) and WT cells to D11 after NFIA induction, and assayed NFIA/B expression every 48 hours from D5. At D7 Nfia expression is low and there was no significant difference between WT and SOX9(-). However, by D9, there was a 40% reduction in Nfia expression in the SOX9(-) that was maintained at D11. Similarly, at D9 and D11 Nfib expression was reduced in SOX9(-) cells relative to WT (Figure 19A). Next, we sought to determine the effect of SOX9 deletion on NFIA and NFIB protein levels using immunofluorescence and flow cytometry. No NFIA expression was detected at D7 in WT or mutant cells by immunofluorescence. At D9, NFIA expression was observed in both WT and SOX9(-) iNPs. There was no obvious reduction in NFIA protein intensity in iNPs lacking SOX9 (Figure 17A). We measured the proportion of SOX2 positive cells (progenitors) that were also NFIB positive by flow cytometry. In WT cells, NFIB is induced from D7 to D9, increasing to 95% at D11. In SOX9(-) cells there were no NFIB-positive cells at D9. At D11, the percentage of NFIB-positive had risen to 65% (Figure 19B). In summary, the levels of Nfia and Nfib expression are reduced in the SOX9(-), demonstrating that SOX9 is required for proper induction of NFIA and NFIB. This phenotype is confirmed for NFIB by flow cytometry. However, no obvious reduction was observed for NFIA by immunofluorescence in SOX9 mutant cells. Whether Nfia and NFIB expression
recovers to WT levels similar to \textit{in vivo} is unclear due to the time resolution of the experiment.

SOX9 and NFIA are both required to co-activate a set of target genes important in gliogenesis (Kang \textit{et al.} 2012). We predicted that the expression of SOX9/NFIA targets would be reduced in SOX9\(^{-}\) cells. To test this, we performed a time course analysis of SOX9\(^{-}\) and WT cultures to D11 and assayed the expression of \textit{Apcdd1} by qPCR. At D7 there was little if any expression of \textit{Apcdd1} in either WT or SOX9\(^{-}\) cells. In WT cells, expression increases between D7 and D9. In contrast, SOX9 null cells at D9 showed a 72\% reduction. By D11, \textit{Apcdd1} expression had increased in SOX9\(^{-}\) cells but was still 54\% of WT (Figure 19C). This reduction is comparable to that observed in NFIA/B\(^{-/-}\) cells suggesting that SOX9 and NFIA/B are equally required for glial lineage activation. Taken together with \textit{Nfia/b} dynamics, SOX9 is required for the proper level of expression of glial genes.

\textit{Nfix} expression, like \textit{Apcdd1}, follows NFIA and NFIB \textit{in vivo} and in iNPs (Matuzelski \textit{et al.} 2017). Although we and others have shown that NFIA and NFIB regulate \textit{Nfix} (this Chapter), SOX9 has not been implicated. To test the effect removal of SOX9 has on \textit{Nfix}, we measured \textit{Nfix} expression by qPCR in WT and SOX9\(^{-}\) iNPs at D5, D7, D9. Consistent with our previous data \textit{Nfix} was not expressed at D5 or D7 in WT or SOX9\(^{-}\) cells (Chapter 3). In WT cells, \textit{Nfix} expression increased at D9. In contrast, SOX9\(^{-}\) had a 60\% reduction in \textit{Nfix} transcription. This demonstrates a role for SOX9 in the regulation of \textit{Nfix} although whether NFIA and NFIB are required for this phenotype is not clear. The level of reduction is comparable to that observed in NFIA/B\(^{-/-}\) cells. This is quantitatively similar to the effect of SOX9\(^{-}\) and NFIA/B\(^{-/-}\) on \textit{Apcdd1} suggesting \textit{Nfix} is regulated by a similar mechanism.

After activation of the glial lineage progenitors migrate away from the progenitor zone and terminally differentiate into GFAP-positive glia. Deletion of SOX9 results in extended neuron differentiation at the expense of gliogenesis (Stolt \textit{et al.} 2003, Kang \textit{et al.} 2012). As the SOX9\(^{-}\) display the expected phenotypes \textit{in vitro} we predicted a defect in terminal glial differentiation. To test this, we differentiated the SOX9\(^{-}\) and WT cells to D15 and measured GFAP by qPCR and flow cytometry. In
WT cells, there is ~3% GFAP-positive cells at D15. In contrast, SOX9\(^{-}\) cells had almost undetectable numbers of GFAP-positive cells and \textit{Gfap} transcript, measured by qPCR, was barely detectable in the SOX9\(^{-}\). Although low, deletion of SOX9 had increased \textit{Gfap} expression compared to NFIA/B\(^{-}\) cells (Figure 19D). In conclusion, SOX9\(^{-}\) has a severe reduction in GFAP-positive glia. This phenotype recapitulates observations \textit{in vivo}, emphasising the requirement of SOX9 in gliogenesis.

\textit{In vivo} conditional knockout of SOX9 results in a delay in NFIA/B and impaired gliogenesis (Stolt \textit{et al.} 2003, Kang \textit{et al.} 2012). Knockout of SOX9 had no effect on the onset of neuron differentiation but exhibited reduced expression of NFIA/B and \textit{Apcdd1}, as well as a defect in terminal glial differentiation. This recapitulates known phenotypes of SOX9 observed \textit{in vivo}. The SOX9\(^{-}\) phenotypes appear comparable to that of NFIA/B\(^{-}\) cells in agreement with previous work which show that SOX9/NFIA cooperate in the activation of the glial lineage (Kang \textit{et al.} 2012). Additionally, we demonstrate a role for SOX9 in NFIX regulation. A complete loss of NFIA/B, \textit{Apcdd1}, or NFIX was not observed in SOX9\(^{-}\) cells suggesting alternate mechanisms exist for their induction. At later time points few glia were generated in the absence of SOX9.

Together, these data validate the \textit{in vitro} differentiation protocol as a model of the neural-to-glial fate transition in the developing spinal cord. The model reproduces the \textit{in vivo} phenotypes of SOX9 and NFIA. It provides new insight into the mechanism underlying the transition by showing the redundancy between NFIA and NFIB. The \textit{in vitro} system can therefore be used to further investigate the mechanism underlying the neural-to-glial fate transition in the developing spinal cord.
Figure 12 Confirmation of NFIA deletion in NFIA and NFIA/B knockouts

NFIA expression in WT, NFIA\(^{-}\), NFIB\(^{-}\), NFIA/B\(^{-}\) cells at D10 (n = 3). (A) Immunofluorescence images for DAPI, SOX9, NFIA, SOX2 in WT and knockout cells. Images are at 20x. Scale bars = 40µM. (B) Quantification of immunofluorescence images. Each point represents the mean NFIA intensity within SOX2-positive nuclei in one image. Error bars represent standard error of the mean. (C) Western blot for NFIA and NFIB, β-TUBULIN (TUB) in WT and knockout cells.
Figure 13 Confirmation of NFIB deletion in NFIA and NFIA/B knockouts

NFIB expression in WT, NFIA\textsuperscript{(−)}, NFIB\textsuperscript{(−)}, NFIA/B\textsuperscript{−/−} cells at D10. (A) Immunofluorescence images for DAPI, SOX2, NFIB, TUBB3 in WT and knockout cells. Images are at 20x. Scale bars = 40µM. (B) Quantification of immunofluorescence images. Each point represents the mean NFIB intensity within SOX2-positive nuclei in one image. Error bars represent standard error of the mean.
Chapter 4 Results

Figure 14 Neuron differentiation is affected in NFIB\(^{(3)}\) but not NFIA\(^{(4)}\), NFIA/B\(^{(4\,4)}\)
The percentage of TUBB3-positive, SOX2-negative cells (neurons) measured by flow cytometry in WT, NFIA\(^{(4)}\), NFIB\(^{(3)}\), NFIA/B\(^{(4\,4)}\) cells at D7. Error bars represent standard error of the mean. \(* p = 0.0167\), Mann-Whitney unpaired \(t\) test.

Figure 15 Glial lineage activation is impaired in NFIA\(^{(4)}\), NFIA/B\(^{(4\,4)}\)
Expression of Apcdd1 and Nfix at D5, D7, D9 in WT, NFIA\(^{(4)}\), NFIA/B\(^{(4\,4)}\) cells measured by qPCR (n = 3-7). Error bars represent standard error of the mean. \(* p = 0.0476\), ** \(p = 0.0012\), Mann-Whitney unpaired \(t\) test.
Figure 16 Abrogation of terminal gliogenesis in NFIA/B^{+/−} cells

Expression of the terminally differentiated glia marker GFAP at D15 in SOX9^{−/−} and WT cells (n = 3-8). (A) Immunofluorescence images for DAPI, GFAP (green), SOX2 (red), TUBB3 (blue). (B) Percentage of GFAP+ cells measured by flow cytometry. * p = 0.0121, ** p = 0.004, Mann-Whitney unpaired t test. (C) Gfap expression measured by qPCR. Values are normalised to the WT within each experiment. (B-C) Error bars show standard error of the mean.
Figure 17 Confirmation of SOX9 deletion in SOX9(-) cells

SOX9 expression in SOX9(-) and WT cells. (A) Immunofluorescence for DAPI, NFIA, SOX9, SOX2 at D7 and D9. Images are at 20x. Scale bars = 40µM. (B) Western blot for SOX9 (red) and -TUBULIN (TUB, green) at D8. In the WT the upper red band indicates SOX9 at ~75kDa. The lower red band is non-specific.
Figure 18 Neuron differentiation is unaffected by deletion of SOX9
TUBB3 expression in WT and SOX9\(^{-}\) (blue) cells (n = 5-9). (A) The percentage of TUBB3+, SOX2- cells (neurons) at D7 measured by flow cytometry. ns p = 0.0599, Mann-Whitney unpaired t test. (B) Tubb3 expression measured by qPCR from D5 to D11. (A-B) Error bars show standard error of the mean.
Figure 19 The expression of glial lineage genes is reduced in SOX9\(^{(+/−)}\) cells

Expression of glial genes in SOX9\(^{(+/−)}\), NFIA/B\(^{(+/−)}\), WT cells through the differentiation (n = 4-12). (A, C) Nfia, Nfib, Apcdd1, Nfix expression measured by qPCR. (B) Percentage of NFIB-positive cells measured by flow cytometry. Cells are first gated on SOX2. (A-C) Measurements are made at D5, D7, D9, D11. (D) GFAP (glia) expression measured by qPCR (left) and flow cytometry (right). ** p = 0.0055, * p = 0.0185, Mann-Whitney unpaired t test. (A-D) Error bars represent standard error of the mean.
Chapter 5. Manipulating TGF-β signalling affects the timing of the neural-to-glial transition

In the developing spinal cord, the neural-to-glial transition is marked by the onset of NFIA and NFIB expression (Chapter 4, (Deneen et al. 2006)). Understanding the mechanism of NFIA/B induction is therefore important in determining how the timing of the transition is controlled. TGF-β signalling has been implicated in the regulation of developmental transitions in multiple regions of the central nervous system. For example, GDF11 is an activator of the TGF-β signalling pathway, and is expressed by neurons across the dorsal-ventral axis between E10 and E12.5 (Shi and Liu 2011). In GDF11 knockout mice the rate of neuron differentiation is slower and the temporal progression of neurons is affected. In addition, activation of TGF-β signalling induces an early transition from HB9-positive somatic motor neurons to SOX10-positive oligodendrocyte progenitors (Dias et al. 2014). It is unclear whether activating TGF-β signalling affects glial genes such as SOX9, NFIA and NFIB. Obtaining precise time resolution in vivo is difficult. Since the in vitro system recapitulates the neural-to-glial transition, it represents a tractable system to test whether TGF-β signalling plays a role in determining the timing of glial gene activation.

In the mouse, all three TGF-β isoforms appear to be expressed in the developing spinal cord at E11.5 (Mecha et al. 2008). TGF-β1 is expressed across the neural tube, TGF-β2 expression is higher in neurons than in progenitors, TGF-β3 is expressed in neurons. However precise time resolution is missing. We sought to resolve the expression pattern of TGF-β1/2/3 in the developing spinal cord across the neural-to-glial transition. To understand this, we performed immunofluorescence at E10.5 and E12.5 in the mouse brachial neural tube using a TGF-β1/2/3 antibody. Little to no staining was observed at E10.5. At E12.5 no expression was observed in the ventricular zone. Adjacent to the progenitors, low levels of TGF-β1-3 were detected in the mantle zone where differentiated neurons reside (Figure 20).
SMAD7 is a transcriptional target of TGF-β signalling and acts as a negative regulator of the pathway (Nakao et al. 1997). SMAD7 expression is therefore a molecular readout of TGF-β signalling. Since TGF-β1-3 ligands are present in vivo during the neural-to-glial transition we hypothesised there might be endogenous TGF-β signalling in vitro. To test this, we measured Smad7 expression daily from D3 to D11 by qPCR. Low levels of Smad7 were detected at D3. From D3 to D5 there was an approximately 400% increase. A further 50% increase was observed between D5 and D7. Smad7 expression then decreased to D11 (Figure 21A).

Smad7 is expressed during the neural-to-glial transition. These data suggest there is endogenous TGF-β signalling in the in vitro system.

Methods to manipulate TGF-β signalling have been well described. Addition of TGF-β2 to in vitro derived hindbrain neural progenitors is sufficient to induce an early transition from motor neurons to serotonergic neurons (Dias et al. 2014). SB431542 (SB) is a chemical inhibitor of TGF-β receptors (Inman et al. 2002). To test the effect of TGF-β2 and SB we took advantage of the endogenous Smad7 expression in vitro. Since we were interested in the effect of TGF-β on the neural-to-glial transition we added TGF-β2 or SB at D5 just prior to the onset of neuron differentiation. iNPs were exposed to TGF-β2 for 48 hours or SB continuously, and Smad7 expression assessed by qPCR. Exposure to TGF-β2 from D5 to D7 increased Smad7 expression by 25% compared to controls at D7. At D9, expression returned to control levels. By contrast, inhibition with SB resulted in a 50% decrease in Smad7 expression at D7. Expression recovered slightly at D9 (Figure 21B). Thus, activating TGF-β signalling induces a transient induction of Smad7 consistent with Smad7 being a transcriptional target of TGF-β signalling. Moreover, the reduction in Smad7 induction when TGF-β signalling is inhibited confirms the presence of endogenous TGF-β signalling in vitro. Together, these data demonstrate we can manipulate TGF-β signalling in iNP cultures.

In the spinal cord, deletion of GDF11 did not affect the timing of neuron differentiation, but reduced the rate at which neurons were produced (Shi and Liu 2011). However, the loss of GDF11 may be compensated for by other TGF-β ligands. To test the effect of TGF-β signalling on neuron differentiation, we exposed iNPs either to TGF-β2 from D5 to D7 or continuous SB from D5 and measured the
neuron marker *Tubb3* by qPCR at D7 and D9. At D7, TGF-β2 treated cells increased *Tubb3* expression by 25%, returning to control levels at D9. In continuous SB treatment there was a 25% reduction in *Tubb3* expression compared to the control at D7 which returned to control levels at D9 (Figure 22B). Manipulating TGF-β signalling has a transient effect on neuron differentiation. Later neurogenic phenotypes are difficult to measure *in vitro* as a lack of neurotrophic factors may affect the survival of differentiated neurons. Nevertheless, these data are consistent with TGF-β signalling controlling the rate of neuron differentiation.

Both *in vivo* and *in vitro* the induction of SOX9 correlates with the onset of neuron differentiation (Chapter 3, (Stolt et al. 2003)). Since manipulating TGF-β signalling had a transient effect on *Tubb3* expression, we hypothesised that it might also affect SOX9 expression. To test this, we measured Sox9 transcription by qPCR at D7 and D9 in iNPs exposed to TGF-β2 for 48 hours or continuously treated with SB from D5. TGF-β2 treated cells at D7 showed a 60% increase in Sox9 expression over the control. In contrast, in iNPs continuously treated with SB Sox9 expression was 50% less than the control at D7. In both cases expression had returned to control levels at D9 (Figure 23). Activation of TGF-β signalling induces a transient increase in Sox9 expression whilst continuous SB treatment reduces the level of Sox9 expression at D7 but is insufficient to block Sox9 expression at D9. This suggests alternative mechanisms might exist for Sox9 induction. Whether TGF-β signalling directly regulates Sox9 or the effects are a consequence of manipulating neuron differentiation is unclear.

In the developing spinal cord, NFIA and NFIB expression marks the onset of the glial lineage. *In vivo* expression begins at E11.5 and *in vitro* at D8 (Chapter 3, (Deneen et al. 2006)). Deletion of SOX9 results in a delay in NFIA and NFIB *in vivo* and *in vitro* (Chapter 4, (Kang et al. 2012)). We predicted that the changes in Sox9 expression from manipulating TGF-β signalling might be sufficient to affect *Nfia* and *Nfib* transcription. To test this, we measured *Nfia* and *Nfib* expression at D7 and D9 using qPCR in cells exposed to 48 hours of TGF-β2 or continuously treated with SB. At D7 *Nfia* expression was increased 100% and at D9 50%, in cells treated with TGF-β2. In continuous SB treatment *Nfia* transcription was reduced to half of the control at both D7 and D9. In TGF-β2 treated cells *Nfib* expression was 100%
higher at D7 compared to controls. At D9, expression was 50% higher. Continuous SB treatment resulted in approximately half the control level of Nfib at D7 which recovered slightly at D9 (Figure 24A). We confirmed this result by antibody staining. As reported previously, at D9 NFIA expression is low in control cells (Chapter 4). An increase in NFIA expression was observed in cells treated with TGF-β2 for 48 hours (Figure 24C). At D10 NFIA is expressed. In cells continuously treated with SB, mean NFIA intensity in progenitors was 50% of the control (Figure 24B, D). Activating or inhibiting TGF-β signalling increases or decreases NFIA and Nfib expression accordingly. The effect on NFIA and Nfib expression is quantitatively similar in iNPs either treated with TGF-β2 or SB. This suggests TGF-β signalling is affecting their expression either directly or by a conserved mechanism. Whether the effect is mediated by SOX9 is not clear.

NFIA and Nfib regulate APCDD1 and NFIX expression, and consequently their transcription follows NFIA and Nfib (Chapter 3, Chapter 4 (Kang et al. 2012, Matuzelski et al. 2017)). We predicted that the changes in the expression of Sox9, Nfia, Nfib resulting from manipulating TGF-β signalling would affect the activation of downstream genes. To test this, we analysed Apcdd1 and Nfix expression by qPCR in cells treated with TGF-β2 or SB. At D7 no difference was detected in Apcdd1 expression in TGF-β2 treated, SB treated and control cells. At D9 Apcdd1 expression in TGF-β2 treated cells was increased 100% compared to the control. In continuous SB treatment Apcdd1 expression remained at control levels at D7 and was reduced by 50% relative to the control at D9. In cells treated with TGF-β2 there was a 100% increase in Nfix expression over the control at D7, which increased to 150% at D9. In cells continuously exposed to SB Nfix expression was half of the control at both D7 and D9 (Figure 25). Manipulating TGF-β signalling affects Nfix expression similar to Nfia/b. This suggests Nfix induction correlates with the kinetics of Nfia/b expression. Activating TGF-β signalling only increases Apcdd1 expression at D9 and continuous inhibition blocked activation of Apcdd1 at these time points. These data suggest that TGF-β manipulations are sufficient to shift the timing of the neural-to-glial transition.

Continuous inhibition of TGF-β signalling in vitro is sufficient to cause a reduction in the glial genes Sox9, Nfia, Nfib, Apcdd1, Nfix. We predicted that changing the
duration of SB treatment would result in intermediate expression of glial genes. To test this, we treated iNPs with SB for either 48 hours or continuously from D5 and assayed Nfia, Nfib expression by qPCR at D9. As described previously continuous treatment with SB resulted in a 50% reduction in Nfia expression. A 30% reduction was observed in cells treated with SB from D5 to D7. In continuous SB treatment Nfib expression was reduced by 35%. In cells treated with SB for 48 hours Nfib expression was reduced by 30% (Figure 26A). We confirmed this result by measuring NFIA intensity in SOX2-positive progenitors at D10 and D11 by immunofluorescence. In cells continuously treated with SB, the mean NFIA intensity in progenitors was 30% of the control at D10 and 45% of the control at D11. In cells treated with SB for 48 hours NFIA intensity was 45% of the control at D10 but had recovered by D11. Therefore, varying the duration of SB treatment shifts the timing of NFIA induction. A 48 hour inhibition of TGF-β signalling is sufficient to delay glial lineage onset.

In the developing spinal cord SOX9 is required for the correct timing of the neural-to-glial transition (Chapter 4, (Kang et al. 2012)). Activation of TGF-β signalling in vitro is sufficient to induce a transient increase in Sox9 expression and subsequently the increased expression of glial genes. To test whether the induction of glial genes was a consequence of increased Sox9 expression, we differentiated WT and SOX9(-) cells and activated TGF-β signalling from D5 to D7 with TGF-β2. We then measured Sox9, Nfia, Nfib, Apcdd1, Nfix induction by qPCR at D8. At this time point, Sox9 was not induced by TGF-β2. WT cells showed a 1.5-2-fold induction of Nfia, Nfib, Apcdd1. Nfix expression was approximately 3-fold higher in treated cells. In SOX9(-) cells Nfia, Nfib, Apcdd1, Nfix all showed similar levels of induction in response to TGF-β activation (Figure 27). Deletion of SOX9 does not reduce the effect activating TGF-β signalling has on glial genes indicating that the increased Sox9 expression is not sufficient to explain the function of TGF-β signalling. The induction of Sox9 expression observed at D7 is not observed at D8. We cannot exclude the possibility that SOX9 is required for the increase in Sox9 expression observed after TGF-β2 exposure.

Manipulating TGF-β signalling in vivo has been shown to affect developmental transitions in the developing midbrain, hindbrain and spinal cord (Dias et al. 2014).
In the developing spinal cord, deletion of the TGF-β ligand GDF11 resulted in a reduced rate of neuronal differentiation (Shi and Liu 2011). Overexpression of a constitutively active TGF-β type I receptor induced a transition from motor neuron progenitors to oligodendrocyte progenitors (Dias et al. 2014). However, whether TGF-β signalling affects the neural-to-glial transition was not clear. In this chapter we show that TGF-β1-3 are present in differentiated neurons at E12.5 during the neural-to-glial transition. We used the in vitro system to test the effect of TGF-β signalling on the activation of glial genes. Activating or inhibiting TGF-β increased or reduced respectively the expression of Tubb3, Sox9, Nfia, Nfib, as well as downstream target genes Apcdd1 and Nfix suggesting these manipulations were affecting the timing of the neural-to-glial transition. We also show that changing the duration of TGF-β inhibition can control the timing of NFIA and NFIB expression. We tested whether the effect of TGF-β was via SOX9 but SOX9\(^{(1)}\) cells responded comparably to WT cells upon TGF-β activation. The ability to shift the dynamics of glial genes by manipulating TGF-β signalling will provide a valuable tool to investigate genes involved in the neural-to-glial transition.
Figure 20 TGF-β1-3 are expressed at E12.5 with NFIA

Immunofluorescence for TGF-β1-3 (blue), NFIA (purple) on transverse sections of mouse brachial neural tubes at E10.5 and E12.5. Images at 20x. Scale bar = 40μM.
Chapter 5. Results

Figure 21 TGF-B signalling is active in vitro

*Smad7* expression measured by qPCR. (A) Expression from D3 to D11 in vitro. **p = 0.0078, Wilcoxon paired t test. (B) Expression in cells either exposed to TGF-β for 48 hours or continuously treated with SB from D5. Values are normalised to untreated cells. (A,B) Error bars represent standard error of the mean.

Figure 22 Manipulating TGF-β signalling affects neuron differentiation

qPCR expression data for *Sox2* and *Tubb3* in iNPs either exposed to TGF-β for 48 hours or continuously treated with SB from D5. Values are normalised to untreated cells. Error bars represent standard error of the mean.
Figure 23 Manipulating TGF-β signalling affects Sox9 expression

Sox9 expression measured by qPCR in iNPs either exposed to TGF-β for 48 hours or continuously treated with SB from D5. Values are normalised to untreated cells. Error bars represent standard error of the mean.
Figure 24 Manipulating TGF-β signalling affects the onset of gliogenesis

Expression of NFIA and NFIB in cells exposed to 48 hours of TGF-β2 of continuously treated with SB. (A) Nfia, Nfib expression measured by qPCR. Values are normalised to the control. Error bars represent standard error of the mean. (B) Mean NFIA intensity in SOX2 positive nuclei from immunofluorescence images. Each data point is the average of 4-6 images taken from one experiment. Data is normalised to the control. Error bars represent standard error of the mean. (C-D) Immunofluorescence for NFIA, SOX2, TUBB3 at D9 (C) and D10 (D). Images at 20x. Scale bars = 40μm.
Figure 25 Manipulating TGF-β signalling affects glial lineage activation

qPCR expression data for *Apcdd1*, *Nfix* in iNPs either exposed to TGF-β for 48 hours or continuously treated with SB from D5. Values are normalised to untreated cells. Error bars represent standard error of the mean.
The timing of the neural-to-glial transition varies with the length of TGF-β inhibition

Analysis of NFIA and NFIB in cells either continuously treated with SB or exposed between D5 and D7. (A) qPCR data for Nfia, Nfib at D9. Error bars show standard error of the mean. * p = <0.05, Mann-Whitney unpaired t test. (B) NFIA immunofluorescence in the same experiment. Images at 20x. Scale bars = 40μm. (C) Mean NFIA intensity in SOX2 positive nuclei from immunofluorescence. Each point represents the data from one image. Data is normalised to the average intensity of control images. Error bars represent standard error of the mean. *** p = <0.0005, Mann-Whitney unpaired t test.
Figure 27 Sox9 is not required for the effect of TGF-β signalling on the neural-to-glia transition

qPCR expression data for Sox9, Nfia, Nfib, Apcdd1, Nfix at D8 in cells exposed to TGF-β2 between D5 and D7. The data is normalised to untreated cells.
Chapter 6. Reconstructing the neural-to-glial transition by single cell sequencing

During the development of the posterior neural tube, neurons and glia are generated sequentially. The onset of gliogenesis is characterised by the induction of NFIA and NFIB (Chapter 4, (Deneen et al. 2006)). Defining the mechanism of their activation moves towards an understanding of the timing of the neural-to-glial transition. Knockout of SOX9 or inhibition of TGF-β signalling delays Nfia, Nfib induction (Chapter 4.2, Chapter 5, (Kang et al. 2012)). In neither experiment is the transition to gliogenesis completely abrogated, suggesting that alternate mechanisms exist. The transcriptomes of single cells harvested through time can be used to reconstruct differentiation trajectories in silico (Peng and Han 2018). Gene expression dynamics plotted along ‘pseudotime’ allows the identification of genes involved in developmental processes. The complexity of the in vivo system means heterogeneity is a problem. The in vitro model recapitulates the expression patterns of glial genes and gives greater control over cellular heterogeneity. We utilised single cell transcriptomics, in combination with the in vitro system, to identify candidate genes involved in the neural-to-glial transition.

Both in vivo and in vitro neurons and glia are generated sequentially (Rowitch 2004). In vitro, neuron differentiation begins at D6. The onset of the glial lineage is characterised by Nfia, Nfib expression which begins at D8 in vitro (Chapter 3). Single cell sequencing from D5 neural progenitors had already been performed in the lab by Andreas Sagner. To reconstruct the gene expression profiles of progenitors transitioning to gliogenesis, we harvested samples from D6 to D9. The Fluidigm-C1 system utilises microfluidics to isolate single cells. However, this process is imperfect, since often multiple cells are captured in individual wells. To ensure only single cells were analysed, each well was imaged and manually checked. Only wells annotated as containing single cells were sequenced. Through this approach, we obtained the transcriptomes of 192 single cells across the neural-to-glial transition; 45 from D5, 41 from D6, 39 from D7, 51 from D8, 16 from D9.
Before proceeding with sequencing, we confirmed the samples were derived from representative differentiations. To test this, we analysed Nfib kinetics by qPCR in iNPs collected daily through the differentiation. To provide a reference, we averaged Nfib expression data from all previous experiments. In all cases, the samples used for single cell sequencing came from experiments with comparable dynamics to the reference data (Figure 28). Thus, the samples used for single cell sequencing follow the typical differentiation trajectory.

In vivo, spinal cord neural progenitors derive from a population of cells which also give rise to trunk mesoderm, a process we recapitulate in vitro (Wilson et al. 2009, Gouti et al. 2014). Additionally, neural progenitors generate neurons from D6 (Chapter 3, (Sagner et al. 2018)). Therefore, in vitro populations comprise neural progenitors, neurons and trunk mesoderm lineages. To remove these populations, Julien Delile utilised an improved version of an analysis pipeline used previously (see Materials and Methods 2.3 (Sagner et al. 2018, Delile unpublished)). The method uses correlation to identify clusters of genes – modules – with concerted patterns of expression in the dataset. This identified 40 gene modules describing 184 cells. 12 cells were not represented by any gene module. 5 gene modules contained neuronal genes such as Tubb3, Dcx, Lhx1, Lhx5, Lbx1. 3 gene modules contained Meox1, Foxc2, Twist1, markers of mesoderm. To identify neuronal and mesodermal cells we re-clustered the data using the genes belonging to the neuronal and mesodermal modules. This identified 25 neurons (14%) and 19 mesodermal cells (10%) (Figure 29). Removal of neurons, mesoderm and unknown cells left 136 single cell transcriptomes.

A variety of methods exist for the reconstruction of developmental lineages from single cell data (Wu et al. 2017). In general, they use subsets of varying genes with dimensionality reduction algorithms to order cells. We hypothesised that by taking a similar approach we could reconstruct the neural-to-glial transition in silico. To achieve this, Julien Delile first selected the 4 gene modules whose averaged expression showed significant positive or negative correlation with sample time (Spearman correlation (r > 0.6). 157 genes made up module 1 and had high expression at D5, decreasing gradually with sample time. Module 2 contained 31 genes, including Sox9, Nfib. Module 3 comprised 35 genes and included Lmx1a.
Both modules label cells collected at later time points. Module 4 contained 15 genes, including \( \text{Dbx1} \), and marked cells harvested at D5 (Figure 30).

Principle Component Analysis (PCA) is a dimensionality reduction technique. We decided to use PCA to try to order the cells by developmental time. However, in PCA each variable is weighted equally, to overcome the bias that the different gene module sizes would have on the PCA, the gene expression within each of the 4 modules was averaged for each cell. After PCA, 52% of the variance was captured along the first principle component (PC1), 20% was captured on PC2, 18% on PC3 and 10% on PC4 (Figure 31A). On PC1 cells were ordered by sample collection day, suggesting this component captured developmental time. To investigate the influence of each of the four gene modules on the pseudotemporal ordering, we examined the loadings of each gene module to understand their contribution to the variance captured on PC1. Gene modules 1 and 4, containing genes decreasing with time, contributed negatively, whilst modules 2 and 3, containing genes increasing with time, contributed positively to PC1 variance (Figure 31B). All four modules contribute to the ordering along PC1.

Single cell sequencing returns the transcriptomes of individual cells. The small amount of starting material means the data are noisy. To capture patterns of gene expression, data are smoothed. A variety of smoothing methods exist, but all methods rely on averaging a series of adjacent points. To identify outliers, we plotted a histogram of cell counts across 30 equal segments of PC1. One cell separated away from the remaining points (Figure 32A). The outlier was removed from the PCA (Figure 32B). There was good overlap between sample days across PC1, suggesting a pseudotemporal ordering could be reconstructed.

To smooth the data, we used a polynomial function fitted with local regression, LOESS (Cleveland and Devlin 1988). A variable of LOESS is span, a smoothing parameter which specifies the bandwidth of the local regression. Smaller span values capture more variance in the data, whereas larger spans remove variance. \( \text{Sox9}, \text{Nfia} \) and \( \text{Nfib} \) are genes induced across the neural-to-glial transition. To avoid over- or underfitting the data, we plotted a range of span values from 0.5 to 1 for \( \text{Sox9}, \text{Nfia}, \text{Nfib} \) across PC1 and qualitatively assessed the effect of the
smoothing parameter. All span values tested gave similar fits. A span value of 1 gave a smooth fit for Sox9, Nfia, Nfib, removing all variation from the data. A span of 0.5 returned fluctuations in the fit, most obviously for Sox9. At 0.7 fewer fluctuations were observed in the fit for Nfia, Nfib, but fluctuations remained up to a span value of 0.8 for Sox9 (Figure 32C). A span value of 0.7 represented a balance between smoothing and capturing variation. The largest variations in fit were observed in regions of PC1 with low cell numbers or where cells had low expression. Applying LOESS across all genes returned their smoothed expression dynamics across pseudotime. We assumed that, similar to other glial genes, candidates would be dynamically expressed across the transition. To remove genes with low expression, we applied a mean expression threshold. 4769 genes had an average expression greater than 1.4 (log value) (47%) (Figure 33A). To find dynamically expressed genes, we calculated the index of dispersion of each smoothed gene expression profile through pseudotime. 176 of 4769 genes (1.7% of the total) including 26 transcription factors (1.9% of the total) had a dispersion score greater than 1.5 standard deviations from the mean index (Figure 33B). Using mean expression and dispersion identified 176 genes with dynamic expression across pseudotime.

To cluster the 176 candidate genes, Julien Delile used Affinity Propagation clustering (Frey and Dueck 2007, Delile unpublished). We manually asked for 12 clusters. 4 clusters contained genes decreasing with time. Gene Ontology analysis identified enrichment of miRNA metabolic process, tetrahydrofolate metabolic process, and ncRNA processing in these clusters. This reflected the presence of Lin28a, Lin28b, Trim71, part of the LIN28/let-7 pathway in these clusters. Two clusters contained Irx3 and Sp8, markers of intermediate neural tube (Figure 34) (Bosse et al. 1997, Li et al. 2014). 4 clusters comprised genes increasing with pseudotime. 3 of these clusters included Sox9, Nfia, Nfib, Slc1a3, Fabp7, genes known to be induced during the neural-to-glial transition (Stolt et al. 2003, Deneen et al. 2006). Also present were Msx1 and Zic3, markers of dorsal neural progenitors (Figure 35). The expression pattern of known transition genes, Sox9, Nfia, Nfib, is consistent with the known induction of these genes during the glial transition. Decreasing Irx3, Sp8 expression in combination with increasing Msx1, Zic3 expression suggests progenitors are acquiring more dorsal fates (Davidson
The incidence of multiple members of the LIN28/let-7 miRNA processing pathway suggests this pathway might be important.

The onset of NFIA and NFIB follows the expression of SOX9 (Kang et al. 2012). This pattern is reproduced in the pseudotemporal reconstruction. 26 transcription factors had dynamic expression across pseudotime, including Sp8, Lin28a, Lin28b, Cux2, Npas3 and Zfp536. To further validate the predictions made by the pseudotemporal reconstruction, we used qPCR to measure the expression dynamics of Sp8, Lin28a, Lin28b, Cux2, Npas3, Zfp536, Sox9, Nfia, Nfib daily from D5 to D9 in independent differentiations. To assess the similarity of the two methods we used Pearson correlation. All genes recapitulated the dynamics observed in silico and had a correlation score above 0.9 with the exception of Lin28a (0.75) and Lin28b (0.80) (Figure 36). The expression pattern of the pseudotemporal reconstruction reproduces in vitro dynamics.

Manipulating TGF-β signalling is sufficient to shift the timing of the neural-to-glial transition. 48 hours of exposure to the ligand TGF-β2 between D5 and D7 induced the expression of Sox9, Nfia, Nfib, Apcdd1. Conversely, continuous inhibition of TGF-β signalling with the inhibitor SB delayed the expression of glial genes. We hypothesised that the validated candidates from the pseudotemporal reconstruction would be similarly affected. To test this, we reanalysed experiments performed in Chapter 5 where TGF-β signalling was activated by TGF-β2 between D5 and D7 or inhibited continuously by the addition of SB. At D7 and D9, the expression of Lin28a, Lin28b, Cux2, Zfp536 was measured by qPCR. Lin28a is a gene decreasing in pseudotime. Exposure to TGF-β2 decreased Lin28a expression by 34% at D7 and by 54% at D9. Inhibition with SB increased Lin28a expression by 82% at D7 and 318% at D9. A related gene, Lin28b, also decreases along pseudotime and returned a similar trend. Treatment with TGF-β2 resulted in a 30% decrease in Lin28b expression at D7 and a 49% decrease at D9. Continuous exposure to SB increased expression of Lin28b by 36% at D7 and by 103% at D9. Cux2 is a gene induced along pseudotime. Activating TGF-β signalling between D5 and D7 with TGF-β2 increased Cux2 expression by 35% at D7 and 26% at D9. In contrast inhibition with SB resulted in a 60% reduction at D7. The trend at D9 was unclear. Zfp536 is also induced along pseudotime. Exposure to TGF-β2 caused a
21% increase in Zfp536. At D9 expression returned to control levels. Continuous treatment with SB reduced Zfp536 by 33% at D7 and 18% at D9. Activating or inhibiting TGF-β signalling manipulates the expression of Lin28a, Lin28b, Cux2, Zfp536. The trends are consistent with the shift in timing of the neural-to-glial transition described in Chapter 5. These data further validate the pseudotemporal reconstruction.

The regulatory network controlling the neural-to-glial transition is not clear. To identify candidate genes, we performed single cell sequencing on in vitro derived neural progenitors and reconstructed the transition in silico. The expression dynamics in pseudotime had a strong correlation with in vitro expression. This included the dynamics of known glial genes. Manipulating TGF-β signalling shifts the timing of the neural-to-glial transition. In these experiments, the expression of validated candidates displayed a trend consistent with the temporal shift. Together these data suggest that the pseudotemporal reconstruction predicts genes involved in the neural-to-glial transition network. A complication is the opposing expression patterns of Irx3 and Msx1. Irx3 marks progenitors of the intermediate neural tube whilst Msx1 demarcates neural progenitors located dorsally. This suggests that alongside a neural-to-glial transition, iNPs are progressively dorsalising. Whether this is a consequence of withdrawing retinoic acid at D5 remains to be investigated. Nevertheless, deconvolving the two processes will require careful validation.
Chapter 6. Results

Figure 28 Validation of the samples used for single cell sequencing

*Nfib* expression in iNPs in experiments used for single cell sequencing measured by qPCR. The reference *Nfib* dynamic (grey) represents the average of all historical data. Differentiations are named for the sample they contributed to the single cell sequencing. Expression is normalised to *Actin*.
Figure 29 Removal of neurons and mesoderm cells from the data set

A heatmap of all cells clustered using neuron and mesoderm gene modules. 5 aggregated gene modules, picked based on \textit{Tubb3, Dcx, Lhx1, Lhx5, Lbx1}, represented neurons (magenta). 3 aggregated gene modules, picked based on \textit{Meox1, Foxc2, Twist1}, represented mesodermal cells (blue). The remaining cells (grey) were kept for further analysis. Each row indicates the standardised level of gene expression (z-score of log-level). Orange represents high expression, purple represents low expression. Each column is a cell. This plot was generated by Julien Delile.
Figure 30 Gene modules correlated with sample time

Heatmap of neural progenitor cells using gene modules correlated with sample time (score > 0.6). Column (cells) order is defined by hierarchical clustering. Each row indicates the standardised level of gene expression (z-score of log-level). Orange equals high expression, blue represents low expression. This plot was generated by Julien Delile.
Figure 31 Principle Component Analysis (PCA) using time gene modules captures developmental time

PCA of single cells using the averaged expression of each gene module that had a correlation score > 0.6 (A) A plot of PC1 and PC2. The percentage variance captured by each PC is shown on the axis labels. Cells are coloured by sample collection day. (B) Plot of the gene module loadings on PC1 and PC2, showing how strongly each gene module influences the principle components. Note: colours are consistent between this figure and Figure 30.
Figure 32 Smoothing single cell expression data using LOESS

(A) Histogram of cell counts in 30 equal bins across PC1. The red bar indicates an outlier separated from the remaining cells on PC1. (B) A plot of the PCA with the outlier cell removed. Box plots show the distribution of samples across PC1. (C) Expression of Sox9, Nfia, Nfib across PC1. Lines show the smoothed expression using different span values for LOESS. The red line shows the final value (0.7) used for further analysis.
Figure 33 Identifying dynamically expressed genes

The application of two selection thresholds to find genes dynamically expressed. Each dot represents a gene. (A) The mean expression (log value) of each gene across pseudotime. The threshold (red line) is set at 1.4. (B) Index of dispersion for all genes across pseudotime. The threshold (red line) is set at 1.5 standard deviations away from the mean.
Figure 34 Genes with decreasing expression across pseudotime

Smoothed expression of the selected genes across pseudotime. Clusters with decreasing expression have been selected. The expression of each gene has been normalised to the maximum. Genes highlighted in red are interesting candidates referred to in the main text.
Figure 35 Genes with increasing expression across pseudotime

Smoothed expression of genes across pseudotime. Clusters of genes increasing in expression have been selected. Data has been normalised to the maximum. Genes highlighted in red are interesting candidates referred to in the main text.
Figure 36 Pseudotemporal reconstruction reproduces *in vitro* dynamics

(A) A comparison between the expression pattern of a selection of candidate genes through pseudotime and across the *in vitro* differentiation. The score is computed by correlating (Pearson correlation) gene expression *in vitro* with the expression in pseudotime. To do so, the expression in pseudotime was averaged across 5 sequential bins. (B) The effect of manipulating TGF-β signalling on *Lin28a*, *Lin28b*, *Cux2*, *Zfp536* expression. Cells were either exposed to TGF-β for 48 hours or continuously treated with SB from D5. Values are normalised to untreated cells. Error bars represent standard error of the mean.
Chapter 7. The effect of NOTCH signalling on the neural-to-glial transition

Single cell sequencing can be used to identify genes involved in developmental decisions. However, analyses tend to return long lists of genes. A challenge is reducing the list of candidates to a testable number. Knowledge of alternate model systems or developmental lineages can be used to prioritise candidates, but this can leave potentially interesting genes unstudied. To move towards a more high-throughput approach, we modified a lentiviral system built on the gateway cloning technology (Katzen 2007) that allows rapid cloning and functional assays. We used this system to test the effect of NOTCH signalling on the neural-to-glial transition.

7.1 A lentiviral reporter system to test candidates

Gateway cloning uses recombination technology to shuttle fragments of DNA between vectors with high efficiency (Katzen 2007). Genes of interest are cloned into small, easy to use entry vectors before being transferred into a destination vector of choice. pLEX_307 is a lentiviral transfer plasmid compatible with Gateway cloning. Lentivirus’ are used in screening approaches to insert DNA into target cells (Škalamera et al. 2011). The multiplicity of infection is dependent on viral titre and cell type but is rarely 100%. Fluorescent proteins can be used to sort populations by Fluorescent Activated Cell Sorting (FACS). We therefore set out to use a lentiviral approach by engineering the pLEX_307 vector to contain either mKate2 or Venus downstream of an Internal Ribosomal Entry Site (IRES) (Jang et al. 1988, Pelletier and Sonenberg 1988). An advantage of this configuration is that the fluorescent protein and the gene of interest are translated in cis.

To test whether the modified lentiviral vectors remained functional, we differentiated ES cells for 2 days in N2B27 and infected cells with lentivirus containing either Venus or mKate2. Analysis was performed 48 hours later by flow cytometry. Cells without virus had no fluorescent protein expression. In cells infected with virus containing the Venus reporter, 5.88% were positive for Venus activity. In the case of cells transfected with the mKate2 virus 4.15% of cells were
positive 48h after infection (Figure 37 Development of a lentiviral reporter system). Thus pLEX-Venus and pLEX-mKate2 express the fluorescent reporters and produce functional lentivirus.

The efficiency of IRES-dependent second gene expression can range from 6 to 100% (Mizuguchi et al. 2000). The pLEX-Venus and pLEX-mKate2 constructs make functional virus but the percentage of positive cells is low. One possibility is that the low efficiency is due to low fluorescent protein expression resulting from an inefficient IRES. To test this, we produced a virus containing eGFP-IRES-mKate2 and used the ratio of eGFP and mKate2 to estimate efficiency. Cells were infected at D2 and analysed 48h later by flow cytometry. 3.51% of cells were positive for the fluorescent proteins. There was a strong correlation between eGFP and mKate2 and of the infected cells, 63% were eGFP, mKate2 double positive. The remaining 37% of infected cells were eGFP positive, mKate2 negative. None of the cells were mKate2 positive and eGFP negative (Figure 37). This suggests that the efficiency of the IRES is ~63%. The strong correlation between the intensity of eGFP and mKate2 suggests that mKate2 intensity is a good correlate of the expression level of the co-transcribed gene.

The modified lentiviral system produces functional virus and the reporter correlates with gene expression. It is unclear whether the percentage of infected cells is sufficient to test candidates from the single cell sequencing. HES1 inhibition of proneural genes, such as ASCL1, results in the maintenance of progenitors at the expense of neuron differentiation (Vasconcelos and Castro 2014). Therefore, constitutive expression of HES1 during the phase of neuron differentiation should inhibit the expression of the neuron marker Tubb3. In vitro neuron production begins at D6. To functionally validate the viral approach, we infected INPs at D4 with either a Hes1-mKate2 or empty-mKate2 virus and sorted equivalent numbers of positive and negative cells at D7. We measured the expression of Tubb3 (neurons) and Sox2 (progenitors) by qPCR. The empty virus did not induce Hes1. Hes1-mKate2 infected cells had a 7-fold increase in Hes1 expression over the uninfected (Figure 38A). In the control virus there was a 20% reduction in Sox2 expression and a 10% reduction in Tubb3 expression compared to the uninfected population. Cells infected with Hes1-mKate2 virus had a 40% decrease in Sox2
expression and an 84% reduction in Tubb3 expression compared to the uninfected population. The expression of the housekeeping gene Gapdh was unchanged between the infected and uninfected populations (Figure 38B). In agreement with previous work, constitutive expression of HES1 inhibits neuron differentiation (Vasconcelos and Castro 2014). The reduction of Sox2 and Tubb3 expression in the control suggests the lentivirus may preferentially infect non-neural populations. Nevertheless, these data confirm the lentiviral approach can be used to test candidates from the single cell sequencing analysis.

7.2 Activation of NOTCH signalling induces glial genes

Single cell sequencing followed by pseudotemporal reconstruction returned a list of 176 genes with a dynamic of expression correlating with the neural-to-glial transition. One cluster representing genes increasing with time contains Sox9, a gene implicated in the regulation of NFIA and NFIB (Kang et al. 2012). Notch1, encoding a receptor of the NOTCH signalling pathway, is represented in the same cluster (Figure 39A). The NOTCH pathway has been proposed to regulate the expression of SOX9 and NFIA (Namihira et al. 2009, Martini et al. 2013). We decided to use lentiviral transfection to investigate the role of NOTCH signalling in the neural-to-glial transition in iNPs.

NOTCH receptor activation results in cleavage of the intracellular domain (ICD), allowing it to translocate to the nucleus and activate transcription (Louvi and Artavanis-Tsakonas 2006). To investigate the role of NOTCH1 in the activation of the glial lineage, we generated a lentivirus containing NOTCH1 ICD (NICD). To validate the NICD virus, we measured the induction of the NOTCH signalling target Hes1. Cells were infected at D4 of the differentiation, at the time of neural commitment. At D7, prior to NFIA and NFIB induction, infected and uninfected cells were sorted and the populations assessed for Hes1 expression by qPCR. Cells infected with NICD virus showed a 52% increase in Hes1 expression (Figure 39B). The activation of Hes1 by overexpressing NICD suggests the virus is functional. NOTCH signalling has a well described role in controlling the balance between progenitor maintenance and neuron differentiation (Louvi and Artavanis-Tsakonas...
Activation of NOTCH signalling blocks proneural gene expression and thus neuron differentiation. To further validate the NICD virus, we measured Sox2 (progenitors) and Tubb3 (neurons) expression by qPCR. The experiment was performed as described above. Overexpression of NICD caused a 31% increase in Sox2 expression and a 75% reduction in Tubb3 expression (Figure 39B). Thus, activation of NOTCH signalling reduces neuron differentiation. It is unclear whether the increase in Sox2 is a consequence of progenitor proliferation or NOTCH signalling inducing its expression.

The activation of the glial lineage involves the induction of NFIA and NFIB, which in iNPs begins at D8 (Chapter 4, (Deneen et al. 2006)). NOTCH signalling has previously been reported to induce both SOX9 and NFIA, but its role in the activation of the glial lineage is unclear (Namihira et al. 2009, Martini et al. 2013). To investigate this, we performed qPCR for glial genes in cells with and without NICD overexpression from day 4 to D7. At D7, iNPs exposed to NICD overexpression had a 130% increase in Sox9 compared to the uninfected control. Nfia and Nfib expression were increased by 93% and 145% respectively, whilst the SOX9/NFIA target, Apcdd1, increased 84% (Figure 39C). Activation of NOTCH signalling is sufficient to induce glial gene expression. These data suggest that activation of the glial lineage can be uncoupled from neuron differentiation.

Although HES1 has a well described role in neuron differentiation, its ability to regulate the neural-to-glial transition is not clear (Louvi and Artavanis-Tsakonas 2006). To understand whether HES1 could mimic the effects of NICD, we measured the expression of glial genes in our HES1 overexpression experiments by qPCR. In cells infected with Hes1-mKate2 virus, we observed a small reduction in Sox9 expression, a 40% increase in Nfia and 51% increase in Nfib. The SOX9/NFIA target, Apcdd1, was increased by 72% (Figure 39E). Constitutive HES1 expression did not affect Sox9 expression, but did induce Nfia, Nfib, Acdd1. Therefore, HES1 does not completely account for the level of glial gene induction by NICD. This suggests that either NICD is directly affecting the neural-to-glial transition or that there are alternative targets involved.
Single cell sequencing analysis returned a list of 26 transcription factors dynamically expressed across the transition. Since we predicted that alternative NOTCH targets were involved, we mined the transcription factor list for potential candidates. Cux2 increases in expression across pseudotime, a pattern confirmed qPCR in vitro (correlation score = 0.9). Activation of NICD in the developing spinal cord, induced Cux2 expression (Iulianella et al. 2009). To test if Cux2 was regulated by NOTCH signalling in iNPs, we measured its expression by qPCR in our NICD overexpression experiments. In cells infected with NICD, Cux2 expression was 192% higher than in the uninfected population. In contrast, two other candidates from the single cell sequencing, Zfp536 and Npas3, were unchanged (Figure 39F). In agreement with in vivo experiments activation of NOTCH signalling is sufficient to induce Cux2 expression (Iulianella et al. 2009). The role of Cux2 in the neural-to-glial transition remains to be tested.

Activation of the glial lineage is characterised by the induction of NFIA and NFIB. Therefore, understanding the timing of the neural-to-glial transition relies on defining the mechanism of NFIA and NFIB induction. In this chapter we report the development of a method for testing candidates that combines the Gateway technology, lentivirus, and fluorescent reporters. NOTCH signalling has been proposed to regulate both SOX9 and NFIA (Namihira et al. 2009, Martini et al. 2013). Notch1 encodes a NOTCH signalling receptor and is induced across pseudotime. We use the lentiviral approach to investigate the role of NOTCH signalling in the neural-to-glial transition by overexpressing a constitutively active form of the receptor (NICD). This results in the induction of glial genes Sox9, Nfia, Nfib, and Apcdd1. Overexpression of a NOTCH signalling target HES1 did not reproduce the induction of glial genes. We mined the single cell dataset for known NOTCH targets and identified Cux2. In agreement with previous work, activation of NOTCH signalling was sufficient to induce Cux2 expression although it remains unclear whether CUX2 has any involvement in the regulation of glial genes (Iulianella et al. 2009).
Figure 37 Development of a lentiviral reporter system

Density plots of flow cytometry data from cells differentiated to D4. Cells were infected with either no virus (Negative, top-left), Venus (top-right), mKate2 (bottom-left) or eGFP-mKate2 (bottom-right) virus at D2. The numbers in the plots indicate the percentage cells in each quadrant.
Figure 38 Constitutive expression of HES1 blocks neuron differentiation

qPCR on D7 cells infected with Hes1-mKate2 or empty-mKate2 virus at D4. Equal numbers of infected (+, solid fill) and uninfected (-, no fill) cells were sorted by FACS. (A) Hes1 expression in infected cells and uninfected cells. (B) Expression of Tubb3, Sox2, Gapdh in infected and uninfected cells. These data are normalised to the average expression of uninfected cells. * p = 0.0286, Mann-Whitney unpaired t test. (A,B) Error bars represent standard error of the mean.
Figure 39 Activation of NOTCH signalling induces glial genes

(A) Expression dynamics of the cluster containing Sox9 and Notch1 from the pseudotemporal reconstruction. The expression of each gene is normalised to the maximum. White is low, black is high expression. (B-F) Gene expression data measured by qPCR in iNPs infected with virus at D4 and equal numbers of cells sorted by FACS at D7. The data is normalised to the uninfected population. Error bars represent standard error of the mean. (B,C,F) iNPs infected with NICD-mKate2 virus. (D,E) iNPs infected with Hes1-mKate2 virus.
Chapter 8. ZFP536 affects the neural-to-glial transition via SOX9

*In vivo* neuron differentiation precedes gliogenesis. The onset of neuron differentiation is at E9 whereas the induction of the early glial markers NFIA and NFIB is at E11.5 (Deneen et al. 2006, Kicheva et al. 2014). The induction of SOX9 is coincident with neuron differentiation (Stolt et al. 2003). ZFP536 is a zinc finger protein and a putative transcription factor (Qin et al. 2009). The combination of ZFP536, SOX10 and OLIG2 is sufficient to reprogram fibroblasts into oligodendrocytes, a subtype of glia (Yang et al. 2013). Additional work has found it to negatively regulate neuron differentiation (Qin et al. 2009). Despite this work, little is known about its expression in vivo or its role in the onset of gliogenesis. We decided to investigate the function of ZFP536 in the neural-to-glial transition.

8.1 ZFP536 is sufficient to activate glial gene expression

To document the expression pattern of *Zfp536* across the neural-to-glial transition in the developing spinal cord, we collected embryos from E9.5 to E12.5 and performed *in situ* hybridisation on transverse sections of the brachial neural tube. No *Zfp536* transcript was detected at E9.5. At E10.5 low expression was seen in the ventricular zone. Expression within the ventricular zone increased further at E11.5. Expression was still present at E12.5, but had decreased (Figure 40). The dynamic of *Zfp536* expression in vivo appeared similar to the reported expression pattern of SOX9 (Stolt et al. 2003). To confirm this, Manuela Melchionda performed SOX9 immunofluorescence across the same time frame. SOX9 emerged ventrally at E9.5, and by E10.5 was expressed in progenitors across the dorsal-ventral axis. The expression of SOX9 was maintained in progenitors until E12.5. These data are consistent with the reported SOX9 expression pattern (Stolt et al. 2003). SOX9 expression appears to overlap with *Zfp536* in the developing spinal cord.

Neuronal differentiation in vivo begins at E9 and in vitro at D6. *Zfp536* expression begins at E10.5. Previous work has shown that constitutive expression of ZFP536 can inhibit RA-induced neuronal differentiation of P19 cells (Qin et al. 2009). To
assess the function of ZFP536 in neuron differentiation, we turned to the lentiviral reporter system and the *in vitro* differentiation protocol (Chapter 7). We infected cells at D4 at the time of neural commitment with a Zfp536-mKate2 virus. At D7, after the onset of neuron differentiation, we sorted the infected and uninfected populations. Expression of Sox2 (progenitors) and Tubb3 (neurons) was assessed by qPCR. Overexpression of ZFP536 resulted in a 16-fold increase in Zfp536 (Figure 41A). Sox2 increased by 87% in infected cells but Tubb3 was unchanged (Figure 41B). Constitutive expression of ZFP536 does not inhibit neuron differentiation.

Since *Zfp536* is expressed across the neural-to-glial transition in iNPs and *in vivo*, we predicted ZFP536 might affect the expression of glial genes. We examined Sox9, Nfia, Nfib and Apcdd1 expression in the altered Zfp536 expression experiment described above. Sox9 expression was increased 3.3-fold in the infected population compared to the uninfected population. *Nfia* and *Nfib* expression was increased by 150% and 83%, respectively, whilst the expression of *Apcdd1* was increased 276% (Figure 41B). Constitutive expression of ZFP536 results in the induction of Sox9 and glial genes. In D7 iNPs, SOX9 is already expressed (Chapter 3), suggesting the effect of ZFP536 on Sox9 expression might be functionally significant. We wanted to confirm this result *in vivo*. Accessibility to the embryo is restricted in the mouse making overexpression experiments challenging. The chick represents a more tractable system to test this. The same ZFP536-mKate2 construct used to make the virus was electroporated unilaterally into the neural tube with the other side acting as an internal control. The embryos were then allowed to develop for 24 hours and SOX9 expression assessed by immunostaining. The electroporated side showed expression of the fluorescent reporter (mKate2 (RFP)). Cells positive for the reporter also showed immunoreactivity for the FLAG-tag attached to the C-terminus of ZFP536. There was a strong correlation between FLAG expression and mKate2. SOX9 staining was observed in a ventral to dorsal gradient on the control side whilst on the electroporated side, ZFP536-FLAG positive cells had increased SOX9 expression. The level of FLAG expression correlated with the intensity of SOX9 (Figure 41C). Constitutive expression of ZFP536 is therefore sufficient to induce SOX9 expression both in iNPs and *in vivo*. Whether the effect on *Nfia, Nfib, Apcdd1* in
iNPs is a consequence of increased SOX9 expression or due to direct regulation by ZFP536 remains to be tested.

Activation of NOTCH signalling is sufficient to activate SOX9 and NFIA expression (Namihira et al. 2009, Martini et al. 2013). In Chapter 7, we demonstrate that NOTCH signalling regulates glial genes in iNPs. Since activation of NOTCH did not induce the expression of Zfp536, we asked whether ZFP536 might be upstream of NOTCH signalling. To test this, we infected cells with ZFP536-mKate2 lentivirus at D4, shortly after neural commitment. At D7 infected and uninfected populations were sorted and the expression of the validated NOTCH targets, Hes1 and Cux2, measured by qPCR. Cux2 expression was induced approximately 4.5-fold over the uninfected population. Hes1 expression was induced 1.7-fold (Figure 41B). Thus, constitutive expression of ZFP536 is sufficient to induce targets of NOTCH signalling.

ZFP536 is a putative transcription factor with dynamic expression across the neural-to-glial transition in vitro. We characterised the expression pattern of Zfp536 in the developing spinal cord, and show it is dynamically expressed across the transition in vivo. Constitutive expression of ZFP536 in vitro was sufficient to induce glial genes as well as Sox9. Although the progenitor marker Sox2 is increased, the neuronal marker Tubb3 is unaffected. This contradicts previous work showing ZFP536 negatively regulates neuron differentiation (Qin et al. 2009). We show that ZFP536 is sufficient to induce SOX9 in vivo. Together these data suggest a role for ZFP536 in the neural-to-glial transition.

### 8.2 ZFP536 is required for the correct timing of the neural-to-glial transition

To test whether removal of ZFP536 would affect SOX9 expression and consequently the transition to gliogenesis, we deleted ZFP536 in ES cells using CRISPR/Cas9. Zfp536 comprises 3 exons, the first and third contributing the majority of the coding sequence. We designed guides to the 5’ region of the third exon. Sequencing identified a clone homozygous for a 10 base pair deletion.
In vivo neuron differentiation begins at E9 and in vitro at D6 (Kicheva et al. 2014). The induction of SOX9 is concurrent with neuron differentiation whilst Zfp536 expression closely follows SOX9 (Stolt et al. 2003). However, deletion of SOX9 did not affect the onset of neuron differentiation Chapter 4. Constitutive expression of ZFP536 was not sufficient to affect the neuron marker Tubb3. We test whether removal of ZFP536 affected neuronal differentiation we measured TUBB3 expression, in WT and ZFP536(i) iNPs every 48 hours from D5 to D11 by qPCR and flow cytometry. At D5 low Tubb3 expression was detected in WT cells. Expression increased from D5 to D7, and again at D9. Expression remained constant to D11. ZFP536(i) cells followed WT dynamics (Figure 42A). We confirmed this result by measuring the fraction of TUBB3-positive, SOX2-negative cells by flow cytometry. ZFP536(i) cells followed WT dynamics (Figure 42B). Deletion of ZFP536 does not affect the onset of neuron differentiation.

SOX9 expression in vivo begins at E9.5, shortly after the onset of neuron differentiation, a schedule recapitulated in vitro. ZFP536 is sufficient to induce SOX9 expression in vivo and in vitro. To test whether ZFP536 is required for SOX9 expression, we collected samples every 48 hours from D5 to D11. Sox9 transcription was measured in WT and ZFP536(i) cultures by qPCR. In WT cells, Sox9 expression increased from D5 to D9, before decreasing at D11. In ZFP536(i) at D5 Sox9 expression was 65% less, at D7 45% less, at D9 23% less. Sox9 expression recovered to WT levels at D11 (Figure 44A). We confirmed this result by immunofluorescence. At D5, no SOX9 was detected in WT or ZFP536(i) cells. As described previously, by D7 SOX9 expression was observed in WT cells. SOX9 expression was maintained to D9. In ZFP536(i) SOX9 expression was reduced at D7 but had recovered at D9 (Figure 43). Therefore, deletion of ZFP536 causes a delay in SOX9 induction.

NFIA and NFIB mark the onset of the glial transition and in vivo their expression begins at E11.5 and in vitro at D8 (Chapter 3, (Deneen et al. 2006)). Deletion of SOX9 results in a delay in NFIA and NFIB induction (Chapter 4, (Kang et al. 2012)). Deletion of ZFP536 causes a delay in SOX9 induction but to test whether this is sufficient to affect NFIA and NFIB induction we measured their expression by
Chapter 8. Results

qPCR. In WT cells, little to no expression was detected at D5. Expression increased slightly at D7 but increased dramatically at D9 and D11. In ZFP536\(^{(c)}\) cells, \(Nfia\) expression was slightly reduced by 17% at D9. At D11 expression had returned to WT levels. \(Nfib\) expression was reduced by 30% at D9 and by 26% at D11 (Figure 44B). To more accurately assay the \(Nfib\) phenotype, we measured the percentage of NFIB-positive cells in SOX2-positive neural progenitors by flow cytometry. There were no NFIB-positive neural progenitors at D7. In WT cells, 26% of neural progenitors were NFIB-positive at D9, and at D11, this had risen to 94%. In ZFP536\(^{(c)}\) cultures, 10% of neural progenitors were positive for NFIB. This recovered to 94% at D11 (Figure 44C). Therefore, deletion of ZFP536 results in a delay in NFIB, but not \(Nfia\), induction. Whether the delay is due to direct regulation by ZFP536, or a consequence of delayed SOX9 expression is not clear.

SOX9, NFIA and NFIB co-activate a cohort of target genes, such as \(Apcdd1\), important in glial lineage activation (Kang et al. 2012). The expression of these targets is reduced in SOX9- and NFIA-deficient embryos at E12.5. To test whether the delayed SOX9, NFIB expression in ZFP536 null iNPs was sufficient to impact the induction of \(Apcdd1\) we collected RNA every 48 hours from D5 to D11 and measured \(Apcdd1\) expression by qPCR. As described previously, \(Apcdd1\) was low at D5 and D7. Transcription was induced at D9, increasing further at D11. In ZFP536\(^{(c)}\) cells expression was reduced by 31% at D9. There was no recovery at D11 (Figure 44D). Deletion of ZFP536 results in reduced \(Apcdd1\) expression. It is unclear whether ZFP536 is directly involved in glial lineage activation, or if the effect is a consequence of delayed SOX9 or NFIB expression.

Terminal glial differentiation is characterised by GFAP expression and follows glial lineage activation by SOX9/NFIA. \(In vitro\) GFAP is expressed by D15. Removal of SOX9 or NFIA/B is sufficient to block terminal differentiation (Chapter 4, (Stolt et al. 2003). We hypothesised that the delay of SOX9, NFIB, as well as the reduction in target gene activation would impact terminal differentiation. To assess this, we measured \(Gfap\) expression by qPCR at D15 in WT and knockout cells. ZFP536\(^{(c)}\) cells had a 50% reduction in \(Gfap\) expression relative to the WT. To confirm this, we measured the percentage of GFAP-positive cells by flow cytometry. Approximately 4% of cells were GFAP-positive in the WT. This was reduced by
45% in ZFP536(-) cells (Figure 44E). Deletion of ZFP536 causes reduced terminal glia differentiation. Whether this is due to the earlier functions of ZFP536 or because it has multiple roles in glial differentiation is not clear.

Removal of ZFP536 results in a delay in SOX9, NFIB expression and a reduction in Apcdd1 and GFAP expression. Typically, knockouts are confirmed using immunocytochemistry. This is difficult for ZFP536 due to the absence of antibodies. To validate the ZFP536(-) result we made an independent deletion of ZFP536 using a different approach. Guides were designed to the 5’ and 3’ regions of exon 1 in order to excise a 2kb fragment of the ZFP536 coding region. A positive clone (ZFP536(-) #2) was taken through the differentiation paradigm and a time course for glial gene expression performed as described above. Sox9 expression, measured by qPCR, was comparable to ZFP536(-) (Figure 45A). The Nfib phenotype was more severe in the ZFP536(-) #2, showing similar kinetics to the SOX9(-) cells. SOX2 expression marks progenitors throughout the neural lineage. The percentage of NFIB-positive in SOX2-positive progenitors at D9 was comparable to that of ZFP536(-). At D11 the percentage of NFIB-positive progenitors had not recovered to the extent of ZFP536(-) (Figure 45B). At day15 we assessed GFAP expression by qPCR and flow cytometry. In both instances the phenotype was comparable to that of ZFP536(-) (Figure 45C). Together these data show that an independent deletion of ZFP536 has a comparable effect on the induction of glial genes, validating the requirement for ZFP536 in glial differentiation. The NFIB phenotype is more severe in ZFP536(-) #2. This could be a consequence of variable differentiation efficiencies between clones.

ZFP536 is sufficient to induce SOX9 expression in vivo. It is also required for the proper timing of SOX9 expression in vitro. To test whether this regulatory relationship is reciprocated, we differentiated WT, SOX9(-), ZFP536(-) cells and assayed their ability to induce ZFP536. Samples were collected every 48 hours from D5 to D11 and Zfp536 expression measured by qPCR. In WT cells, expression was low at D5 and increased to D11. At D5, D7, D9 Zfp536 expression in SOX9(-) did not change from WT levels. At D11, Zfp536 expression was reduced by 29%. Zfp536 expression was reduced at all time points in ZFP536(-) cells (Figure 46). Thus, SOX9 does not regulate the expression of Zfp536. A small reduction is
observed at D11 after glial lineage activation. This might suggest that at later time points SOX9 regulates Zfp536. Whether this is in combination with NFIA similar to how other glial genes are regulated at this time is unclear.

ZFP536 is sufficient to induce SOX9 and glial gene expression. In this section we delete ZFP536 and show that it is required for the correct schedule of glial gene expression. ZFP536\(^{(c)}\) iNPs had a delay in SOX9 and NFIB as well as reduced expression of Apccd1. At D15 there was also a reduction in the percentage of GFAP-positive glia. It is not clear whether these phenotypes are a result of delayed SOX9 expression or ZFP536 having multiple roles throughout glial differentiation. However, ZFP536 is not regulated by SOX9 as Zfp536 expression is not affected in the SOX9\(^{(c)}\). Removal of ZFP536 does not affect the percentage of neurons generated *in vitro*. This data moves towards a model where the timing of the neural-to-glial transition is uncoupled from the onset of neuron differentiation.

8.3 ZFP536 affects the neural-to-glial transition via SOX9

NFIA and NFIB mark the onset of the glial lineage. Deletion of SOX9 delays NFIA and NFIB induction, and abrogates terminal gliogenesis (Stolt *et al.* 2003, Deneen *et al.* 2006). *In vivo*, SOX9 expression begins at E9.5 and Zfp536 at E10.5. Constitutive expression of ZFP536 induces Sox9 and glial genes whilst ZFP536 null iNPs had a delay in SOX9, NFIB induction and reduced gliogenesis. It is not clear whether the later phenotypes are a consequence of delayed SOX9 expression. To define the epistasis between SOX9 and ZFP536, we used CRISPR/Cas9 to delete SOX9 in ZFP536\(^{(c)}\) cells. We targeted Sox9 with the same gRNA used to make the SOX9\(^{(c)}\) cell line (Chapter 4). Sequencing identified a homozygous clone with a 29bp deletion.

Neuron differentiation begins at E9 *in vivo* and *in vitro* at D6. SOX9 and ZFP536 expression follows neuron differentiation. Neither removal of SOX9 nor ZFP536 affects the onset of neuron differentiation. We predicted the double knockout would also have no neurogenic phenotype. To assess neuron differentiation, we measured *Tubb3* expression by qPCR in WT and SOX9/ZFP536\(^{(c)}\). As described
previously, in the WT Tubb3 expression was low at D5 and increased to D11. The dynamics of Tubb3 expression in SOX9/ZFP536(-/-) cells was comparable to WT cells. We confirmed this result by measuring the proportion of TUBB3 positive, SOX2 negative cells (neurons) using flow cytometry (Figure 47). The onset of neuron differentiation is therefore unaffected upon combined deletion of SOX9 and ZFP536.

NFIA and NFIB are molecular correlates of the glial transition. *In vitro* their expression begins at D8/9. Deletion of SOX9 results in a delay in NFIA and NFIB induction (Chapter 4, (Kang et al. 2012)), whilst the removal of ZFP536 causes a delay in NFIB expression. It is unclear whether the effect of ZFP536 on NFIB is a consequence of delayed SOX9 expression. To test this, we measured Nfib expression in WT, SOX9(-/-), ZFP536(-/-), SOX9/ZFP536(-/-) iNPs across the neural-to-glial transition. Samples were collected every 48 hours from D5 to D11 and Nfib measured by qPCR. As seen previously, SOX9(-/-) and ZFP536(-/-) had reduced Nfib expression compared to the WT. Nfib levels in the SOX9/ZFP536(-/-) at D9 were comparable to the SOX9(-/-). At D11 there was a recovery of Nfib expression to ZFP536(-/-) levels (Figure 48A). To confirm this result, we measured the proportion of NFIB positive in the SOX2 positive population (progenitors) by flow cytometry. 26% of cells were positive for NFIB in WT cells, 10% were positive in ZFP536(-/-) and 3% were positive in SOX9(-/-) cells at D9. In SOX9/ZFP536(-/-) cells at D9, 4% of neural progenitors were positive. At D11 the percentage of NFIB-positive progenitors recovered to WT levels. NFIB induction upon combined deletion of ZFP536 and SOX9 is comparable to that seen in SOX9(-/-) cells (Figure 48B). Therefore, the delay in NFIB induction in ZFP536(-/-) cells is a consequence of delayed SOX9 expression. At D11, there is a greater recovery of NFIB expression in the SOX9/ZFP536(-/-) compared to the SOX9(-/-). What is compensating for SOX9/ZFP536 at later stages is not clear.

The function of SOX9, NFIA and NFIB is to co-activate a group of genes important in glial lineage activation (Chapter 4, (Kang et al. 2012)). One target, Apcdd1, is induced *in vitro* at D9, following NFIA and NFIB induction. The expression of Apcdd1 is reduced upon deletion of SOX9 and ZFP536, although the SOX9(-/-) phenotype is more severe. To test whether ZFP536 functions in glial lineage
activation, we analysed *Apcdd1* expression by qPCR in WT, SOX9\(^{(\cdot)}\), ZFP536\(^{(\cdot)}\), SOX9/ZFP536\(^{(\cdot/\cdot)}\) cells. *Apcdd1* expression in SOX9\(^{(\cdot)}\) and ZFP536\(^{(\cdot)}\) was reduced by 71% and 31% respectively at D9. Although expression increased at D11, the level was less than WT cells. In the SOX9/ZFP536\(^{(\cdot/\cdot)}\) *Apcdd1* expression showed a 65% reduction (Figure 48C). The dynamic of *Apcdd1* in SOX9/ZFP536\(^{(\cdot/\cdot)}\) cells was comparable to that in the SOX9\(^{(\cdot)}\) cells. The lack of any further reduction suggests that ZFP536 promotes glial lineage activation via SOX9.

Following SOX9/NFIA mediated lineage activation cells migrate away from the progenitor zone and terminally differentiate. Terminal differentiation is characterised by the onset of GFAP expression. *In vitro*, GFAP is expressed at D15. Deletion of SOX9 severely reduces the number of GFAP-positive glia whilst removal of ZFP536 causes a ~50% reduction. It is unclear whether the glial phenotype in ZFP536\(^{(\cdot)}\) cells is due to a direct function of ZFP536 or a consequence of earlier events. To assess this, we measured *Gfap* expression in SOX9/ZFP536\(^{(\cdot/\cdot)}\) cells at D15. As discussed previously deletion of SOX9 results in a 96% decrease in *Gfap* expression. Removal of ZFP536\(^{(\cdot)}\) results in a 51% decrease. SOX9/ZFP536\(^{(\cdot/\cdot)}\) cells show a reduction of 93% in *Gfap* expression (Figure 48D). Combinatorial deletion of SOX9 and ZFP536 therefore did not further reduce terminal differentiation with respect to the SOX9\(^{(\cdot)}\) cultures. This suggests the reduced gliogenesis in ZFP536\(^{(\cdot)}\) cells is dependent on SOX9.

ZFP536 is required for the correct timing of the neural-to-glial transition. Deletion of ZFP536 causes a delay in SOX9, NFIB induction and a reduction in *Apcdd1* expression *in vitro*. Knockout of SOX9 causes a delay in NFIB induction and reduction of *Apcdd1*. Although the phenotypes are similar the SOX9\(^{(\cdot)}\) is more severe. In this section, we investigated the epistasis of SOX9 and ZFP536 by deleting SOX9 in ZFP536\(^{(\cdot)}\) ES cells using CRISPR/Cas9. Combinatorial deletion of SOX9 and ZFP536 resulted in comparable *Sox9*, *Nfib*, *Apcdd1* phenotypes to the SOX9\(^{(\cdot)}\). This suggests the functions of ZFP536 in gliogenesis are dependent on SOX9. Whether ZFP536 is only regulating SOX9 expression or interacting with SOX9 to activate glial genes remains to be tested. An experiment to test this would be to delete ZFP536 after NFIB induction. Combined deletion of SOX9 and ZFP536 does not completely block the induction of NFIB or the activation of *Apcdd1*
suggesting alternative mechanisms exist. Inhibition of TGF-beta signalling delays the induction of the glial lineage (Chapter 5). Assaying glial gene induction in SOX9/ZFP536\textsuperscript{(-/-)} cells in the context of TGF-beta inhibition might block the transition completely. Additionally, activation of NOTCH signalling is sufficient to induce glial gene expression (Chapter 7, (Namihira et al. 2009, Martini et al. 2013)). This implies a separate network can control the timing of the transition although at what point they converge is not clear. The advantage of this configuration might be to impart temporal robustness on the neural-to-glial transition. Nevertheless, the ZFP536\textsuperscript{(-/-)} phenotype provides evidence that the pseudotemporal reconstruction correctly predicts genes involved in the neural-to-glial transition. Preliminary data suggest that deletion of another candidate, NPAS3, also has a gliogenic phenotype.
Figure 40 Zfp536 and SOX9 expression in vivo

The expression pattern of Zfp536 and SOX9 in brachial neural tubes of E9.5, E10.5, E11.5 and E12.5 mice. Images were taken at 10x magnification (A) Zfp536 expression detected by in situ hybridisation. (B) SOX9 expression detected by immunofluorescence, performed by Manuela Melchionda. Note: the SOX9 staining is not in the same embryos as the Zfp536 in situ.
Figure 41 Constitutive expression of ZFP536 in vitro and in vivo

(A,B) iNPs were infected with Zfp536-mKate2 virus at D4 and equal numbers of cells sorted by FACS at D7. Gene expression was measured by qPCR. Error bars represent standard error of the mean. (A) Zfp536 expression in infected cells (+) and uninfected cells (−). (B) Gene expression data is normalised to the uninfected population. (C) Electroporation of Zfp536-FLAG-IRES-NLS-mKate2 (the same vector used to make the virus) into chick embryos and incubated for 24 hours. Immunofluorescence images for DAPI, SOX9, RFP (mKate2), and FLAG (tag attached to ZFP536). Images were taken at 20x magnification. Scale bars = 40µM. Chick electroporations were done with the help of Despina Stamatakis and Katherine Exelby.
Figure 42 Neuron differentiation is unaffected by deletion of ZFP536

TUBB3 expression in WT, SOX9\(^{(-)}\), ZFP536\(^{(-)}\) iNPs at D5, D7, D9, D11 (n = 7-11). Error bars represent standard error of the mean. (A) Tubb3 expression measured by qPCR. (B) Percentage of TUBB3 positive, SOX2 negative (neurons) measured by flow cytometry.
Figure 43 SOX9 expression is delayed in ZFP536<sup>(↑)</sup>

Immunofluorescence images for DAPI, SOX9, SOX2, TUBB3 in WT, SOX9<sup>(↑)</sup>, ZFP536<sup>(↑)</sup> iNPs at D5, D7, D9. Images were taken at 20x magnification. Scale bars = 40µM.
Figure 44 Deletion of ZFP536 reduces glial gene expression

Expression of glial genes in WT, SOX9\(^{(-)}\), ZFP536\(^{(-)}\) iNPs (\(n = 7\)-12). Error bars represent standard error of the mean. (A-D) Samples were collected at D5, D7, D9, D11. (A,B,D) Gene expression measured by qPCR. (C) Percentage of NFIB positive progenitors (gated on the progenitor marker SOX2) measured by flow cytometry. (E) GFAP expression at day 15 measured by qPCR (left) and flow cytometry (right).
Figure 45 Glial phenotypes are recapitulated in an alternate deletion of ZFP536

Expression of glial genes in WT, SOX9\(^{(c)}\), ZFP536\(^{(c)}\), ZFP536\(^{(c)}\) #2 (alternative strategy - deletion in exon 1) iNPs (n = 3-12). Error bars represent standard error of the mean. (A,B) Samples were collected at D5, D7, D9, D11. (A,B) Gene expression measured by qPCR. (B(right)) Percentage of NFIB positive progenitors (gated on the progenitor marker SOX2) measured by flow cytometry. (E) GFAP expression at day 15 measured by qPCR (left) and flow cytometry (right).
Chapter 8. Results

Figure 46 Zfp536 induction is unaffected by deletion of SOX9
Zfp536 expression measured by qPCR in WT, SOX9\(^{-}\), ZFP536\(^{-}\) iNPs at D5, D7, D9, D11 (n = 6-9). Error bars represent standard error of the mean.

Figure 47 Combined deletion of SOX9/ZFP536 does not affect the onset of neuron differentiation
TUBB3 expression in WT, SOX9\(^{-}\), ZFP536\(^{-}\), SOX9/ZFP536\(^{-/-}\) iNPs at D5, D7, D9, D11 (n = 3-11). Error bars represent standard error of the mean. (Left) Tubb3 expression measured by qPCR. (Right) Percentage of TUBB3 positive, SOX2 negative cells (neurons) measured by flow cytometry.
Figure 48 No additional reduction of glial genes upon combined deletion of ZFP536 and SOX9

Expression of glial genes in WT, SOX9\(^{(-)}\), ZFP536\(^{(-)}\), SOX9/ZFP536\(^{(-/-)}\) iNPs (n = 3-12). Error bars represent standard error of the mean. (A,C) Samples were collected at D5, D7, D9, D11. (B) Samples were collected at D7, D9, D11. (D) Samples were collected at D15. (A,C,D) Gene expression measured by qPCR. (B) Percentage of NFIB positive progenitors (gated on the progenitor marker SOX2) measured by flow cytometry.
Chapter 9. Discussion

In the developing spinal cord, neural progenitors sequentially generate neurons and glia. \textit{In vivo} neuron differentiation begins at E9 whilst the earliest glial makers, NFIA and NFIB, are expressed at E11.5 (Kessaris \textit{et al.} 2001, Rowitch 2004). Although various mechanisms have been described, molecular detail of the gene regulatory network underlying the neural-glial transition is limited. In this study we attempt to address this as well as building on existing understanding. We began by developing a differentiation protocol which converts ES cells to gliogenic progenitors. Progenitors recapitulated all the transcriptomic features of the neural-to-glial transition observed \textit{in vivo}. Using genetic and biochemical manipulation of known genes and signalling pathways we show that the sequence of steps described \textit{in vivo} is recapitulated \textit{in vitro}. We utilised single cell sequencing in combination with the \textit{in vitro} model system to reconstruct the developmental trajectory of transitioning progenitors \textit{in silico}. Plotting transcriptional dynamics through pseudotime produced a molecular description of the transition. This led to the identification of ZFP536, a putative transcription factor previously implicated in gliogenesis. ZFP536 is expressed at the relevant time points \textit{in vivo} and manipulating its expression changes the timing of the neural-to-glial transition. In this section we discuss the major findings from our work.

9.1 An \textit{in vitro} model of the neural-to-glial transition

\textit{In vivo} neurons are produced by progenitors from approximately E9 to E12.5 after which they generate glia (Rowitch 2004). NFIA and NFIB are molecular correlates of glial lineage onset. \textit{In vivo} their expression begins at E11.5 and continues throughout gliogenesis (Deneen \textit{et al.} 2006). SOX9 is required for the correct timing of NFIA and NFIB induction and is also expressed throughout gliogenesis (Kang \textit{et al.} 2012). SOX9 expression is coincident with the onset of neuron differentiation beginning at E9.5 (Stolt \textit{et al.} 2003). SOX9 and NFIA collaborate to activate downstream target genes important in glial differentiation. One example is APCDD1 and its expression begins at E12.5 following NFIA and NFIB induction (Kang \textit{et al.} 2012). NFIB has also been shown to activate a related transcription
factor, NFIX. In the developing spinal cord NFIX expression begins at E12.5 (Matuzelski et al. 2017). In summary, a series of transcriptional changes occur with a defined temporal schedule.

Due to their amenability to biochemical and genetic manipulations in vitro models have been used to investigate biological mechanisms. The discovery of conditions to indefinitely expand embryonic stem cells has led to the development of differentiation protocols that recapitulate aspects of embryonic development. A recently described method induces neural progenitors (iNPs) with a primitive spinal cord identity (Gouti et al. 2014). In order to recapitulate the neural-to-glial transition in vitro we cultured iNPs in a minimal medium and recorded the transcriptional changes of genes known to have specific temporal dynamics in vivo. Induction of neural progenitors was at D4 and neuron differentiation began at D6. SOX9 expression appeared at D6, coincident with differentiated neurons. NFIA and NFIB were induced at D8, and the downstream targets emerged at D9. In short, the in vitro system recapitulated the in vivo temporal schedule.

The delay of approximately 2 days observed in vivo between the onset of neuron differentiation, SOX9 expression and the induction of NFIA and NFIB is comparable in vitro. The recapitulation of this time in iNPs argues against systemic input and may suggest the timing of the transition is intrinsically controlled. Cell autonomous transitions are well described in Drosophila neuroblasts and increasing evidence suggests they are important in neural development (Toma et al. 2016). An alternative possibility is that the timing is controlled at the tissue level. The in vitro method produces iNPs with high efficiency (>70% SOX2-positive at D5) but also generates some paraxial mesoderm. Therefore, tissue level control must be either neural or mesodermal derived. BRA is required for the induction of paraxial mesoderm in NMPs (Gouti et al. 2017). Assaying glial gene expression in BRA deficient cells would demonstrate the requirement for paraxial mesoderm derived signals in the timing of the transition.

GFAP is a marker of differentiated astrocytes and cells expressing GFAP appear at E17.5 after progenitors have migrated away from the ventricular zone and undergo terminal differentiation. We observed GFAP-positive cells by D15 in vitro
suggesting iNPs were capable of terminal differentiation. However, the efficiency of their generation was low (~2%) suggesting a limiting component. This was not a problem of timing as iNPs cultured to D17 did not generate a higher proportion of astrocytes (data not shown). Activation of JAK/STAT signalling is required for terminal astrocytic differentiation. Phosphorylated STAT3 appears coincident with GFAP expression in vivo and directly binds the Gfap promotor to induce its transcription (Hong and Song 2014). LIF is a ligand which activates the JAK/STAT signalling pathway. Exposure to LIF after NFIA induction increased the percentage of GFAP-positive cells by 150% at later time points. This suggests that JAK/STAT signalling is limiting in vitro. Whether this represents the maximal efficiency of gliogenesis or whether there are additional missing components is not clear. In vivo progenitors migrate away from the ventricular zone into the mantle zone before they begin to express GFAP. In the mantle zone nuclei are less densely packed. One hypothesis is that the in vitro system does not recapitulate this organisation at later stages. Preliminary experiments where progenitors are split into dishes at lower density and exposed to LIF suggest the proportion of glia can be increased (data not shown).

9.2 iNPs recapitulate the phenotypes of glial genes and gives additional insight into the molecular mechanism

SOX9 and NFIA/B are essential components of the underlying network controlling glial lineage activation. Mice deficient for either SOX9, NFIA, or NFIB have severe defects in gliogenesis (Stolt et al. 2003, Deneen et al. 2006, Kang et al. 2012). Together SOX9 and NFIA activate a cohort of target genes, including Apcdd1, important for terminal differentiation. Constitutive expression of Apcdd1 can rescue gliogenesis when Nfia is inhibited (Kang et al. 2012). Deletion of SOX9 results in a delay in NFIA induction, reduced Apcdd1 expression, and a reduction of differentiated astrocytes (Stolt et al. 2003). NFIA deficient mice also have reduced Apcdd1 expression as well as a reduction of GFAP-positive glia (Deneen et al. 2006). NFIB knockout mice have a similar gliogenic phenotype. This along with their structural similarity has led to the proposal that functional redundancy exists between NFIA and NFIB. However, although deletion of NFIA or NFIB results in a
reduction in NFIX expression only NFIB was capable of activating the *Nfix* promotor (Matuzelski et al. 2017). This has led to the proposal that NFIA and NFIB function via distinct mechanisms. It is unclear whether NFIB also interacts with SOX9 to activate target genes such as *Apcdd1*. In Chapter 4 we functionally validate the *in vitro* system by recapitulating the phenotypes of SOX9, NFIA and NFIB deficient mice. We also present new insight into the underlying mechanism controlling glial gene induction.

Deletion of NFIA *in vivo* results in a reduction of SOX9/NFIA target genes, such as *Apcdd1* and a reduction of astrocytes at later stages. In iNPs deficient for NFIA *Apcdd1* expression was still observed but it was reduced compared to WT cells consistent with the *in vivo* phenotype. Whether this reduction represents a delay in SOX9/NFIA target gene induction is unclear without samples taken at later time points. Alternatively, additional mechanisms might exist for glial lineage activation. NFIB has been proposed to be redundant with NFIA despite functional differences. We show that combined deletion of NFIA and NFIB reduces *Apcdd1* expression further. This demonstrates that NFIB can partially compensate for NFIA in glial lineage activation. Whether NFIB also requires SOX9 is not clear.

Conditional deletion of SOX9 reduces NFIA and *Apcdd1* expression in the developing mouse spinal cord (Kang et al. 2012). SOX9 regulation of NFIB has yet to be demonstrated. Removal of SOX9 in iNPs delays NFIA and NFIB expression and caused a reduction in *Apcdd1* expression. Intriguingly, the level of reduction was comparable between SOX9 (<sup>-</sup>) and NFIA/B (<sup>-/-</sup>) hinting at a relationship between SOX9 and NFIA/B in glial lineage activation. Previous work has shown that SOX9 and NFIA co-regulate glial target genes (Kang et al. 2012). How much the delay in NFIA and NFIB contributes to the SOX9 (<sup>-</sup>) phenotype is not clear. At later stages *Apcdd1* recovers in both SOX9 (<sup>-</sup>) and NFIA/B (<sup>-/-</sup>) cells suggesting compensatory mechanisms. Combinatorial deletion of SOX9, NFIA and NFIB would determine whether these involve NFIA/B and SOX9 respectively. However, SOX9 or NFIA alone was not sufficient to activate the *Apcdd1* promotor region. Although distal regulation cannot be dismissed this may suggest alternative inputs for glial lineage activation.
NFIB but not NFIA has been reported to directly regulate the promoter of another NFI family gene, Nfix (Matuzelski et al. 2017). However, deletion of either NFIA or NFIB in mice reduced NFIX expression. Deletion of NFIA did not affect the transcription of Nfix in vitro, contradicting the in vivo data. Whether this represents a difference in the regulation of NFIX transcription and translation, or a greater capacity for NFIB compensation in vitro is not clear. Combined deletion of NFIA and NFIB reduced Nfix expression, confirming the role of NFIB in regulating Nfix.

Without data from NFIB deficient cells, it is difficult to interrogate the function of NFIA in Nfix induction. Although Nfix levels were reduced in NFIA/B\(^{-/-}\), expression still increased at later stages indicating alternative regulatory inputs. SOX9 is a potential candidate as it regulates NFIA and NFIB as well as downstream glial genes (Kang et al. 2012). Nfix expression was reduced in SOX9-deficient cells demonstrating that SOX9 regulates Nfix. The reduction of Nfix was comparable between SOX9\(^{-/-}\) and NFIA/B\(^{-/-}\) as it was for Apcdd1. This suggests a similar dependency on SOX9, NFIA and NFIB. It is not clear if SOX9 and NFIA/NFIB co-regulate Nfix expression.

At later stages the glia progenitors also generate GFAP-positive astrocytes. Deletion of either NFIA or NFIB in vivo reduces the number of cells expressing GFAP (Deneen et al. 2006). It has been hypothesised that the residual gliogenesis is a consequence of redundancy between the two genes. In agreement with previous work, NFIA\(^{-/-}\) cells generated fewer GFAP-positive astrocytes compared to WT. Moreover, we found that deletion of both NFIA and NFIB abrogates terminal astrocytogenesis. This demonstrates functional redundancy between NFIA and NFIB as well as emphasising the requirement for these two genes in terminal differentiation.

Since NFIA and NFIB expression correlates with the termination of neuron differentiation we propose that NFIA and NFIB are the molecular correlates of gliogenesis. Previous work has suggested that constitutive expression of downstream target genes can restore gliogenesis when Nfia is inhibited. However, we find that the residual activation of glial genes downstream of NFIA and NFIB in NFIA/B\(^{-/-}\) cells is not sufficient to compensate for their loss. NFIA and NFIB are expressed throughout the astrocyte lineage, suggesting NFIA and NFIB could have
functions at later stages. For example, NFI proteins were found to bind the Gfap promotor and NFIA was required for its demethylation (Namihira et al. 2009).

Previous work has demonstrated that deletion of SOX9 results in reduced astrocytogenesis (Stolt et al. 2003). We also show that cells lacking SOX9 produce barely detectable numbers of GFAP-positive cells. This corroborates the requirement for SOX9 in gliogenesis. Since deletion of SOX9 has multiple consequences upstream of terminal glial differentiation, it is difficult to directly implicate SOX9 in terminal differentiation. It would be interesting to conditionally delete SOX9 at later stages and assess the severity of the GFAP phenotype.

The SOX-E subfamily consists of Sox8, Sox9, Sox10. The residual astrocyte production in SOX9(-) cells could be attributed to redundancy with other SOX-E family members. Both SOX8 and SOX10 are expressed later than SOX9, at E11 and E11.5 respectively. SOX10 is required for oligodendrogenesis and its expression is restricted to the progenitor domain from which oligodendrocytes are generated. SOX8 is expressed across the dorsal ventral axis making it a more likely candidate (Stolt et al. 2005). However, combined deletion of SOX8 and SOX9 did not significantly change the expression of glial genes, arguing against functional redundancy (data not shown).

9.3 Signalling pathways affecting the neural-to-glial transition

9.3.1 TGF-β signalling links neuron differentiation to the neural-to-glial transition

TGF-β signalling has been implicated in controlling the timing of developmental transitions in the midbrain, hindbrain and spinal cord (Dias et al. 2014). In the developing spinal cord, Gdf11 is expressed in newly born neurons adjacent to the progenitor domain from E10 to E11.5, just prior to NFIA induction (Shi and Liu 2011). At E12.5 we found that TGF-β1-3 were also present in neurons across the dorsal-ventral axis. Therefore, these TGF-β ligands are present during the neural-to-glial transition.
In the developing hindbrain, TGF-β signalling has been proposed to control the timing of a transition from motor neuron to serotonergic neuron production (Dias et al. 2014). Deletion of GDF11 reduced the rate of neuronal differentiation (Shi and Liu 2011). We show that TGF-β2 signalling regulates the expression of the neuronal marker *Tubb3*. Activation of TGF-β2 induced *Tubb3* expression. At earlier stages Sox2 expression is unaffected suggesting that the progenitor pool is maintained. This is consistent with neurons arising from asymmetric cell divisions (Noctor et al. 2004). At later stages Sox2 expression is decreased after cells have been exposed to TGF-β2 signalling, whilst *Tubb3* expression is at WT levels. This might suggest that progenitors generate a finite number of neurons and that the neurogenic phase ends prematurely in cultures exposed to TGF-β2. It would be interesting to perform a lineage tracing experiment to measure the length of neuron differentiation in cells with manipulated TGF-β signalling. Nevertheless, we suggest that these data are consistent with TGF-β affecting the rate of neuronal differentiation.

Previous work has demonstrated that activation of TGF-β signalling *in vivo* and *in vitro* induces an early transition from motor neuron to oligodendrocyte production (Shi and Liu 2011, Dias et al. 2014). It is not clear whether this is also true for the transition from interneuron to astrocyte generation. We show that by manipulating TGF-β signalling we change the schedule of the neural-to-gliarial transition. Sox9, *Nfia*, *Nfib* are precociously induced in response to TGF-β activation. The downstream markers, *Apcdd1* and *Nfix*, were also precociously induced but only after the appropriate delay, suggesting an early transition to gliogenesis. We demonstrate that the early induction of glial genes is not a consequence of increased Sox9 expression. SOX9 deficient cells had levels of induction comparable to WT cells upon stimulation with TGF-β2. This hints at alternative regulatory inputs into NFIA and NFIB. NOTCH signalling has been proposed to directly regulate NFIA (Namihira et al. 2009). Whether TGF-β affects glial genes through NOTCH signalling is not clear.

Continuous inhibition of TGF-β signalling from D5 with SB resulted in a delay in the neural-to-gliarial transition. In these conditions NFIA had yet to recover by D11. A 48 hour exposure to SB from D5 to D7 reduced NFIA expression at D10 but the levels
recovered to WT levels by D11. This intermediate phenotype suggests that sustained TGF-β signalling is required for the correct timing of the neural-to-glial transition. Previous studies had proposed TGF-β signalling regulates the rate of neuronal production (Shi and Liu 2011). This raises the possibility of a link between the rate of neuron differentiation and glial lineage activation, although how this mechanism might integrate into the glial transcriptional network is not clear. However, an implication of this model is that it is not critical when the 48 hour inhibition of TGF-β signalling happens, as long as it falls within the phase of neuron differentiation it will affect glia production. Further work will be necessary to test this directly.

Continuous inhibition of TGF-β signalling was not sufficient to block the activation of glial genes, as Nfia and Nfib began to recover at D11. The reason for this could be technical. SB may not be sufficient completely inhibit TGF-β signalling, or TGF-β signalling is active prior to D5. Another possibility is that alternative signalling pathways affect glial gene expression. The BMP signalling pathway also converges on SMAD4 before regulating transcription. BMPs have also been used to induce glial differentiation of NS cells (Conti et al. 2005, Pollard et al. 2006). We tested whether dual inhibition of TGF-β and BMP signalling was sufficient to block the neural-to-glial transition. However preliminary data indicated that NFIA and NFIB were still induced in these conditions (data not shown). Whether this means alternative signalling pathways are involved is not clear.

Differentiating neurons activate NOTCH signalling in neighbouring cells. NOTCH has previously been shown to regulate SOX9 and NFIA (Kang et al. 2013, Martini et al. 2013). Whether TGF-β affects glial genes through NOTCH remains unclear. Performing TGF-β manipulations in the presence of NOTCH signalling inhibitors would test whether NOTCH is required for the effects of TGF-β. Nevertheless, these data are consistent with a coupling between neuron differentiation and gliogenic onset. This may explain how the timing of the in vitro system so closely resembles that of the developing mouse spinal cord.
9.3.2 NOTCH signalling uncouples neuron differentiation and the neural-to-glial transition

NOTCH signalling plays an essential role in the development of the mammalian CNS. Inhibition of NOTCH signalling results in increased neuron differentiation and an exhaustion of the progenitor pool (Imayoshi et al. 2010). Conversely, activating the NOTCH pathway maintains a cells progenitor status and blocks neuron differentiation. It has also been shown to activate Sox9, Nfia expression and increases gliogenesis (Namihira et al. 2009, Martini et al. 2013). We recapitulate these findings by engineering a lentiviral reporter system to overexpress the NOTCH1 intracellular domain. Activation of NOTCH signalling inhibited neuron differentiation and induced Sox9, Nfia, Nfib expression at D7. This resulted in the induction of the downstream target, Apcdd1. In contrast to TGF-β signalling, NOTCH pathway activation allows the uncoupling of neuron differentiation and gliogenesis. This hints at NOTCH acting downstream of TGF-β in the neural-to-glial transition although experiments to test this directly are required.

NOTCH has been shown to directly regulate Hes genes and at later time points Nfia (Namihira et al. 2009). NFIA negatively regulates Hes gene expression (Piper et al. 2010). Hes genes have been shown to be important for developmental progression of neural progenitors. Overexpression of Hes5 caused accelerated astrogensis at later stages (Bansod et al. 2017). It was not clear whether Hes genes directly affect glial gene expression. We show that overexpression of HES1 only moderately induces Nfia and Nfib transcription, and did not activate Sox9 expression. Therefore, we find no evidence for HES1 directly affecting glial genes, suggesting the astrogenic phenotype is downstream of Hes. What diverts the transcriptional response to NOTCH signalling is not clear.

9.4 Reconstruction of the neural-to-glial transition by single cell sequencing identifies candidate genes

In chapters 4-7, we characterise the in vitro system as a faithful model of the neural-to-glial transition. It recapitulates known phenotypes resulting from the
deletion of glial genes or by manipulating extrinsic signalling cues. This motivated us to attempt to reconstruct the transition. We harvested cells across the transition and obtained their transcriptomes by single cell sequencing. Reconstruction of pseudotime identified 176 candidate genes with a dynamic expression pattern, including 26 transcription factors. This list described the expression patterns of known genes such as Sox9, Nfia, Nfib. It also described the expression of novel genes whose dynamic was confirmed in vitro. Therefore, the analysis provides an accurate molecular description of transcriptional activity across the neural-to-glial transition in vitro.

The method we used to perform these experiments were SMART-seq in combination with the Fluidigm C1 system. This is reported to have greater transcriptional accuracy relative to alternative methods, however single cell sequencing data is noisy (Ziegenhain et al. 2017). We wanted to confirm that the pseudotemporal ordering returned genes relevant to the neural-to-glial transition. Since manipulating TGF-β signalling shifts the timing of the transition in iNPs, we tested the effect of manipulating TGF-β signalling on a selection of candidate genes. The selection of genes showed gene expression changes consistent with the shift in glial gene expression. This links the pseudotemporal reconstruction to the timing of glial lineage activation. We begin with a literature review of a selection of candidates before discussing the role of ZFP536 in the neural-to-glial transition.

9.4.1 Lin28a, Lin28b, Trim71

miRNAs have been directly implicated in the control of genes involved in neural development. For example, miR-153 negatively regulates Nfia transcripts (Tsuyama et al. 2015). We find that three genes involved in miRNA processing, Lin28a, Lin28b (Lin28), Trim71, decrease in expression along pseudotime. TRIM71 modulates LIN28B activity by targeting it for ubiquitylation and degradation (Lee et al. 2014). Trim71 expression decreases later than Lin28a and Lin28b suggesting it is not required after LIN28 downregulation. LIN28A has been shown to promote cell cycle progression. Consistent with this, constitutive expression of LIN28A can partially rescue neural progenitor proliferation in the absence of SOX2 (Li et al.
Preliminary data suggests that forced expression of LIN28A from D4 reduced the expression of the neuronal marker Tubb3 at D7. Additionally, we show that manipulating TGF-β has the opposite effect on Lin28a and Lin28b compared to the effect on Tubb3 and glial genes. Specifically, in conditions where TGF-β signalling is activated and Tubb3 expression induced, Lin28a, Lin28b are downregulated. Whether LIN28A, LIN28B are regulating neuron differentiation or maintaining a progenitor state though increased cell cycle progression is not clear. It would be interesting to analyse the effect of LIN28A overexpression in TGF-β and NOTCH manipulations. Measuring cell cycle dynamics in these conditions would test the role of LIN28 in neural progenitors.

TGF-β manipulations also change the timing of the neural-to-glial transition. This implicates LIN28 in glial gene induction. Previous work has shown that constitutive expression of LIN28A was compatible with neuron differentiation but blocked gliogenesis in vitro (Balzer et al. 2010, Cimadamore et al. 2013). The primary function of LIN28A and LIN28B is to negatively regulate the maturation of the let-7 miRNA family (Piskounova et al. 2011, Tsialikas and Romer-Seibert 2015). Global miRNA profiling of Neural Stem cells undergoing astrocytogenesis identified several let-7 family members upregulated as differentiation progressed (Gioia et al. 2014). Inhibition of let-7b/g in in vitro derived human neural progenitors reduced gliogenesis (Patterson et al. 2014). These data appear to contradict the neurogenic phenotypes and suggest LIN28 might negatively regulate gliogenesis. Another possibility is that LIN28 promotes progenitor maintenance at the expense of both neuron differentiation and gliogenesis. In our preliminary overexpression experiments we observe a modest reduction in Sox9 at D7 without affecting Nfia, Nfib. NFIA and NFIB expression is first detected at D8 in iNPs. D7 may be too early for Nfia, Nfib phenotypes. We did not test the effect of LIN28 at later stages.

9.4.2 Trim8

The JAK/STAT3 signalling pathway is required for the terminal differentiation to GFAP-positive glia. Protein Inhibitor of Activated STAT3 (PIAS3) has been reported to inhibit STAT3 from activating target gene expression (Chung et al. 1997). TRIM8
has been shown to negatively regulate PIAS3 allowing STAT3 to affect transcription (Okumura et al. 2010). Trim8 expression increases across pseudotime. This suggests that TRIM8 enables progenitors to respond to JAK/STAT signalling following glial lineage acquisition. However, Trim8 expression appears far in advance of terminal differentiation making the role of TRIM8 unclear. The expression pattern of Trim8 both in vitro and in vivo has yet to be reported.

### 9.4.3 Cux2

CUX2 has previously been identified to control cell cycle progression in the developing spinal cord (Iulianella et al. 2008). It is not expressed at E9.5 but at E10.5 and E11.5 it is present in progenitors and differentiating neurons across the dorsal-ventral axis (Iulianella et al. 2008). We find Cux2 is induced across pseudotime, an expression pattern confirmed in vitro by qPCR. Its expression appears to follow Sox9 expression but precede Nfia, Nfib consistent with the reported expression pattern in vivo. CUX2 is regulated by NOTCH signalling and is required for dorsal interneuron formation (Iulianella et al. 2009). Consistent with this, we show that Cux2 is induced upon activation of NOTCH signalling. Its expression is also activated after exposure to TGF-β2 possibly due to increased neuron differentiation in these conditions. A role for Cux2 in the neural-to-glial transition has yet to be reported.

### 9.4.4 Npas3

Npas3 has an expression pattern similar to Sox9 in silico and in vitro. In the developing spinal cord Npas3 expression appears to follow neuron differentiation (Shin et al. 2010, Shin and Kim 2013). However, there were no detectable progenitor or neuronal phenotypes after constitutive expression of NPAS3 (Shin and Kim 2013). Gliogenesis was not assayed in this study. The function of NPAS3 in neural development remains unknown.
9.5 ZFP536 affects SOX9 expression and hence the timing of the neural-to-glial transition

*Zfp536* has an expression pattern similar to *Sox9* in pseudotime and *in vitro*. It also responds comparably to *Sox9* to manipulation of TGF-β signalling. Specifically, *Zfp536* expression is transiently induced after TGF-β pathway activation or reduced when signalling is inhibited. Previous work has shown that ZFP536 negatively regulates neuronal differentiation of P19 cells (Qin *et al.* 2009). Additionally, a combination of SOX10, OLIG2, ZFP536 is sufficient to reprogram fibroblasts into oligodendrocyte progenitor cells (Yang *et al.* 2013). This work suggested a role for ZFP536 in oligodendrogenesis but whether it has a more general role in gliogenesis is not clear. In the developing spinal cord oligodendrocytes are generated from the OLIG2-positive motor neuron domain. The expression pattern of ZFP536 *in vivo* is not well described. We hypothesised that if it was oligodendrocyte specific its expression should be restricted ventrally. However, *Zfp536* is expressed across the dorsal-ventral axis. Expression begins at E10.5, and increases at E11.5. No expression was detected at E9.5. *In vivo* SOX9 expression begins at E9.5 shortly after the onset of neuron differentiation. Therefore, *Zfp536* appears to follow SOX9 expression and neuron differentiation *in vivo*.

The spatial and temporal expression pattern of *Zfp536* encouraged us to investigate its role in neural progenitors. We found that ZFP536 is not sufficient to alter *Tubb3* expression suggesting neuronal differentiation is unaffected. These data contradict published work from P19 cells where constitutive ZFP536 expression inhibits neuronal differentiation (Qin *et al.* 2009). The phenotype could be time point and context dependent, or ZFP536 regulation of neuron differentiation may be indirect and assaying later time points could reveal neurogenic phenotypes. However, removal of ZFP536 did not impact the timing or proportion of differentiated neurons *in vitro*. We conclude that ZFP536 is not required for neuron differentiation. This is consistent with *Zfp536* expression in neural progenitors *in vivo*. 

149
The absence of a neuronal phenotype led us to test its role in glial gene induction. To test this, we constitutively expressed ZFP536 \textit{in vitro} using a lentiviral reporter. ZFP536 was sufficient to induce the progenitor marker Sox2, the glial genes Sox9, Nfia, Nfib and their downstream target Apcdd1. This assay does not rule out the effect being a consequence of progenitor proliferation expanding progenitors at the expense of other cell types. However, the expression of the neuronal marker Tubb3 was unaffected suggesting that the number of neurons was not decreased. The effect on Sox9 was significant given that at this stage SOX9 is already expressed. Repeating the experiment at earlier time points, when SOX9 is not expressed, would test whether ZFP536 is sufficient to induce SOX9.

We validated the effect of ZFP536 \textit{in vivo}. ZFP536 overexpression in chick neural tubes was sufficient to activate SOX9 expression, in agreement with our \textit{in vitro} data. We observed SOX9 induction in progenitors across the dorsal-ventral axis suggesting the effect of ZFP536 on SOX9 is not spatially restricted. In addition, the transfected neural tubes were not obviously expanded and we did not observe clusters of electroporated cells. This suggests that overexpression of ZFP536 does not noticeably affect progenitor proliferation. Collectively, our \textit{in vitro} and \textit{in vivo} data demonstrate ZFP536 regulation of SOX9.

To test whether ZFP536 is required for SOX9 expression we deleted it in ES cells using CRISPR/Cas9 and assayed the effect on astrocyte generation. We observed a reduction in SOX9 expression which recovered at later time points. This was sufficient to delay NFIB induction and reduce Apcdd1 expression. In population assays we did not observe any effect on Nfia expression. Additionally, at later stages the percentage of differentiated astrocytes was reduced. Therefore, ZFP536 is required for the proper timing of SOX9 expression and the correct schedule of glial genes. SOX9 was required for ZFP536 to affect the neural-to-glial transition as the combinatorial deletion of SOX9 and ZFP536 was not more deleterious than removal of SOX9 alone. It is not clear whether ZFP536 acts autonomously to affect SOX9 expression or if it is dependent on SOX9 for its function. Performing coimmunoprecipitation experiments would begin to dissect these two possibilities. Additionally, if SOX9 is required for ZFP536 function inducing ZFP536 prior to SOX9 induction or in SOX9\textsuperscript{(1)} cells would have no effect on SOX9 and glial gene
expression. A paralog of ZFP536, ZFP219, shares 39% sequence homology and has been suggested to form a transcription factory complex with SOX9 to regulate chondrogenesis (Takigawa et al. 2010). Whether their similarity is sufficient to retain SOX9 binding properties is not clear. The in vivo expression pattern of ZFP219 has yet to be described. However, its expression was not detected in our single cell sequencing dataset. It is tempting to speculate that ZFP219 and ZFP536 confer lineage specificity to SOX9 function.

9.5.1 What regulates Zfp536?

In vitro Zfp536 expression appeared coincident with Sox9 but its expression seemed to follow SOX9 in vivo. This discrepancy is likely due to the limitations of each technique. Specifically, qPCR is a population assay and in situ hybridisation is not reliably quantitative. Nevertheless, the expression dynamics suggested Zfp536 might be regulated by SOX9. However, we detected no change in Zfp536 expression upon deletion of SOX9. Manipulating TGF-β signalling shifts the timing of glial gene induction and Zfp536 expression. Previous work has suggested TGF-β signalling controls the rate of neuronal production (Shi and Liu 2011). Commitment to neuronal differentiation results in the activation of NOTCH signalling in neighbouring cells. Whilst activation of the NOTCH signalling pathway induces Sox9, Nfia, Nfib, it was not sufficient to induce Zfp536 expression. This suggests that TGF-β signalling does not affect glial genes through NOTCH.

9.6 The timing of the neural-to-glial transition

How developing systems precisely control the temporal generation of cell types is a fundamental question in developmental biology. Timing mechanisms can operate at three distinct levels: systemically, at the tissue level, and cell autonomously. Examples of all three can be found in neural development (Toma et al. 2016). In this study we investigate the timing of a developmental transition and find evidence of both tissue level and cell autonomous control. In this section, we discuss our findings in the context of developmental timing mechanisms.
9.6.1 Cell autonomous control of the transition

In the *Drosophila* nerve cord a well-defined transcriptional cascade controls the sequential generation of neuronal subtypes. The downregulation of Hunchback by Seven-up triggers the sequential induction of Krüppel, Pdm1, Castor and Grainyhead (Pearson and Doe 2003, Kanai *et al.* 2005). There is emerging evidence of transcriptional cascades functioning in mammalian neural development although the molecular detail is limited. In the developing spinal cord, the neural-to-glial transition is under transcriptional control. SOX9 induction of NFIA and NFIB results in the activation of APCDD1, NFIX and terminal gliogenesis (Kang *et al.* 2012, Matuzelski *et al.* 2017). SOX9 expression appears two days prior to NFIA and NFIB induction. Additionally, the chromatin around the *Nfia* locus appears to be organised when SOX9 is first expressed (Glasgow *et al.* 2017). Therefore, it is unclear what controls the delay between SOX9 expression and NFIA and NFIB induction.

To obtain a detailed description of the molecular events during the transition we utilised single cell transcriptomics and pseudotemporal reconstruction. This identified ZFP536, a transcription factor which regulates the timing of glial lineage acquisition. The mechanism of ZFP536 function is either to regulate the expression of SOX9 or to act via SOX9 to regulate glial genes. The former explanation implies the level of SOX9 is a determinant of NFIA and NFIB induction. The latter suggests that the regulatory partners effect the efficiency of SOX9 target gene induction. Although subtly different, both hypotheses support an “accumulation to threshold” model whereby the rate of SOX9 production controls the timing of glial lineage activation. This is similar to the mechanism of proneural genes such as ASCL1. Accumulation of ASCL1 in neural progenitors results in the down regulation of *Hes* genes and commitment to a neuronal fate (Imayoshi *et al.* 2013).

An “accumulation to threshold” model for SOX9 function requires a detailed understanding of Sox9 regulation. In this study we describe ZFP536 as being important. In somatic tissues SOX9 has been described to autoregulate itself (Mead *et al.* 2013). It is not clear whether SOX9 autoregulation is present in neural tissues. The function of ZFP536 may be to promote SOX9 autoregulation,
increasing the rate of SOX9 accumulation. This would imply that SOX9 autoregulation is dependent on cofactors. However, there are two alternative possibilities: that SOX9 autoregulation is inefficient or that negative regulatory inputs buffer SOX9 expression. miR-124 has been proposed to inhibit Sox9 in neural progenitors undergoing neuron differentiation (Cheng et al. 2009). It is unclear if a similar mechanism exists in spinal cord neural progenitors. Nevertheless, these data motivate further investigation of the regulatory inputs affecting Sox9 expression.

In an “accumulation to threshold” model, the threshold is a key determinant of the timing of a transition. We demonstrate that NFIA and NFIB are the molecular correlates of glial lineage activation. What sets the threshold determining their expression? One possibility is that negative regulatory inputs must be overcome. In the cortex, miR-153 was found to inhibit Nfia and Nfib and its downregulation is required for NFIA and NFIB expression (Tsuyama et al. 2015). An alternative possibility is that the accessibility of regulatory elements controls the timing of their expression. However, previously identified regulatory regions around Nfia have been shown to associate with the transcriptional start site in advance of Nfia transcription (Glasgow et al. 2017). Although alternative regulatory elements may exist, these data suggest that the chromatin organisation is not a limiting factor.

### 9.6.2 Tissue level control of the transition

Our study also indicates evidence of tissue level control of the neural-to-glial transition. Previous work has implicated TGF-β signalling in the control of developmental timing. The source of TGF-β ligands in the developing spinal cord remains to be properly investigated, but it has been suggested that they are produced by differentiated neurons (Shi and Liu 2011). In the developing hindbrain, activation of TGF-β induces a transition between two phases of neuron differentiation (Dias et al. 2014). In iNPs, exposure to TGF-β2 promotes the early activation of glial genes. It has been proposed that TGF-β signalling affects the rate of neuronal differentiation (Shi and Liu 2011). This might suggest that the length of
neuron differentiation and/or total neuronal production may control the timing of glial gene expression.

What links neuronal production to the transcriptional network controlling gliogenesis? As progenitors commit to a neuronal fate they activate NOTCH signalling in neighbouring cells. As well as having well described roles in neural progenitor maintenance, we and others have demonstrated that NOTCH signalling also induces glial genes (Namihira et al. 2009, Martini et al. 2013). NOTCH might regulate glial genes directly, through validated targets such as HES1, or through novel targets such as CUX2. Nevertheless, NOTCH signalling may provide the mechanistic link between the control of neuron differentiation by TGF-β signalling and the activation of the glial transcriptional network.

In summary, the data presented in this study along with previously published work indicate the timing of the neural-to-glial transition is under both cell autonomous and tissue level control. How these multiple regulatory inputs are integrated and their relative weights requires further investigation. The application of mathematical models in combination with misexpression experiments will further illuminate the underlying mechanism.

9.7 Conclusions

In this study, we highlight the potential of in vitro model systems for interrogating the molecular mechanisms underlying developmental processes. We extend a previously characterised differentiation protocol which takes ES cells to neural progenitors with a spinal cord identity (Gouti et al. 2014). This minimal system recapitulates the schedule of gene expression dynamics observed in the developing spinal cord. Specifically, progenitors undergo neuronal differentiation, upregulate glial progenitor markers, and generate astrocytes. Furthermore, the system recapitulated the known phenotypes of key genes required for gliogenesis. We used the in vitro model as a tool to expand our understand of the molecular processes underlying the neural-to-glial transition.
Previous work has demonstrated that TGF-β signalling controls the timing of a transition between two waves of neuronal differentiation in the developing hindbrain. It was also shown to affect the timing of gliogenesis at later stages in the ventral spinal cord (Dias et al. 2014). However, whether this was a general mechanism affecting glial lineage onset was not clear. We build on our current understanding by showing that manipulation of TGF-β signalling affects the timing of glial lineage onset. We hypothesise that this mechanism is a feature of the transition as TGF-β ligands are expressed along the dorsal-ventral axis in neurons (Mecha et al. 2008, Shi and Liu 2011). Previous work has suggested that the function of TGF-β signalling is to control the rate of neuronal production (Shi and Liu 2011). This suggests a systematic coupling between neuron differentiation and the onset of the glial lineage. As neurons differentiate they activate NOTCH in the neighbouring cell and maintain its progenitor status. NOTCH also induces glial gene expression. We report, in agreement with previous work, that by activating the NOTCH signalling pathway neuron differentiation and gliogenesis can be mechanistically uncoupled.

Although genes such as SOX9, NFIA and NFIB have been shown to be required for gliogenesis, molecular detail of the network regulating their expression is limited. We use single cell transcriptomics combined with pseudotemporal reconstruction to provide a complete description of the transcriptional dynamics of \textit{in vitro} transitioning progenitors. This identified ZFP536, a putative transcription factor, which controls the expression of SOX9 and subsequently the timing of the neural-to-glial transition. ZFP536 is upstream of both SOX9 and NOTCH. Our data does not discriminate between ZFP536 acting autonomously to affect SOX9 expression or functioning with SOX9 to orchestrate the neural-to-glial transition.

How progenitors generate distinct cell types with temporal precision is a fundamental question in developmental biology. An increase in cortex size is one of the evolutionary mechanisms associated with the increased cognitive abilities of humanoids (Rakic 2009). Therefore, the control of developmental transitions has important functional and evolutionary consequences. A variety of mechanisms have been proposed to encode time in developing systems (Toma et al. 2016, Ebisuya and Briscoe 2018). In our investigation of a transition in the developing
spinal cord we find evidence for both tissue level and cell intrinsic control. A gene regulatory network in progenitors controls the acquisition of glial fate (Kang et al. 2012). It also integrates signalling cues from differentiated cell types (Namihira et al. 2009). The presence of multiple mechanisms controlling temporal precision may impart robustness to developmental decisions.
Appendix A

Script to generate pseudotemporal ordering from time gene modules

```r
library('ggplot2')
library('gplots')
library('apcluster')
library('RColorBrewer')
library('viridis')
setwd("/Users/watsont/Documents/SingleCellSeq/201807_JulienAnalysis/LogData")
load("/Volumes/lab-briscoej/working/julien/Muscle/inst/extdata/all_available_180426/analysis_Gliogenesis_Simple/D5toD9_180618_40gms_D5rep2/final_workspace.Rdata")
all.Data <- log(readcounts_normalized+1)
# all.Data <- readcounts_normalized
pdf(paste0(format(Sys.Date(),"%Y%m%d"),'_TimeGeneModulesHeatmap.pdf'),width=40,height =40)
gplots::heatmap.2(all.Data[unlist(genemodules2),], trace="none", Colv=FALSE, Rowv=FALSE,
  # RowSideColors=genemodules2[1:4],
  col=colorRampPalette(c("whitesmoke", "black")),
  lhei=c(0.5,20),
  lwid=c(0.2,10),
  margin=c(1.0,5.0),
  key = FALSE,
  cexRow=.50)
de.v.off()

pdf(paste0(format(Sys.Date(),"%Y%m%d"),'_TimeGeneModules.pdf'), width=10, height=10,
useDingbats=FALSE)
par(mfrow=c(2,2))
for (gm in 1:length(genemodules2)){
  plot(colMeans(all.Data[unlist(genemodules2[gm])]), pch=16, col=factor(getPopName(all.Data)), ylab='Mean Expression', main=gm)
}
de.v.off()

# average gene modules
m1 <- colMeans(all.Data[unlist(genemodules2[1])],)
m2 <- colMeans(all.Data[unlist(genemodules2[2])],)
m3 <- colMeans(all.Data[unlist(genemodules2[3])],)
m4 <- colMeans(all.Data[unlist(genemodules2[4])],)
GM.df <- rbind(m1,m2,m3,m4)

# generate PCA
TimePCA <- prcomp(t(GM.df), center=TRUE, scale=TRUE)
my_palette <- rev(viridis(n=5))

#### The functions getPopName and plotPCA were written by Julien Delile ####
```

157
getPopName <- function(mat){
  return( factor(unlist(lapply(colnames(mat), function(x) strsplit(x,"_"))[1][1]))) # second "1"
}

plotPCA <- function(pca.rot, x1="PC1", x2="PC2", cellnames=FALSE){
  # Plot PCA
  library(ggplot2)
  library(viridis)
  data.pca.plot=data.frame(row.names = rownames(pca.rot))
  data.pca.plot$x=pca.rot[,x1]
  data.pca.plot$y=pca.rot[,x2]
  data.pca.plot$pt.size=2
  # Show population names...
  p=ggplot(data.pca.plot,aes(x=x,y=y))+geom_point(aes(colour=factor(getPopName(t(pca.rot)))),size=3)+scale_colour_viridis(direction=-1, discrete=TRUE)+theme_classic()
  #
  p=ggplot(data.pca.plot,aes(x=x,y=y))+geom_point(aes(colour=(unname(x)[,1]+.0001),size=3))#
  +scale_colour_gradient2()#
  # Show gene levels: "Gene", Population name
  #
  p=ggplot(data.pca.plot,aes(x=x,y=y))+geom_point(aes(colour=as.numeric(Progenitor.Data[Gene,.size=3])+scale_colour_gradient2()
  p2=p+xlab(x1)+ylab(x2)+scale_size(range = c(2, 2))
  p3=p2+theme(axis.title.x = element_text(face="bold", colour="#990000", size=24), axis.text.x =
  element_text(angle=90, vjust=0.5, size=16))+theme(axis.title.y = element_text(face="bold",
  colour="#990000", size=24), axis.text.y = element_text(angle=90, vjust=0.5,
  size=16))+guides(colour = guide_legend(override.aes = list(size=3)))+theme(legend.text =
  element_text(size = 10, face = "bold"))+theme(legend.title=element_blank())+theme_classic()
  p4=p3+theme(legend.title=element_blank())
  # Use the following lines to print cell names in plot
  if(cellnames==TRUE){
    cell.names=factor(unlist(lapply(rownames(pca.rot), function(x) strsplit(x,"_"))[1][2]))) # 2 for second item
    data.pca.plot$cell.names = cell.names
    p4+geom_text(data=data.pca.plot, mapping=aes(x=x, y=y, label=cell.names), size=3,
    vjust=2, hjust=0.5)
  }else{
    print(p4)
  }
}

df(pdf(paste0(format(Sys.Date(),"%Y%m%d"),'_Julien_TimePCA_PC1PC2.pdf'), useDingbats =
FALSE)
PCA <- plotPCA(data.frame(pca_res$x), x1="PC2", x2="PC1", cellnames=FALSE)
deVoff()

df(pdf(paste0(format(Sys.Date(),"%Y%m%d"),'_Tom_TimePCA_PC1PC2.pdf'), useDingbats =
FALSE)
PCA <- plotPCA(data.frame(TimePCA$x), x1="PC2", x2="PC1", cellnames=FALSE)
deVoff()
all.Data <- all.Data[which(rowSums(all.Data > 0) >= 10),]

# xlab="PC2 (20%)", ylab="PC1 (51%)"

# Identify and remove outlier cells on PCA
pdf(paste0(format(Sys.Date(),"%Y%m%d","_HistogramAcrossPC1.pdf"), height=10, width=10, useDingbats = FALSE))
PC1hist.data <- hist(pca_res$x[,"PC1"], breaks=seq((min(pca_res$x[,"PC1"])), (max(pca_res$x[,"PC1"])), length.out=30), col=viridis(n=30), plot=TRUE)
PC1 <- pca_res$x[,"PC1"][which(pca_res$x[,"PC1"] <= PC1hist.data$breaks[which(PC1hist.data$counts == 0)])[1]]) # remove outlier
all.Data <- all.Data[,names(PC1)]
dev.off()

# Replot PCA without outlier
pdf(paste0(format(Sys.Date(),"%Y%m%d","_TimePCAMinusOutlier_PC1PC2.pdf"), useDingbats = FALSE))
PCA <- plotPCA(data.frame(pca_res$x[names(PC1),]), x1="PC2", x2="PC1", cellnames=FALSE)
dev.off()

# Plot the distribution fo samples through pseudotime without outlier
pdf(paste0(format(Sys.Date(),"%Y%m%d","_DistributionSamplesPseudotime.pdf"), height=10, width=5, useDingbats = FALSE))
test=data.frame(PC1=PC1, t=getPopName(all.Data))
 ggplot(test, aes(x=t, y=PC1, fill=t)) + geom_boxplot() + scale_fill_viridis(direction=-1, discrete=TRUE) + geom_jitter(shape=16, position=position_jitter(0.1)) + theme_classic()
dev.off()

some_genes = c('Lin28a', 'Sox9', 'Zfp536', 'Nfia', 'Nfib', 'Ttyh1', 'Nfix')
some_genes = 'Nfib'

# Loop to find best smoothing function and span.
pdf(paste0(format(Sys.Date(),"%Y%m%d","_Span_SomeGenes_PC1.pdf"), width=10, height=10, useDingbats=FALSE))
par(mfrow=c(3,3))
for(Gene in some_genes){

  matplot(PC1,all.Data[Gene,], pch=16, main=Gene)
  for(span in seq(.5, 1, length.out=5)){
    df = data.frame("PT"=PC1, "TF"=all.Data[Gene,])
    lo <- loess("TF~PT", data = df, span = span)
    fit <- predict(lo, newdata=data.frame("PT"=seq((min(PC1)), (max(PC1)), length.out=20)))
    lines(seq((min(PC1)), (max(PC1)), length.out=20),fit / max(fit) * max(all.Data[Gene,]), col=sample(viridis(n=15)))
  }
}

# for(f in seq(.5, 1, length.out=5)){
#   fit2 = lowess(PC1,all.Data[Gene,], f=f)
#   lines(fit2$x, fit2$y / max(fit2$y) * max(all.Data[Gene,]), col="blue")

159
# Compute smooth function for all genes - written by JULIEN DELILE
fit_list <- lapply(rownames(all.Data), function (gn) {
  x = all.Data[gn,]
  df = data.frame("PT"=PC1, "TF"=x)
  lo <- loess("TF~PT", data = df, span = .6)
  fit <- predict(lo, newdata=data.frame("PT"=seq((min(PC1)), (max(PC1)), length.out=20)))
  ss.dist <- sum(scale(x, scale=FALSE)^2)
  ss.resid <- sum(resid(lo)^2)
  Rsq <- 1-ss.resid/ss.dist
  return(list("fitted.data"=fit, "Rsquare"=Rsq))
})

# Fitted value dataframe
df2 = do.call(rbind, lapply(fit_list, function(x) x$fitted.data))
rownames(df2) <- rownames(all.Data)
df2[which(df2 < 0)] <- 0 # some predict value are negative

# Filter out genes based on dispersion
dispersion_all = apply(as.matrix(df2) , 1 , function(x) var(x)/mean(x))
dispersion_cutoff = mean(dispersion_all) + 3/2*(sd(dispersion_all)) # calculate 1.5 sd away from mean

# Filter genes with low mean expression
mean_level_filtered = names(which(rowMeans(df2) > 1.4)) #4785 genes survive
Appendix

plot(rowMeans(df2), pch=16)
abline(h = 1.4, col = "red")
plot(sort(rowMeans(df2)), pch=16)
abline(h = 1.4, col = "red")
hist(rowMeans(df2))
dev.off()

# Combine fit, dispersion and expression lists to get final selection
selected_genes = Reduce(intersect, list(rsquare_filtered, dispersion_filtered,
mean_level_filtered))    #121 genes left

The following code was written by JULIEN DELILE
# Order clusters on heatmap
data.sc = df2[selected_genes,]

# These function:
#   * identifies peak regions
#   * filters them
#   * returns a peak id mask
filter_split_discretize_peaks <- function(curve, genename, plot=FALSE, min_level_ratio=.2,
peak_size_ratio=.1){
  #curve=mat_path_filtered2['Myc',]
  #curve=mat_path_filtered2['Isl1',]
  #curve=mat_path_filtered2['Hmbox1',]
  # Step 1: set low level to zero
  zero_thres = min_level_ratio * max(curve)
  curve_zeroed=curve
  curve_zeroed[which(curve<zero_thres)] <- 0
  # Step 2: identify peaks between zero-level regions
  peak_id = matrix(0,nrow=1, ncol=length(curve))
  curr_peak=1
  for(i in 1:length(curve)){
    if(curve_zeroed[i] != 0){
      peak_id[i]=curr_peak
    }else{
      if(i-1 != 0){
        if(curve_zeroed[i-1] != 0){
          curr_peak <- curr_peak+1
        }
      }
    }
  }
  # compute "size" (i.e. area/timestep) for all peaks
  peaks_size=as.data.frame(tapply(curve_zeroed, peak_id, FUN=sum)) # Sum curve values by peak_id groups
  # Remove peak "0" from peaks_size dataframe if it exists
  if(0' %in% rownames(peaks_size) ){
    peaks_size=as.data.frame(peaks_size[-which(rownames(peaks_size)=="0"),])
  }
}

161
# Step 3: keep only top-2 peaks if more peaks exist
if(dim(peaks_size)[1] > 2){
    peaks_size=as.data.frame(peaks_size[tail(order(peaks_size),n=2),])
}

# Step 4: Exclude second peak if it is too small
if(dim(peaks_size)[1] == 2){
    total_peaks_size=sum(peaks_size)
    peaks_size_threshold=peak_size_ratio*total_peaks_size # threshold for filtering out small
    peaks is a ratio of the total peaks "sizes"
    peaks_size = as.data.frame(peaks_size[which(peaks_size > peaks_size_threshold), ])
}

kept_peak_list=rownames(peaks_size)
num_peak=length(kept_peak_list)

# Gaussian fitting
#peak_values=NULL
pc=1
peak_mask = rep(0,length(curve))
for(pid in kept_peak_list){

    peak_pos=which(peak_id==pid)
    peak_length=length(peak_pos)
    peak_avg=mean(curve[peak_pos])
    tab <- data.frame(x=peak_pos, curve=curve[peak_pos])

    # NOTE: initial sigma should be evaluated from a priori knowledge (ex: sigma = 5h)
    # this would require to rescale the pseudo-time axis with hour units for example,
    res<-NULL
    try{
        res <- nls( curve ~ k*exp(-1/2*(x-mu)^2/sigma^2),
                    data = tab,
                    start=c(mu=peak_pos[as.integer(.5*peak_length)],sigma=5,k=peak_avg),
                    control=nls.control(maxiter=1000, minFactor=1e-9)
                ),
        silent=TRUE
    }

    if(!is.null(res)){
        v <- summary(res)$parameters,"Estimate"
        #start/end = mu +/- 2 sigma
        peak_start = max( peak_pos[1], as.integer(v[1]-2*v[2]) )
        if(peak_start < tail(peak_pos,n=1)){  # this case may happen with late genes
            peak_start = peak_pos[1]
        }
        peak_end = min( tail(peak_pos,n=1), as.integer(v[1]+2*v[2]) )
        if(peak_end < peak_start){
            peak_end=tail(peak_pos,n=1)
        }
        fit_outcome = "succeeded"
    }
    else {
        # if gaussian fit does not work, try to fit a
        peak_start = peak_pos[1]
        peak_end = tail(peak_pos,n=1)
    }

    plot(curve~x, data=tab)
    #plot(function(x) v[3]*exp(-1/2*(x-v[1])^2/v[2]^2),col=2,add=T,xlim=range(tab$x) )
} else {
    # if gaussian fit does not work, try to fit a
    peak_start = peak_pos[1]
    peak_end = tail(peak_pos,n=1)
fit_outcome = "failed"
}

print(paste0('peak ',pid,' : start ',peak_start,' / end ',peak_end,' / Gaussian function fitting ',fit_outcome))

# peak_curve = rep(0,length(curve))
# peak_curve[peak_start:peak_end] = curve[peak_start:peak_end]
peaks_mask[peak_start:peak_end] = pc

# peak_name = genename
# if(num_peak==2){
# peak_name=paste0(peak_name, "_peak", toupper(letters[pc]))
#
}

# peak_values=rbind(peak_values, matrix(peak_curve, nrow=1, ncol=length(curve),
dimnames=list(peak_name) ) )
pc <- pc + 1

# Set gene level to zero for out-of-peak region
curve_zeroed2 = curve_zeroed
curve_zeroed2[which(!(peak_id %in% kept_peak_list))] <- 0

# Optional display of the process
if(plot==TRUE){
  par(mfrow=c(4,1))
  plot(curve)
  plot(curve_zeroed)
  plot(curve_zeroed2)
  plot(peak_mask)
  par(mfrow=c(1,1))
}
return(peak_mask)

# apply peak filtering/splitting/binarizing function to each rows/genes
peak_mask=do.call(rbind, parallel::mclapply(rownames(data.sc),
  function(x){filter_split_discretize_peaks(data.sc[x,])}, mc.cores=12) )
rownames(peak_mask)=rownames(data.sc) # should be merged with previous line with a proper R command !

# Build peak dataset
peakA_mask = peak_mask
peakA_mask[which(peak_mask==2)] <- 0

peakB_mask = peak_mask
peakB_mask[which(peak_mask==1)] <- 0
peakB_mask[which(peak_mask==2)] <- 1
rownames(peakB_mask) = paste0(rownames(peakB_mask), "_2nd_peak")

peak_all = rbind(peakA_mask, peakB_mask)
# remove empty peakB lines
peak_all <- peak_all[-which(rowSums(peak_all) == 0),]

rownames(peak_starts) = rownames(peak_all)
peak_ends = as.matrix(apply(peak_all,1,function(x){tail(which(x==1),1)}))
rownames(peak_ends) = rownames(peak_all)
peak_lengths = as.matrix(rowSums(peak_all == 1))

# # Build pseudo time data by peak
# pt.genes.exp.peak.only.A = path.profile.allgenes.smooth
# pt.genes.exp.peak.only.A[which(peakA_mask==0)] <- 0
# pt.genes.exp.peak.only.B = path.profile.allgenes.smooth
# pt.genes.exp.peak.only.B[which(peakB_mask==0)] <- 0
# rownames(pt.genes.exp.peak.only.B) = paste0(rownames(pt.genes.exp.peak.only.B),
# "_2nd_peak")
# pt.genes.exp.peak.only = rbind(pt.genes.exp.peak.only.A, pt.genes.exp.peak.only.B)
# pt.genes.exp.peak.only <- pt.genes.exp.peak.only[which(rowSums(pt.genes.exp.peak.only)
# == 0),]

peak_maxs = as.matrix(apply(data.sc, 1, which.max))
rownames(peak_maxs) = rownames(data.sc)

# Normalize gene profile between x and 1
mydata <- t(apply(data.sc, 1, function(x)(x/(max(x)))))

pdf(paste0(format(Sys.Date(), "%Y%m%d"),'_SomeGenesSmoothed_PC1.pdf'), width=20,
    height=20)
par(mfrow=c(3,3))
for (Gene in some_genes){
    plot(mydata[Gene,], type="l", main=Gene, cex.main=3.0, xlab="Pseudotime", ylab="Smoothed
    Expression", cex.lab=1.5)
}
dev.off()

# Method5 / Affinity propagation
# library(apcluster)
# d.apclus <- apcluster::apcluster(apcluster::negDistMat(r=2), mydata, q=.0) # q controls
# number of clusters
# d.apclus <- apclusterK(apcluster::negDistMat(r=2), mydata, K=12,verbose=TRUE) #
# with target number of clusters
# cat("affinity propagation optimal number of clusters: ", length(d.apclus@clusters), "n")
# heatmap(d.apclus)
# plot(d.apclus, mydata[,1:15])
# nclusters=length(d.apclus@clusters)
# clusters=matrix(0,nrow=dim(mydata)[1], ncol=1, dimnames=list(rownames(mydata),
# c('cluster_id')))for(i in 1:nclusters){
    for(j in d.apclus@clusters[[i]]){
        clusters[j] = i
    }
}

# plot clusters
gg_color_hue <- function(n) {
    hues = seq(15, 375, length = n + 1)
    hcl(h = hues, l = 65, c = 100)[1:n]
}
# mycol = gg_color_hue(nclusters)
# getPalette <- colorRampPalette(RColorBrewer::brewer.pal(11, "Spectral") ) # required to get more than 9 colors
# mycol = getPalette(nclusters)
mycol = viridis(nclusters)

clusters_ordered=unlist(lapply(1:nclusters, function(i){ rep(i, times=length(d.apclus@clusters[[i]])) }))

my_palette <- colorRampPalette(c("whitesmoke", "black"))(n = 1000)
# my_palette <- viridis(n=1000)

gplots::heatmap.2(mydata[rownames(unlist(d.apclus@clusters)),], trace="none", Colv=FALSE, Rowv=FALSE,
   RowSideColors=mycol[clusters_ordered],
   col=my_palette,
   scale="row"
)
dev.off()

# ###########################################
# Sort clusters by mean start and mean length
# ###########################################

# get the gene ids for each cluster
clusters_idx=lapply(1:nclusters, function(x){ which(clusters==x) })
# get the average starting point for each cluster
clusters_meanstart=as.matrix(unlist(lapply(clusters_idx, function(x) as.integer(mean(peak_starts[rownames(mydata)[x],])))), byrow=TRUE,nrow=1)
clusters_meanlength=as.matrix(unlist(lapply(clusters_idx, function(x) as.integer(mean(peak_lengths[rownames(mydata)[x],])))), byrow=TRUE,nrow=1)
clusters_meanmax=as.matrix(unlist(lapply(clusters_idx, function(x) as.integer(mean(peak_maxs[rownames(mydata)[x],]])))), byrow=TRUE,nrow=1)

# get the rank of each cluster in term of average starting point
#clusters_meanstart_rank=rank(t(clusters_meanstart), ties.method = "random")
#clusters_meanlength_rank=rank(t(clusters_meanlength), ties.method = "random")

mat_clust_id=as.matrix(clusters)
mat_clust_meanstart=apply(mat_clust_id,1,FUN=function(x){ clusters_meanstart[x] })
mat_clust_meanlength=apply(mat_clust_id,1,FUN=function(x) clusters_meanlength[x ] }
mat_clust_meanmax=apply(mat_clust_id,1,FUN=function(x){ clusters_meanmax[x ] })#
mat_clust_meanstart_rank=apply(mat_clust_id,1,FUN=function(x) clusters_meanstart_rank[x ] })#
mat_clust_meanlength_rank=apply(mat_clust_id,1,FUN=function(x) clusters_meanlength_rank [x ]})#
mat_peakstart=peak_starts[rownames(mat_clust_id)]
mat_peaklength=peak_lengths[rownames(mat_clust_id)]

mat_merged=cbind(mat_clust_id , mat_clust_meanstart_sorted, matPeakstart, mat_peaklength)
mat_merged=cbind(mat_clust_id , mat_clust_meanmax, mat_clust_meanstart, mat_clust_meanlength)

# sort by 1. cluster max, 2. cl. start, 3. cl. id
mat_reordered=mat_merged[order(mat_merged[,2],mat_merged[,3],mat_merged[,1]),]
# sort by 1. cl. start, 2.cl. length, 3. cl. id
# mat_reordered=mat_merged[order(mat_merged[,3],mat_merged[,4],mat_merged[,1]),]
clusters_rank=unique(mat_reordered[,1])

# plot sorted clusters
# gplots::heatmap.2(peak_all[rownames(mat_reordered),], trace="none", Colv=FALSE, Rowv=FALSE,
#                  RowSideColors=mycol[mat_reordered[,1]],
#                  col=my_palette
# )
# gplots::heatmap.2(pt.genes.exp.peak.only[rownames(mat_reordered),], trace="none", Colv=FALSE, Rowv=FALSE,
#                   RowSideColors=mycol[mat_reordered[,1]],
#                   col=my_palette,
#                   scale="row"
# )
# gplots::heatmap.2(pt.genes.exp.peak.only.norm[rownames(mat_reordered),], trace="none", Colv=FALSE, Rowv=FALSE,
#                   RowSideColors=mycol[mat_reordered[,1]],
#                   col=my_palette
# )
# gplots::heatmap.2(mydata[rownames(mat_reordered),], trace="none", Colv=FALSE, Rowv=FALSE,
#                   RowSideColors=mycol[mat_reordered[,1]],
#                   col=my_palette
# )
# ylGnBl=colorRampPalette('ylGnBl5')
# brewer.pal(299,ylGnBl)
# display.brewer.pal(299,"YlGnBu")
# my_pal=colorRampPalette(brewer.pal(n=9,"YlGnBu"))

# Plot all clusters
ordered.data = mydata[rownames(mat_reordered),]
mycol2=rev(mycol)[order(unique(mat_reordered[,1]))]
ordered.rowsidecolors = mycol2[mat_reordered[,1]]
keptgenes = rownames(ordered.data)

# plotPopNGenesGraph(gf2, lay2, as.data.frame(data$smoothdatalog_v2_d1), keptgenes)

pdf(paste0(format(Sys.Date(),"%Y%m%d"),'_SelectedGenesPseudotime.pdf'),width=10,height=30, useDingbats=FALSE)
gplots::heatmap.2(ordered.data[keptgenes,], trace="none", Colv=FALSE, Rowv=FALSE,
                   RowSideColors=ordered.rowsidecolors,
                   col=my_palette,
                   lhei=c(0.5,20),
                   lwid=c(0.2,10),
                   margin=c(1.0,5.0),
                   key = FALSE,
                   cexRow=1,
                   xlab= "Pseudotime")
dev.off()

# # Plot specific cluster
# cl_rank = 4
Appendix

# cl_id=clusters_rank[cl_rank]
# cl_id=mat_reordered['Nfia',1]

# keptgenes_ids=which(mat_reordered[,1] == cl_id)
# keptgenes=rownames(mat_reordered)[which(mat_reordered[,1] == cl_id)]

# pdf("tmp/test5.pdf",width=7,height=7)
# gplots::heatmap.2(ordered.data[,1][keptgenes], trace="none", Colv=FALSE, Rowv=FALSE,
#                   RowSideColors=order.rowsidecolors[keptgenes_ids],
#                   col=colorRampPalette(c("white","black")),
#                   lhei=c(0.5,20),
#                   lwid=c(0.2,10),
#                   margin=c(2.0,6.0),
#                   key = FALSE,
#                   cexRow=1.0)
# dev.off()

# Tf-only

TFList <- read.table("~/Documents/SingleCellSeq/gene lists/TF_mouse_150722.dat")
TFList <- TFList[,3]
TFList <- intersect(rownames(data.sc),TFList)
tfnames = rownames(mat_reordered)[rownames(mat_reordered) %in% TFList]

ordered.data = mydata[rownames(mat_reordered[tfnames]),]
ordered.rowsidecolors = mycol2[mat_reordered[tfnames,1]]
keptgenes_ids = which(rownames(ordered.data) !="") # all genes
keptgenes = rownames(ordered.data)[keptgenes_ids]

pdf(paste0(format(Sys.Date(),"%Y%m%d"),"_SelectedTFsPseudotime.pdf"),width=10,height=10
useDingbats=FALSE)
gplots::heatmap.2(ordered.data[,], trace="none", Colv=FALSE, Rowv=FALSE,
                   RowSideColors=order.rowsidecolors,
                   col=my_palette,
                   density.info="none", # turns off density plot inside color legend
                   lhei=c(0.5,20),
                   lwid=c(0.2,10),
                   margin=c(5.0,10.0),
                   key=FALSE,
                   cexRow=1.4,
                   cexCol=1.5,
                   xlab= "Pseudotime")
dev.off()

# cl_id=mat_reordered[c('Nfia', 'Nfib', 'Sox9', 'Rfx4'),1]

# keptgenes_ids=which(mat_reordered[,1] == cl_id)
# keptgenes=rownames(mat_reordered)[which(mat_reordered[,1] == cl_id)]
Appendix B

Script to compare in silico and in vitro expression data

```r
expt.data <- read.table("/Users/watsont/Documents/SingleCellSeq/201807_JulienAnalysis/LogData/20180714_SingleCellCandidates_qPCRMeans.txt", header=TRUE, row.names=1)

# subset out day 5 to day 9 to match single cell time points
expt.data2 <- expt.data[2:6]

# Normalise data to max
expt.data.norm <- t(apply(expt.data2, 2, function (x) {x/(max(x))}))

# mydata is from Appendix A script
some_genes <- intersect(rownames(ordered.data), rownames(expt.data.norm))

# Plot heatmap of candidates validated by qPCR
pdf(paste0(format(Sys.Date() ,"%Y%m%d"), "Validation_qPCR_14July2018.pdf"), width=10, height=10)
gplots::heatmap.2(as.matrix(expt.data.norm[some_genes,]), trace="none",
                  Rowv=as.dendrogram(hc.df2.norm.gdfit.final),
                  Rowv=FALSE,
                  Colv=FALSE,
                  # key=FALSE,
                  # dendrogram ="row",
                  col=colorRampPalette(c("whitesmoke", "black")),
                  density.info="none", # turns off density plot inside color legend
                  # scale ="row",
                  cexRow=1.0,
                  cexCol=1.0,
                  srtCol=0,
                  xlab="Day")
dev.off()

# Plot heatmap of candidates validated by qPCR along Pseudotime
pdf(paste0(format(Sys.Date() ,"%Y%m%d"), "Validation_PseudoTime_14July2018.pdf"),width=10,height=10)
gplots::heatmap.2(mydata[some_genes,], trace="none", Colv=FALSE, Rowv=FALSE,
                  # RowSideColors=mycol[mat_reordered[tfnames,1]],
                  col=colorRampPalette(c("whitesmoke", "black")),
                  density.info="none", # turns off density plot inside color legend
                  # main="The Network Controlling the Neural-to-Glial Fate Switch",
                  # lhei=c(0.5,20),
                  # lwid=c(0.2,10),
                  # margin=c(5.0,10.0),
                  # key=FALSE,
                  cexRow=1.0,
                  cexCol=1.0,
                  srtCol=0,
                  xlab="Pseudotime")
dev.off()
```

168
# calculate the correlation between pseudotime and expt data
mydata5 <- cbind(rowMeans(mydata[,1:4]), rowMeans(mydata[,5:8]), rowMeans(mydata[,9:12]), rowMeans(mydata[,13:16]), rowMeans(mydata[,17:20]))

for (Gene in some_genes){
    a[Gene] <- cor(expt.data.norm[Gene,, mydata5[Gene,, method="pearson"] # chose pearson over spearman because searman compares ranks instead of the raw data
}

# plot pseudotime vs expt data
pdf(paste0(format(Sys.Date(),"%Y%m%d","_ExptVsSC.pdf"), width=15, height=15))
par(mfrow=c(3,3))
for (Gene in some_genes){
    plot(expt.data.norm[Gene,, mydata5[Gene,, main=Gene, cex.main=3.0, xlab="qPCR", ylab="Single Cell", cex.lab=1.5, pch=16)
    abline(lm(mydata5[Gene, ~ expt.data.norm[Gene,]))
}
dev.off()}
Reference List


