

Expression and functional analysis of SOX3 in murine neurogenesis



Nicholas Rogers

B.Sc. (Biotechnology) (Honours)

A thesis submitted in fulfilment of the requirements for the degree of Doctor of
Philosophy

Discipline of Biochemistry

School of Molecular and Biomedical Science

The University of Adelaide, Australia

Thesis Amendments

Introduction:

Page 12- "Originally" not "Original"

Page 20- "severe" not "server"

Page 22- "affected" not "effected"

Introduction figure 10 (legend)- "Antigen peptide sequence used to generate" not "Binding site of"

Page 27- "C57BL/6" not "black6"

Page 28- "effector" not "effectors"

Page 29- "A recent" not "A Recent"

Materials and Methods:

Pages 37- "C57BL/6" not "c57Bl/6"

Page 41- "Wild type Male on the Sox3 null background" not "Male Sox3 null +/Y"

Acknowledgments:

"Patience" not "patients"

Contents

Abstract.....	5
Original publications.....	7
Declaration.....	8
Acknowledgements.....	9
Abbreviations:.....	10
Introduction.....	12
1.1 SoxB1 Subgroup members and early central nervous system development.....	12
1.1.1 Sox Family of transcription factors.....	12
1.1.2 Expression of SoxB1 Subgroup members during early central nervous system development.....	13
1.1.3 Early Telencephalic development.....	14
1.1.4 Early embryonic development and the generation of neurogenic progenitors.....	15
1.1.5 Function of SoxB1 genes in early neurogenesis.....	18
1.2 SOX3 and adult neurogenesis.....	21
1.2.1 Adult neurogenesis.....	21
1.2.2 SoxB1 subgroup members and neural progenitors in the adult mouse brain.....	22
1.2.3 Summary of Sox3 expression:.....	23
1.3 SOX3/Sox3 function and CNS development.....	24
1.3.1 The role of SOX3 in human diseases.....	24
1.3.2 Polyalanine expansions within the SOX3 open reading frame.....	24
1.3.3 SOX3 deletions and Human disease.....	25
1.3.4 Sox3 function in murine CNS development.....	26
1.3.5 Sox3 and Gastrulation.....	27
1.3.6 Sox3 and Septo-optic dysplasia.....	27

1.3.7 Sox3 and morphogenic signalling within the developing spinal cord	29
1.4 SOX3 DNA binding and functional effects.....	30
1.4.1 SOX3 binding in neural progenitor cells	30
1.4.2 SOX3 DNA binding and partner factors	31
1.5 Studying Sox3/SOX3 within a neural progenitor population	32
1.5.1 In vivo derived neuroprogenitors	32
1.5.2 Neural progenitor differentiation from embryonic stem cells	33
1.5.3 Summary of Sox3 function:.....	34
Aims	35
Aim1: Study SOX3 expression in the developing and developed CNS	35
Aim 2: Elucidate Sox3 function in central nervous system development:	35
Methods	36
Results	43
Result Chapter 1	43
Expression of the murine transcription factor SOX3 during embryonic and adult neurogenesis	43
Results Chapter 2.....	53
Mechanistic Insight into the Pathology of Polyalanine Expansion Disorders Revealed by a Mouse Model for X Linked Hypopituitarism	53
<i>Dbx1</i> is a direct target of SOX3 in the spinal cord	63
Discussion.....	81
Sox3 expression throughout dorsal telencephalic development.....	81
SOX3 expression within the adult neurogenic niches	82
Dynamic expression of SOX3 in the developing ventral diencephalon	83
SOX3 expression outside of the neurogenic niches within the adult brain	84
Generation of Sox3 null embryonic stem cells	85
Sox3 null chimera generation and implications for Sox3 disease models	85
SOX3 expression in cultured neural progenitor differentiation	86

SOX3 function within neural progenitors.....	87
Sox3 putative targets and potential function.....	88
<i>Dbx1</i> as a downstream target of <i>Sox3</i> function in the developing spinal cord	90
Recovery of DBX1 expression within <i>Sox3</i> null embryos.....	91
DBX1 as a direct target of SOX3.....	92
Future Directions	94
<i>Sox3</i> expression and role in adult neurogenic zones	94
SOX3 expression in non neurogenic adult brain zones.....	95
Comprehensive SOXB1 expression profiling	95
Functional redundancy between <i>SoxB1</i> genes	96
<i>Sox3</i> function and D-V patterning within the developing spinal cord.....	97
SOX3 Binding and function within spinal cord development.....	98
SOX3 binding sites and functional effects.....	99
<i>Dbx1</i> and hypothalamic development.....	100
Concluding remarks	102
Bibliography	103
Appendix 1:	110
Appendix 2:	111
Appendix 3:	112

Abstract

The Sox (SRY-related HMG box) family of proteins are transcription factors. There are, in total, 30 different genes in the Sox family. Each Sox protein contains a HMG box (high-mobility-group) which functions as a DNA binding domain. The HMG box is highly conserved (>50% identity) throughout the entire Sox family. *Sox3* belongs to the SoxB1 subgroup.

SOX3 has been associated with human CNS related disorders. Duplication and mutations of *SOX3* have been identified in patients with X-linked hypopituitarism (XH). Afflicted XH patients suffer from varying levels of mental retardation and pituitary hormone deficiencies which can lead to short stature.

Previous studies have shown that *Sox3* is expressed in nascent neuroprogenitor cells and is functionally required in mammals for development of the dorsal telencephalon and hypothalamus. Using a SOX3-specific antibody, data within my thesis shows that murine SOX3 expression is maintained throughout telencephalic neurogenesis and is restricted to progenitor cells with neuroepithelial and radial glial morphologies. In addition, characterisation of SOX3 expression within the adult neurogenic regions indicates that it is a lifelong marker of neuroprogenitor cells.

In contrast to the telencephalon, *Sox3* expression within the developing hypothalamus is up-regulated in developing neurons and is maintained in a subset of differentiated hypothalamic cells through to adulthood.

In addition, using genome wide expression analysis examining a *Sox3* null neural progenitor population, I identified a number of putative *Sox3* targets. The data identified *Dbx1* as a robust *Sox3* target with *Dbx1* down-regulation, at both the

mRNA and Protein level, within Sox3 null mice at early stages of CNS development. I also independently confirm a number of SOX3 binding sites surrounding *Dbx1*, with one site showing clear enrichment *in vivo*. In addition, correlation between these putative targets and that of a previously published SOX3-ChIP data set show a clear enrichment for SOXB1 binding sites near the mis-regulated genes suggesting they are direct targets of *Sox3*.

Taken together, data presented within my thesis identifies new regions of *Sox3* expression and putative Sox3 targets. This data helps advance our knowledge of *Sox3* regulation and function within CNS development.

Original publications

This thesis is based on the following publications:

Results Chapters 1 (Rogers et al, 2013):

Rogers, N., Cheah, P. S., Szarek, E., Banerjee, K., Schwartz, J. and Thomas, P. (2013). Expression of the murine transcription factor SOX3 during embryonic and adult neurogenesis. *Gene Expr Patterns* **13**, 240-8

Results Chapter 2 (Hughes et al, 2013):

Hughes, J., Piltz, S., **Rogers, N.**, McAninch, D., Rowley, L. and Thomas, P. (2013). Mechanistic insight into the pathology of polyalanine expansion disorders revealed by a mouse model for X linked hypopituitarism. *PLoS Genet* **9**, e1003290.

Results Chapter 2 (Rogers et al, Manuscript):

Rogers, N., McAninch, D. And Thomas P. Dbx1 is a direct target of Sox3 in the spinal cord. (Manuscript)

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name and to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time

Nicholas Rogers

November 2013

Acknowledgements

First, I would like to thank my supervisor Paul Thomas for giving me the opportunity to do a PhD within his laboratory. I would also like to thank him for his advice, support and patients throughout its course.

I am also thankful for the scientific expertises and opinions provided by members of the Thomas laboratory. To James, who helped me with general protocols including ES cell culture and Southern blots. Sandie Piltz, who helped me so much with my mouse work and was an invaluable source of knowledge. Dale, thanks for the help with ChIP experiments, statistics and bioinformatics, you have found your calling. To Pike-see, thank you for your help with histology, central nervous system anatomy and general expertise for our paper together.

Next I would like to thank everybody within the Thomas lab (Eva, Edwina, James, Adi, Dale, Sandie, Pike-see Daniel, Bryan, Shaun and Kakoli) for their expertises and wonderful insights, You made my PhD an enjoyable experience.

In addition, I would also like to thank the transient population of people who have enjoyed social evenings within the discipline of biochemistry. Without you all, I wouldn't have had so much fun on Friday nights, and probably had more productive Saturday mornings!

I would also like to thank my mother and father for their love and support throughout all of my years of study. Without your encouragement none of this would have been possible. To my brother Stuart and his wife Claire, thanks for your unconditional support.

Finally Kristie, thank you for being supporting, patient, tolerant, helpful and awesome throughout my PhD. I could not have done this without you.

Abbreviations:

3V	Third ventricle
ANOVA	Analysis of variance
ARC	Arcuate nucleus
cDNA	complementary DNA
ChIP	Chromatin immunoprecipitation
CNS	Central nervous system
DNA	Deoxyribonucleic acid
dpc	Days post coitum
ES	Embryonic stem
HMG	High mobility group
H&E	Haematoxylin and eosin staining
IHC	Immunohistochemistry
Kb	Kilobase(s)
LOF	Loss of function
LV	Lateral ventricle
ME	Median eminence
MPOA	Medial preoptic area
Ne	Neurohypophysis
NP	Neural progenitors
P	Postnatal
PCR	Polymerase Chain Reaction
PL	Pial Layer
qRT-PCR	quantitative Real Time PCR
RIN	RNA integrity number
RNA	Ribonucleic acid
SGZ	Sub granular Zone
SOX	Sry-related HMG Box

SRY	Sex determining region Y
SVZ	Sub ventricular zone
TV	Telencephalic vesicles
VZ	Ventricular Zone
VD	Ventral Diencephalon
XH	X-linked hypopituitarism

Introduction

1.1 SoxB1 Subgroup members and early central nervous system development

1.1.1 Sox Family of transcription factors

The Sox family of transcription factors have been shown to control cell fate throughout normal development. The Sox genes encode 20 proteins in mammals, originally identified by the sequence homology with founding members SRY (Sex determining region Y) with each SOX protein sharing over 50% homology over the HMG (High-mobility group) DNA binding domain (Pevny and Lovell-Badge, 1997; Wright et al., 1993). Each SOX protein is subdivided into a smaller group as defined in Bowles et al., 2000. SOX proteins within a subgroup share a degree of amino acid homology outside of the HMG domain, however little homology outside of the HMG domain is shared between SOX proteins of different subgroups (Bowles et al., 2000). The Sox family subgroup that shares the highest sequence homology with SRY is the SoxB1 subgroup.

1.1.2 Expression of SoxB1 Subgroup members during early central nervous system development

The SoxB1 Subgroup of the Sox family is comprised of three genes: *Sox1*, *2* and *3*. The three SoxB1 subgroup members share a high degree (approximately 90%) of sequence homology over the 79 amino acid HMG domain (Collignon et al., 1996; Kamachi et al., 2000) Figure 1).

SoxB1 expression begins at an early stage in mouse embryonic development where *Sox2* is detected at the morula stage (2.5 days post coitum (dpc)) and is the first of the SoxB1 members expressed (Avilion et al., 2003; Bowles et al., 2000). *Sox3* expression appears in the epiblast and extra-embryonic ectoderm at 6.5 dpc in the mouse embryo and is considered one of the earliest neural markers in vertebrates (Brunelli et al., 2003; Wood and Episkopou, 1999). *Sox1* is first expressed in early ectodermal cells of the developing central nervous system (CNS) at 7.5 dpc (Pevny et al., 1998).

SoxB1 subgroup members are widely expressed in the developing CNS (Collignon et al., 1996; Pevny and Lovell-Badge, 1997; Uchikawa et al., 1999; Wood and Episkopou, 1999; Woods et al., 2005). A comprehensive analysis of SoxB1 gene expression in mouse embryos from 6.5-9dpc reported that all three transcripts are present during early CNS development (Wood and Episkopou, 1999). The data also indicated that the three SoxB1 subgroup members have highly overlapping expression patterns. It was observed that at 6.5 dpc both *Sox2* and *Sox3* have similar expression patterns. *Sox1* and *Sox3* expression becomes restricted to a subset of *Sox2* expressing cells from around 8 dpc (Figure 2).

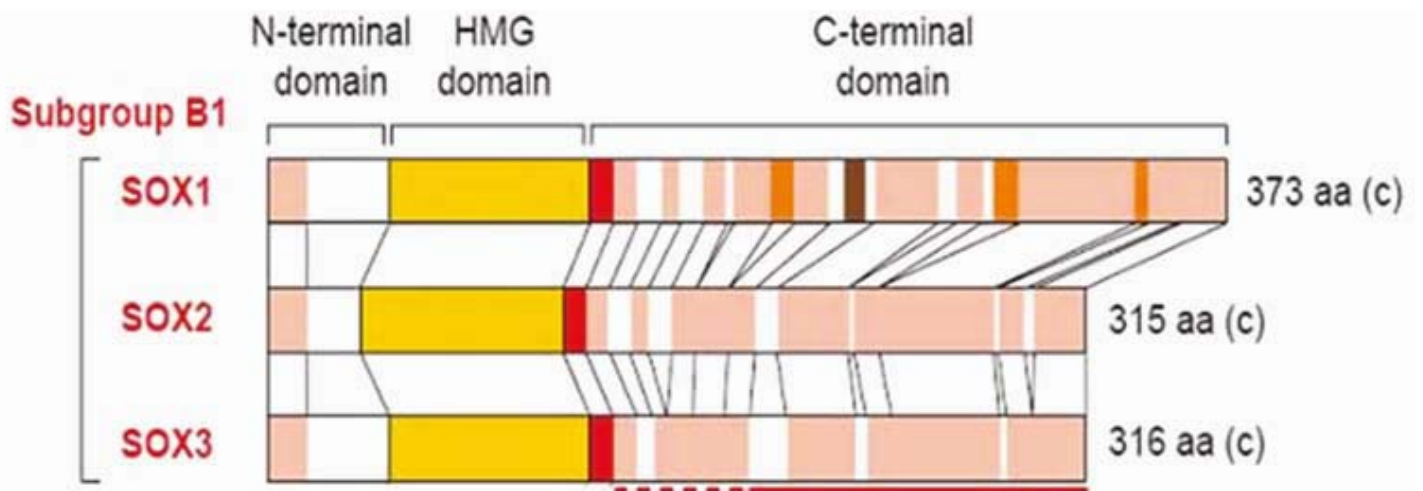


Figure 1: Sequence Homology between SoxB1 subgroup members. The HMG box, coloured yellow, has the highest degree of sequence homology between SoxB1 members. Other regions of high sequence homology indicated in pink. (Kamachi et al., 2000)

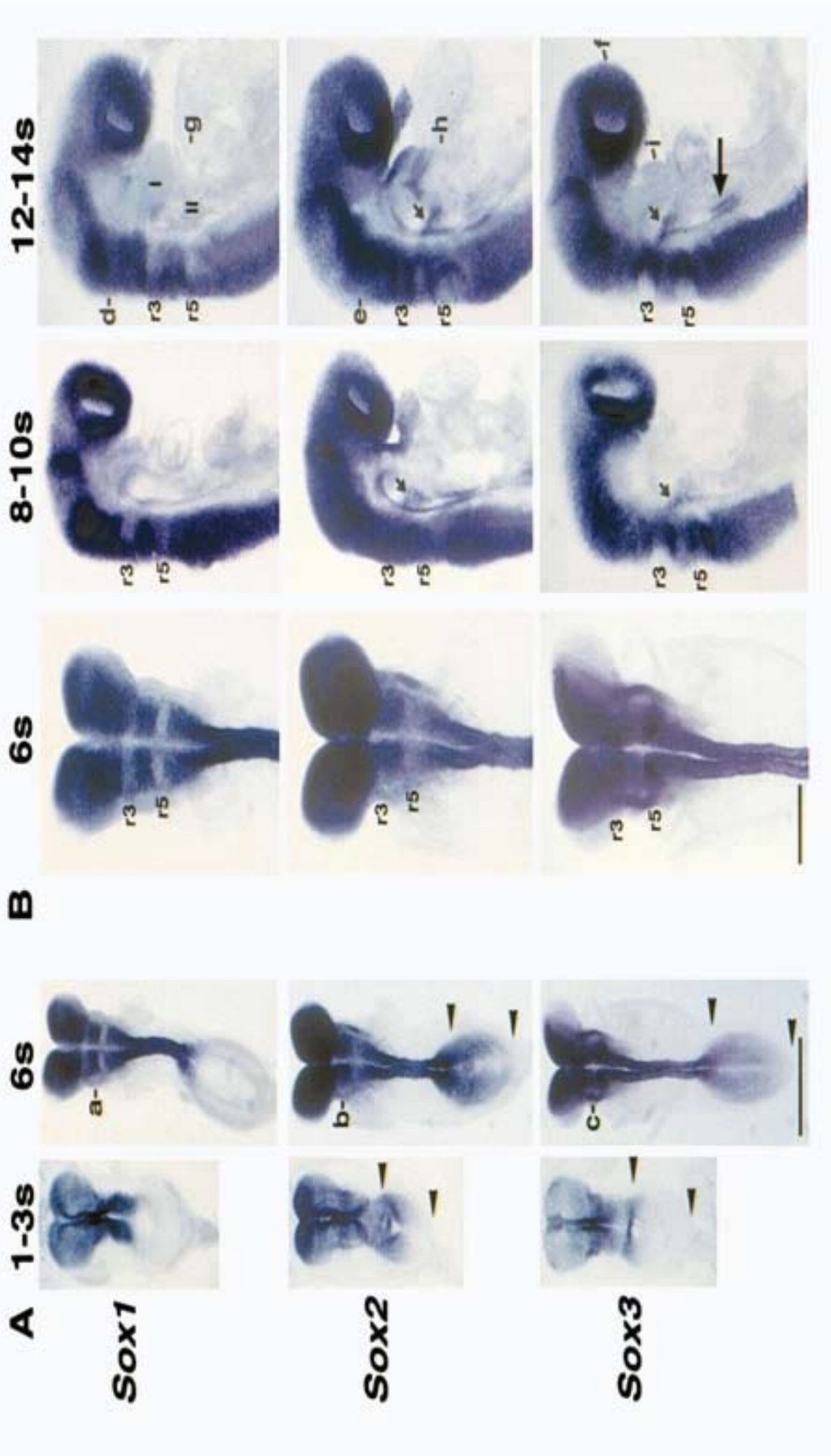


Figure 2: Expression of SoxB1 subgroup members between 8-9 dpc. 1-6 somites dorsal view. 8-14 somites dorsal view. 8-14 somites lateral view. 8-14 somites lateral view. 8-14 somites lateral view. Note overlapping expression patterns. S:Somites (Wood and Episkopou, 1999).

Sox3 nucleotide/amino acid sequence is highly conserved throughout mammalian species. The coding sequence for Sox3 is composed of a single-exon found on the X chromosome. At the onset of early gastrulation Sox3 found in the epiblast is restricted to anterior ectoderm which gives rise to the central nervous system (CNS). The extra-embryonic expression becomes restricted to the chorion. Sox3 expression in the chorion is downregulated by the late streak stage (approx 7.5 dpc). Sox3 expression is maintained in the developing CNS from 7.5-9.5 dpc, and is expressed within the ventral diencephalon at 12.5 dpc in the region of the presumptive hypothalamus and infundibulum (Solomon et al., 2004). It has also been shown to be expressed within the forebrain of 9.5 dpc embryos in zones such as the developing telencephalic vesicles (Collignon et al., 1996); Figure 2) indicating that Sox3 is expressed within progenitor cell regions (see below) that are likely to undergo neurogenesis as the CNS develops. Unpublished data from the Thomas lab indicates that SOX3/Sox3 is expressed within the developing telencephalic vesicles between 10.5-14.5 dpc (Figure 3 & 4) in regions which develop to form the hippocampus, neo cortex and sub ventricular zone (Figure 3). However Sox3 expression during neurogenesis has not been comprehensively studied and is critical to elucidating the role of Sox3 in CNS development.

1.1.3 Early Telencephalic development

Analysis of Sox3 expression in early CNS development shows that Sox3 is initially expressed within the neural progenitor population (Wood and Episkopou, 1999). However, no studies have looked specifically at Sox3/SOX3 at the height of

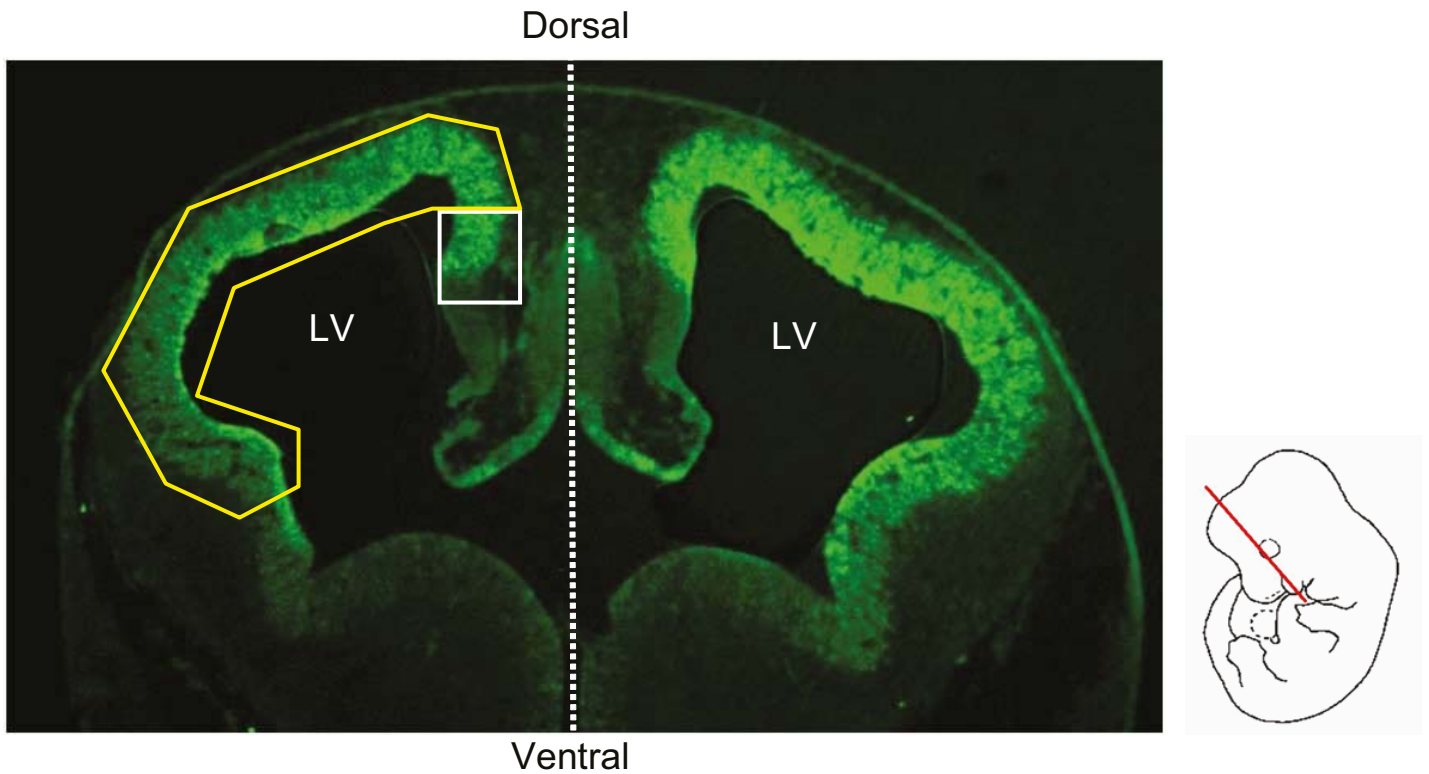


Figure 3: SOX3 is expressed in the ventricular zone of the developing dorsal telencephalon. SOX3 immunofluorescence of the telencephalon (coronal section) of E12.5 embryo. SOX3 (Green) at the ventricular zone of dorsal telencephalon. The white dotted line is the dorsal-ventral midline, separating the left and right hemispheres. Region framed by yellow line: presumptive cerebral cortex; region framed by white line: presumptive hippocampus. The relative position of the section on mouse is illustrated at the right bottom corner. LV: lateral ventricles. (K.Lee Rubicon 2007, N. Rogers, unpublished data)

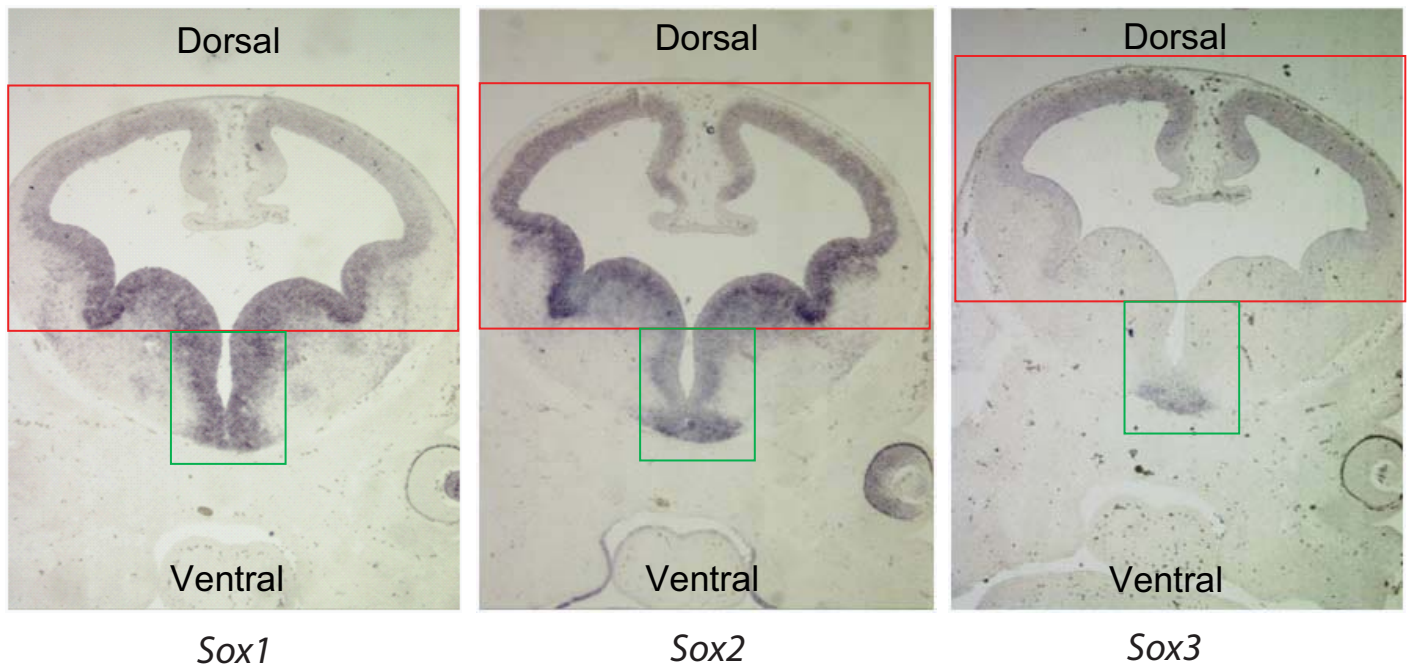


Figure 4: *Sox1*, *Sox2*, *Sox3* expressions partially overlap in developing telencephalon and diencephalon. In situ hybridisation of *Sox1*, *Sox2* and *Sox3* expressions in E12.5 embryos (coronal sections). *Sox1*: high expression at VZ of ventral telencephalon and diencephalon, low expression at VZ of dorsal telencephalon; *Sox2*: High expression at VZ of developing telencephalon and diencephalon; *Sox3*: expression at dorsal telencephalon and ventral telencephalon. Note: the telencephalon is framed by red lines; the diencephalon is framed by green lines. (Thomas Lab, unpublished data)

neurogenesis or within telencephalic development, a region/time with a large population of neural progenitor expansion.

The telencephalon is derived from neuroectoderm located at the anterior end of the developing neural plate (Figure 5). The telencephalon is anterior and dorsal to the diencephalon, and together make up the prosencephalon or forebrain of the developing CNS. Patterning of the telencephalic dorsal midline roof plate, results in slower expansion of cells inhabiting the midline, giving rise to two separate vesicles (Left & Right) during neural tube closure (E9) and telencephalic expansion. The dorsal telencephalic vesicles give rise to the neo cortex and hippocampus (Monuki and Walsh, 2001; Wilson and Rubenstein, 2000; Zaki et al., 2003). The telencephalon gives rise to structures known to have abnormal development within *Sox3* null mice (Discussed in more detail in Introduction 1.3.4). However no studies have investigated SOX3 expression throughout the development of the neo cortex or hippocampus.

1.1.4 Early embryonic development and the generation of neurogenic progenitors

Although neurogenesis occurs throughout the early developing CNS, one focus of my PhD studies has been to characterise SOX3 expression during dorsal telencephalic development in particular within neural progenitors. Embryonic neurogenesis within the developing murine telencephalon begins at around 9th embryonic day (E) in embryogenesis. Neurons are generated, on mass, through until E18 (Figure 6). The peak of neuronal production occurs at around E14. After E18 gliogenesis begins producing first astrocytes followed by oligodendrocytes (Figure 6).

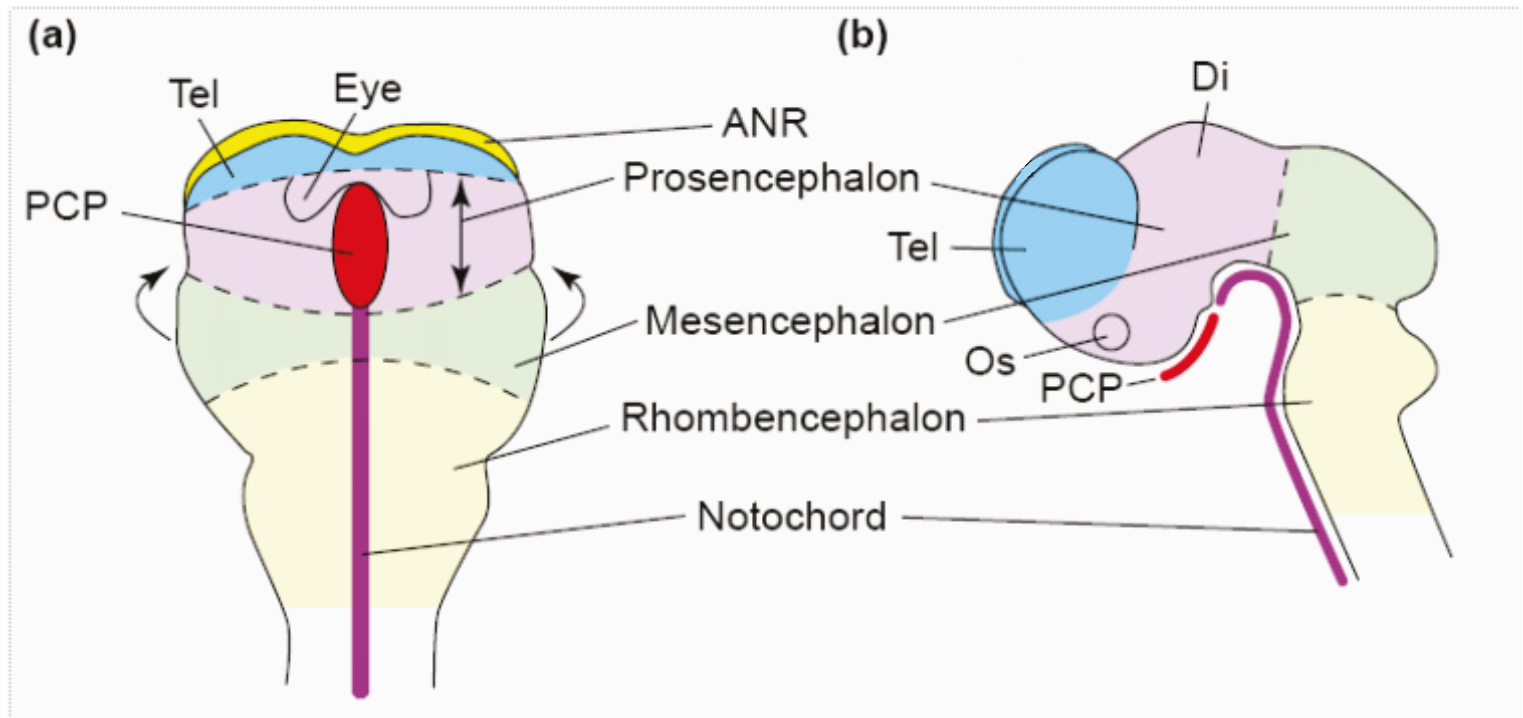


Figure 5: Early murine telencephalic development (a) Folds of the early anterior neural plate indicated by arrows (~E8.5) (b) Side view of the developing murine brain at E10.5 clearly showing the developing telencephalic vesicles as a part of the prosencephalon. Tel: Telencephalon, Di: Diencephalon, PCP: prechordal Plate, ANR Anterior neural ridge (Zaki et al., 2003)

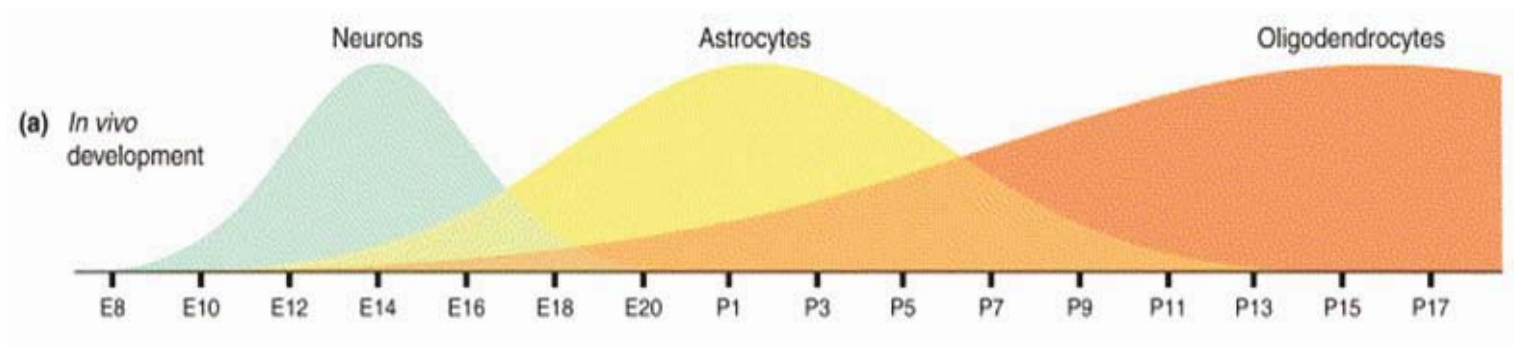


Figure 6: Neurogenesis within the developing telencephalon. (a) In vivo timeline of neurogenesis within the telencephalon (~E9-E18). Clearly showing that gliogenesis primarily occurs after E18. (Sauvageot and Stiles, 2002)

However a small amount of neurogenesis occurs postnatally, within the adult mouse brain (see below). All neurons born at around E14 are derived from neural progenitor cells (NP).

Within the developing telencephalon there are three sub types of neural precursor cells that can undergo the neurogenic pathway and produce new neurons. They are; neuroepithelial cells (NE), radial glial cells (RG) and basal progenitors (BP) (Bulfone et al., 1999; Chanas-Sacre et al., 2000; Gotz and Huttner, 2005; Guillemot, 2005; Hodge et al., 2008; Zimmer et al., 2004) (Figure 5 & 7).

Neuroepithelial cells within the developing telencephalon first undergo a round of division leading to neurogenesis. The division occurs symmetrically or asymmetrically, producing either two identical neuroepithelial cells or a radial glial cell and basal progenitor cell. In the case of very early neurogenesis neurons can be produced directly from these early divisions.

During early neurogenesis it has been hypothesised that radial glial cells are the main precursor cells for neurogenesis within the dorsal telencephalon (Anthony et al., 2004; Malatesta et al., 2003). Radial glia are generated from neuroepithelial asymmetric divisions and in contrast to neuroepithelial cells, no longer express proteins that form tight junctions. Radial glial cells also express markers commonly associated with active astrocytes such as brain-lipid-binding protein (BLBP) and Glial fibrillary acidic protein GFAP (although not in rodents)(Hartfuss et al., 2001; Kriegstein and Gotz, 2003). Radial glial cells can give rise to both neurons and basal progenitors.

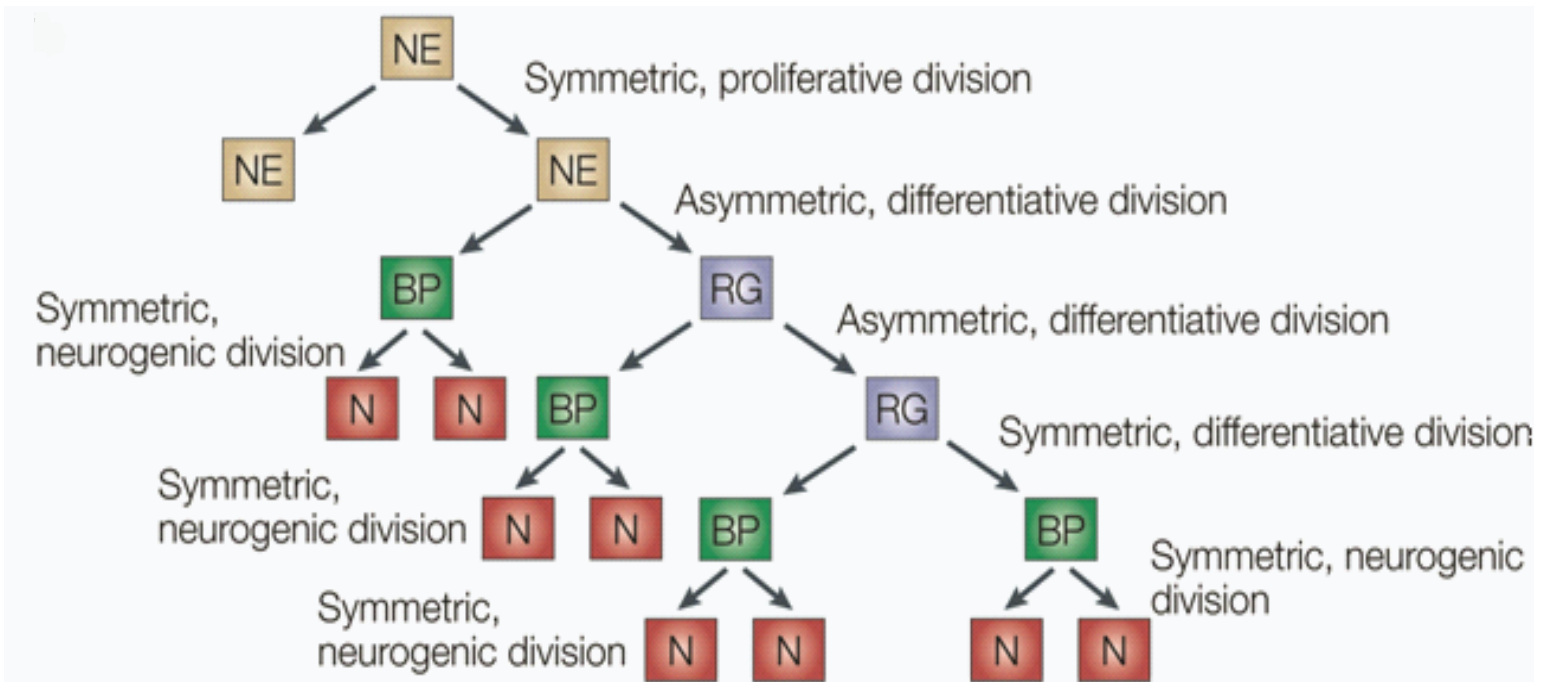


Figure 7: Overview of neurogenic division that occur in the developing telencephalon. Neuroepithelial cells (NE) give rise to both Radial glial (RG) and Basal Progenitor (BP) cells which in turn give rise to neurons (N) early in neurogenesis. (Gotz and Huttner, 2005)

Neuroepithelial and radial glial cells both have apical-basal polarity. This polarity is thought to determine the type of division that occurs within the cells (Gotz and Huttner, 2005; Huttner and Brand, 1997), Figure 8). These two progenitor types also undergo interkinetic nuclear migration, the process whereby the cell's nucleus migrates along the apical-basal axis through its cell cycle. This gives the impression of a single cell layer being multiple layers thick. As such, the location of the neural progenitors at the early stages of neurogenesis is identified in the ventricular zone (Figure 8). Most of the neural progenitors surrounding the telencephalic vesicles are located in the ventral zone, defined by those cells expressing SOX3 in figure 4.

Basal progenitors do not have the same apical-basal polarity as neuroepithelial and radial glial cells. Basal progenitors undergo symmetric divisions producing either two daughter basal progenitor cells or two new born neurons (Haubensak et al., 2004). Due to this type of division and their location, at the radial side of the ventral zone, they are thought to be cells responsible for the amplification of the number of neurons produced and in addition give rise to the sub-ventricular zone (see below) of the adult murine brain.

Although *Sox3* expression has been observed within early neural progenitor cells during the initiation of neurogenesis no data on *Sox3*/SOX3 expression over the course of neurogenesis has been performed (Wood and Episkopou, 1999).

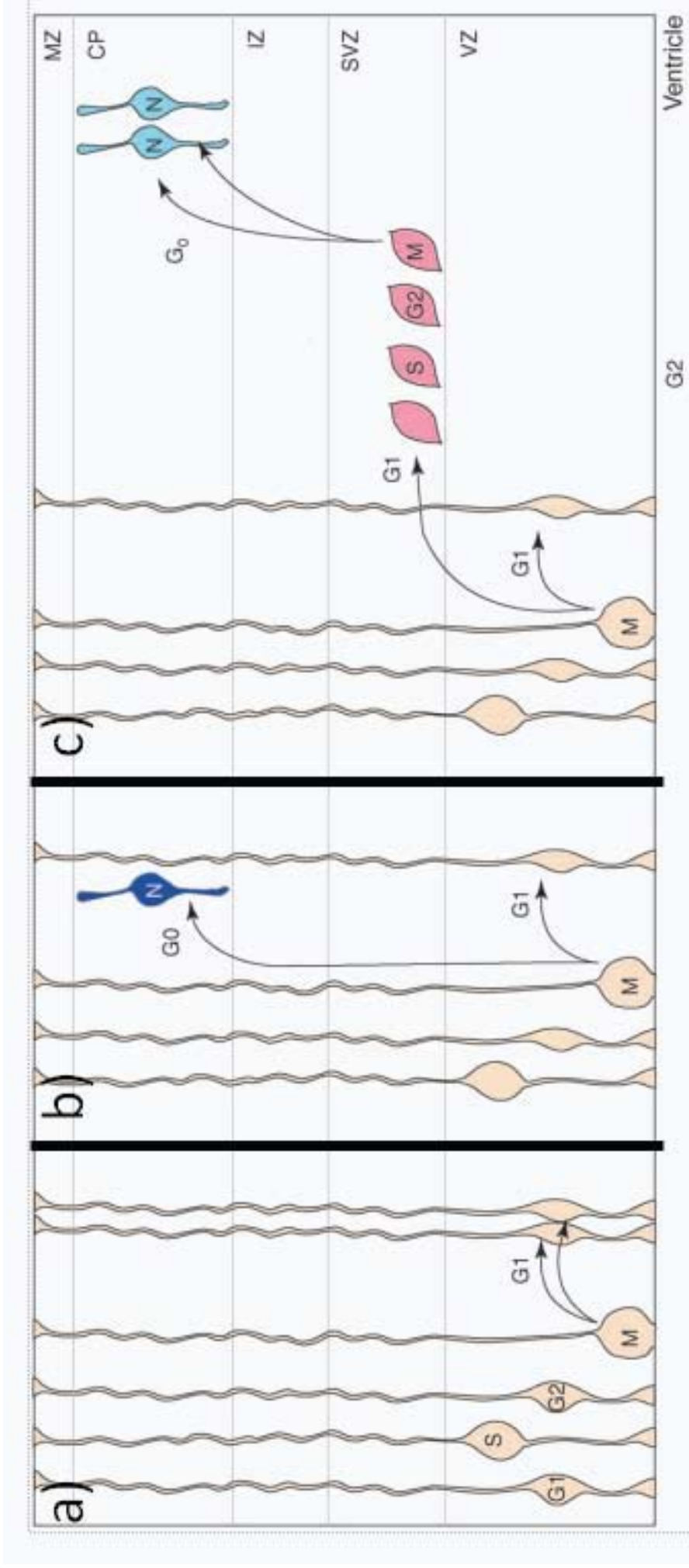


Figure 8: Asymmetric/symmetric division of neural progenitors (a) Symmetric division producing two NP cells. (b) Asymmetric division producing a new born neuron and a new NP cell (c) Asymmetric division, resulting in a NP cell and a BP cell that can then undergo symmetric division. VZ: ventricular zone, SVZ: subventricular zone (Guillemot et al 2005).

1.1.5 Function of SoxB1 genes in early neurogenesis

SoxB1 genes have been shown to have restricted expression within neural progenitors during the early stages of embryonic development. However functional studies are required to elucidate the role that the SOXB1 proteins play throughout neurogenesis.

One study focusing on the developing chick spinal cord provided initial data on SoxB1 function (Bylund et al., 2003). *In ovo* electroporation was used to over express *Sox1-3* in the spinal cords of stage 10 chick embryos. The authors analysed the expression of markers for actively dividing cells (Neural progenitors (NP)) through incorporation of BrdU and neuronal differentiation (NeuN, Tuj1). No change in the neural progenitor cells was observed, however a decrease in expression of the markers for differentiated neurons was observed within Sox1-3 over expressing tissue. The results suggest that the role of the SoxB1 members in the developing CNS is one of neural progenitor cell maintenance and the study concluded that over-expression of Sox1-3 inhibits vertebrate neurogenesis (Bylund et al., 2003).

An additional finding within the 2003 Bylund paper looked at the ability of Sox3 to activate or repress its target genes (Bylund et al., 2003). When the C-terminal end of SOX3 was replaced with an obligator activator domain (viral protein VP16), leaving the HMG box intact, the over expression mimicked the effect of SOX3 over expression. Replacing the C-terminal end of the SOX3 open reading frame with a known repressor domain (*D.melanogaster* engrailed protein) resulted in cells over expressing the construct, to be terminally differentiated within 45hours. The data

suggested that SOX3 acts as a transcriptional activator and that repression of SOX3 target genes results in differentiation of neural progenitors.

However it is worth noting that Sox3 repression of a target gene has been shown to be required for endogenous gastrulation development (Acloque et al., 2011), Introduction 1.3.5). In addition, SOX2 has been shown to act as both a repressor (Botquin et al., 1998) and an activator (Ambrosetti et al., 1997) suggesting that the trans-repressive/trans-activation function of a SoxB1 gene may be context dependent and therefore dictate the transcription factors role in each cell type.

A recent publication looking at the SoxB1 members and their regulation in neural progenitor differentiation, found that the SoxB1 members promote the progenitor state and inhibit differentiation independent of the notch signalling pathway (Holmberg et al., 2008). This was achieved through *in ovo* electroporation of stage 10 chick neural tubes it was found that notch repression of neural differentiation required SOX3, although the SoxB1 mechanisms for neural progenitor maintenance is independent of the notch pathway.

Further evidence supporting a critical role for Sox3 in the CNS, stems from studies involving targeted mutagenesis of Sox3. Studies on Sox3 null mice, indicated that Sox3 is required for formation of the hypothalamic-pituitary axis (Discussed in detail below in introduction section 1.3.4)(Rizzoti et al., 2004) . Postnatal Sox3 null mice

also suffer from dysgenesis of the corpus callosum and hippocampus, which are both structures derived from the telencephalic vesicles.

The overlapping expression patterns and high degree of sequence homology between the three proteins suggests a possibility that some functional redundancy between the proteins might occur. A study on zebrafish embryo development, looking at compound quadruple SoxB1 knockout embryos found that individual knock downs of *Sox2/3* resulted in no gross abnormalities within the developing embryos, however compound quadruple null embryos had the server phenotypes thus providing evidence for functional redundancy (Okuda et al., 2010). This implies that if one of the members of SoxB1 is restricted another member could functionally compensate, thus making it difficult to study phenotypes of individual proteins.

In situ hybridisation data on *SoxB1* has been comprehensively analysed between embryonic stages 2.5dpc to 9.5dpc and demonstrates *SoxB1* expression within the early neural progenitor cells. When this data is considered with the functional *SoxB1* over expression data, it suggests that *SoxB1* genes are required for neural progenitor maintenance. In addition, these data advocate a role for *Sox3* in endogenous neurogenesis however *Sox3* has not been comprehensively studied either within the murine models (in particular forebrain development) or within the adult neurogenic niches.

1.2 SOX3 and adult neurogenesis

1.2.1 Adult neurogenesis

There are two known areas of neurogenesis within the adult mammalian brain: the sub-ventricular zone (SVZ) and the sub-granular zone (SGZ) ((Zhao et al., 2008) Figure 9). The SVZ is located along the edges of the developed lateral ventricles. The SGZ is a layer of cells located within the dentate gyrus of the hippocampus. Both of these neurogenic niches harbour neural progenitor cells that can generate new neurons.

Within the neurogenic niches there appears to be at least two distinct populations of neurogenic progenitor cells within each region (Fukuda et al., 2003; Morshead et al., 1994; Weiss et al., 1996). The two populations will be referred to as the quiescent neural progenitors (qNP) and the amplifying neural progenitors (aNP) (Fukuda et al., 2003; Suh et al., 2007). The qNP cells are slow dividing and few in number, whereas the aNP cells proliferate rapidly producing numerous new born neurons. There is no definitive evidence to determine which of these cell types is a true “stem” cell.

However, it is hypothesised that the qNP population gives rise to the aNP population, as ablation of qNP cells (GFAP expressing/dividing cells) stop the generation of new neurons within both the SVZ and SGZ (Garcia et al., 2004).

Neuroblasts born in the SVZ migrate along the rostral migratory stream to become interneurons in the olfactory bulb. The new born neurons pass through a number of

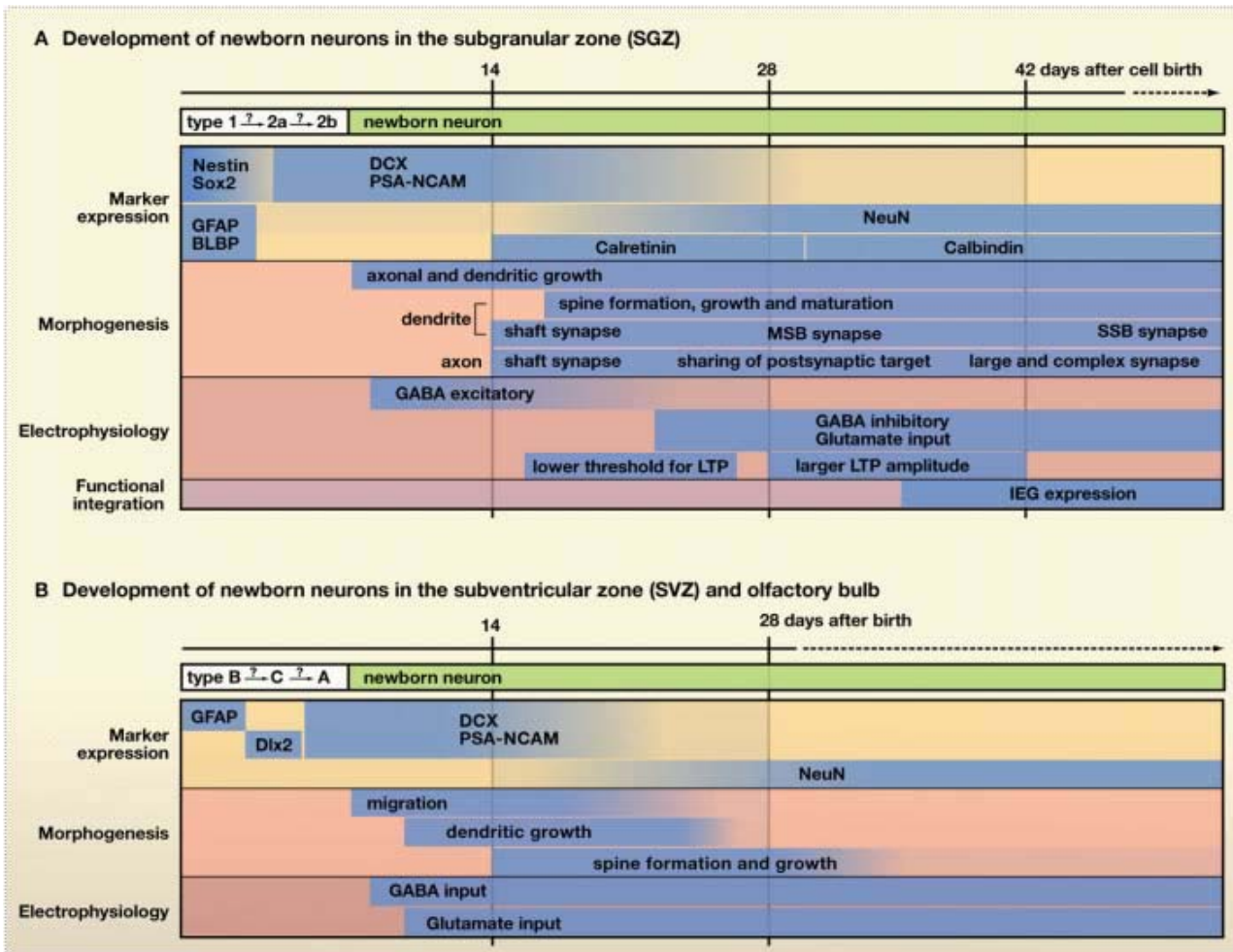
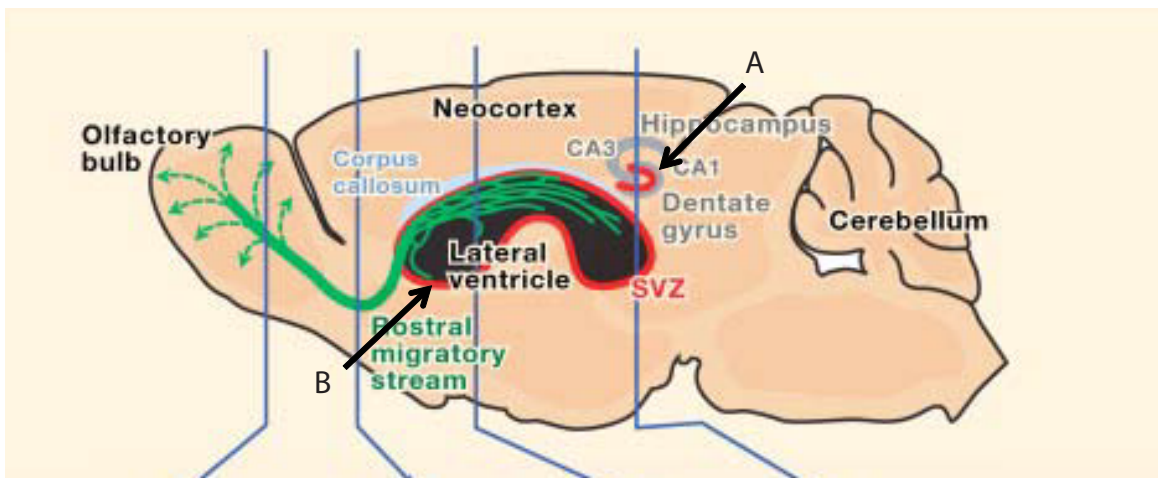


Figure 9: Location and timing of adult neurogenesis. A) SGZ, Type1 cells are thought to be the qNP population, 2a cells are thought to be the amplifying population of neural progenitors (aNP), those responsible for expansion of the growing population, 2b cells are thought to be migrating neurons that are partially differentiated. B. SVZ, type B cells are thought to be the qNP population of SVZ(although not exactly the same as the SGZ population), type C cells are the aNP and type A cells are partially differentiated cells that are migrating (Zhao et al., 2008)

different maturation stages before becoming integrated mature neurons (Lois and Alvarez-Buylla, 1994; Petreanu and Alvarez-Buylla, 2002). This is similar to the SGZ where new granular cells are born and mature over a longer period of time (around 7 weeks) (Esposito et al., 2005; Zhao et al., 2006) (Figure 9). However, these two distinct zones of neurogenic cells have individual molecular identities which have been demonstrated with both marker studies and neurosphere assays ((Bonaguidi et al., 2008; Craig et al., 1996; Reynolds and Weiss, 1992; Walker et al., 2008)Figure 9).

The exact role of the neural progenitors and newly born neurons within the adult brain is largely unknown although behavioural studies have indicated that neurogenesis in the SGZ may play a role in spatial learning and memory (Farmer et al., 2004; van Praag et al., 1999). As a result any effect that may reduce the progenitors' ability to regenerate, particularly in the hippocampus, would most likely have a profound effect on an individual. Moreover neurogenesis within the adult mouse can be effected strongly by environment stimuli and genetic background (Kempermann and Gage, 2002; Zhao et al., 2008).

1.2.2 SoxB1 subgroup members and neural progenitors in the adult mouse brain

SOX3 expression in the adult CNS has been comprehensively analysed in a paper published in 2006 by Wang and colleagues (Wang et al., 2006). The authors focused on the neurogenic regions (SVZ, SGZ) of the adult mouse brain. SOX3 expression was compared with markers for differentiated neurons (Tuj1, NeuN) and NP cells

(GFAP) using immunohistochemistry. The authors also used BrdU incorporation to identify NP cells. It was reported that BrdU/SOX3 overlap occurred within the SGZ. SOX3 was also found to be expressed within neurospheres derived from the SVZ and was down-regulated with differentiation.

The researchers used a SOX3 antibody that was raised against a C-terminal epitope of the *Xenopus Sox3* protein ((Wang et al., 2006), Figure 10). However the antibody has been tested within the Thomas laboratory on *Sox3* null tissue and is not SOX3 specific (Results Chapter 1:(Rogers et al., 2013)).

As a result data produced within the 2006 Wang paper needs to be re-examined and the expression patterns, particularly within the adult neurogenic niches need to be identified with a SOX3 specific antibody.

1.2.3 Summary of Sox3 expression:

Sox3 expression has been widely studied within early CNS development. SOX3 expression, however, has not been comprehensively studied across neurogenic stages of brain development. Furthermore, no definitive data showing SOX3 expression within the adult neurogenic niches has been performed. Data showing that *Sox3* plays a role in neural progenitor maintenance and differentiation prompts further studies the endogenous expression within these regions (Results Chapter 1: Rogers et al 2013).

MSOX1 P QHYQ-GAGAGVNGTVPLTHI

MSOX2 YQSGP-VPGTAKYGTLP LSHM

MSOX3 H QHYQ-GAGTAVNGTVPLTHI

XSOX3 - QHYQSAAGP GVNGTVPLTHI -Wang et al (2006) Sox3 Ab amino acid recognition

Figure 10: Binding site of the Wang et al (2006) SOX3 antibody and comparison between SoxB1 subgroup

1.3 SOX3/Sox3 function and CNS development

1.3.1 The role of SOX3 in human diseases

X-linked mental retardation affects 1/600 males (Laumonnier et al., 2002). X-linked hypopituitarism (XH) is a mental retardation syndrome that is associated with altered SOX3 dosage. SOX3 has been implicated as a component in the development of a certain type(s) of X-linked mental retardation (Lagerstrom-Fermer et al., 1997; Laumonnier et al., 2002; Solomon et al., 2002; Woods et al., 2005). A number of XH patients have duplications within their X-chromosome, the common feature within these duplications is the SOX3 gene. These data provided initial evidence that SOX3 function is important for normal central nervous system development.

1.3.2 Polyalanine expansions within the SOX3 open reading frame

The SOX3 open reading frame contains a number of polyalanine tracts outside of the HMG domain. The polyalanine tracts are highly conserved between mice and humans and expansion within these tracts has been associated with diseases in a number of different proteins (Hughes and Thomas, 2013).

Evidence for the importance of SOX3 function in CNS development came when a polyalanine expansion (+11 alanines) was identified within the human SOX3 open reading frame. The polyalanine track expansion can result in mental retardation and growth hormone deficiency (Laumonnier et al., 2002) suggesting that SOX3 is required for both CNS and pituitary development. Another polyalanine expansion

mutation (+7 alanines) was identified in three siblings with panhypopituitarism and pituitary defects (Woods et al., 2005). The +7 alanine expansion was shown to reduce transcriptional activity within an In vitro system suggesting a reduction in *SOX3* function.

All XH patients have growth hormone deficiency, in some cases panhypopituitarism, resulting in short stature if not treated with hormone replacement therapy. Mild mental retardation is not found in all patients with the *SOX3* abnormalities (Woods et al., 2005). However the polyalanine expansions appear to be playing a role in affecting Sox3 function. To truly elucidate the affect of these mutations, studies looking at the functional impact of polyalanine expansions on Sox3 need to be conducted (Result Chapter 2:(Hughes et al., 2013)).

1.3.3 *SOX3* deletions and Human disease

Human male patients with *SOX3* deletions have been identified (Helle et al., 2013; Stevanovic et al., 1993; Vencesla et al., 2007). A 6Mb deletion, including *SOX3*, was identified in a human patient who suffered from intellectual disability and haemophilia B (Stevanovic et al., 1993). Two more recent studies have identified smaller X chromosome deletions (approximately 2Mb) encompassing the *SOX3* gene. One of the *SOX3* null individuals has minor facial abnormalities however both patients were diagnosed to have varying levels of intellectual disability and developmental delay. (Helle et al., 2013; Vencesla et al., 2007). A greater knowledge of Sox3 function may one day help in the early identification of and the discovery of novel therapies to aid, afflicted individuals. These studies highlight the critical role of endogenous *SOX3*

function in normal CNS development. However few functional studies have been performed to elucidate the endogenous role that Sox3 plays in development.

1.3.4 Sox3 function in murine CNS development

To further investigate the function of Sox3 and its effect on the developing CNS a number of transgenic and *Sox3* null mouse lines have been constructed. The *Sox3* null mouse model most pertinent for the studies of the effects of Sox3 on neurogenesis, was originally generated by the Lovell-Badge laboratory (Rizzoti et al., 2004). Using homologous recombination the endogenous Sox3 open reading frame was completely replaced with a cassette containing GFP (Figure 11), thereby generating a Sox3 null allele.

Sox3 null male mice have variable phenotypes, the most severe of which is dwarfism, hypopituitarism and craniofacial defects (Rizzoti et al., 2004; Rizzoti and Lovell-Badge, 2007). Heterozygous females generally appear to be grossly normal although some have mild craniofacial defects and bifurcated Rathke's pouch.

Additional morphological defects that were originally identified in the *Sox3* null mice include dysgenesis of the corpus callosum and a failure of the dorsal hippocampal commissure to cross the midline (Figure 12). The corpus callosum and dorsal hippocampal commissure are both derived from the telencephalon. These phenotypes were incompletely penetrant, and background dependant with the

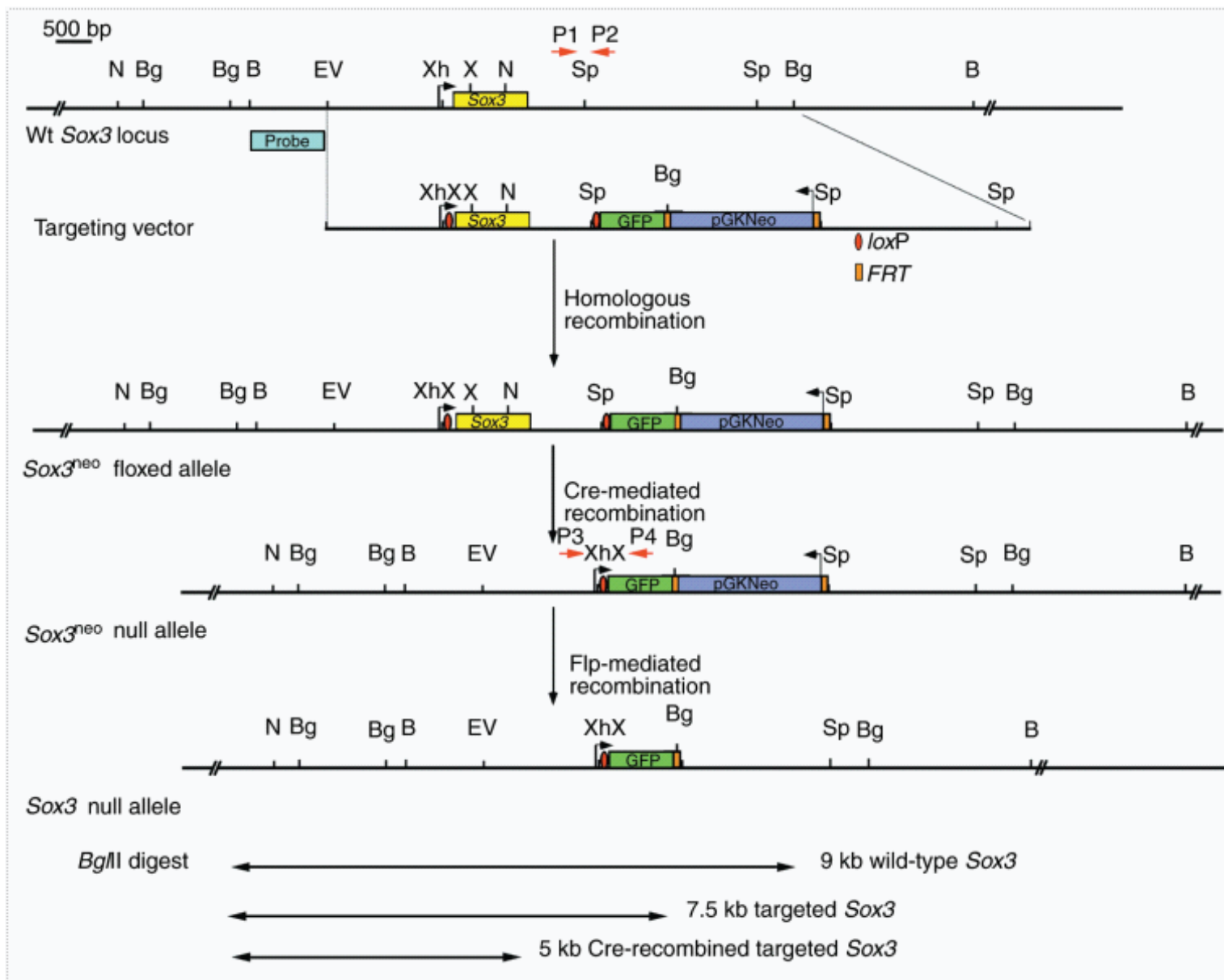
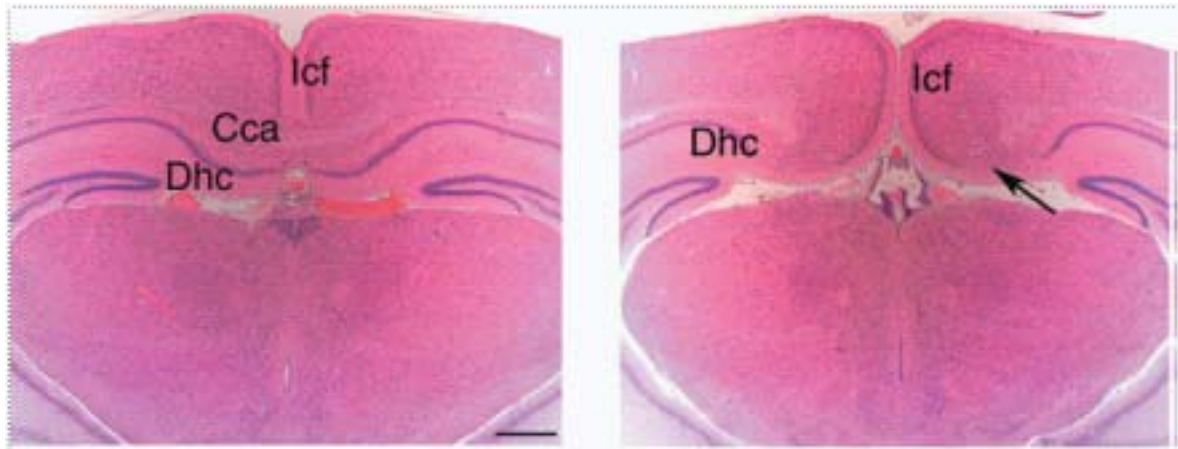


Figure 11: *Sox3* null mouse construct showing replacement of the SOX3 open reading frame with the GFP reporter construct. A small amount of the *Sox3* 3' regulatory sequence was deleted. (Rizzoti et al, 2004).

A)



B)



Figure 12: *Sox3* null phenotypes A) Left panel, Coronal section of a wild type mouse brain showing the dorsal hippocampal commissure and the corpus collosum crossing the mid-line. A) Right panel, *Sox3* null phenotype showing the dysgenesis of both the corpus collosum (Cca) and dosal hippocampal commissure (Dhc). B) Two *Sox3* null mice, one clearly showing dwarfism while the other appears phenotypical normal. (Rizzoti et al,2004)

abnormal dorsal hippocampal commissure and corpus callosum phenotypes not observed within *Sox3* null mice on a predominantly black6 background.

1.3.5 Sox3 and Gastrulation

Another critical role for Sox3 within early development was identified when generating the Sox3 null mice. Direct generation of Sox3 null chimeras, through injection of the 129sv/ev XY Sox3 null ES cells into a 3.5 dpc blastocysts resulted in an early embryonic lethal phenotype (Rizzoti et al., 2004). The phenotype was later found to be a defect in gastrulation. Sox3, at gastrulation is required to repress expression of *Snail1* within ectodermal lineages at the anterior pole of the incipient CNS (Acloque et al., 2011). The resultant repression of *Snail1* inhibits the ectodermal cells from undergoing an epithelial-to-mesenchymal transition, a transition that normally occurs on the posterior side of the gastrula where *Snail1* is expressed. Repression of *Snail1* by Sox3 is required for normal development and highlights the need for more functional studies into Sox3 function within early neural progenitor cells. In addition, the gastrulation phenotype also provides an assay of Sox3 function within the early CNS development (Result Chapter 2: Hughes et al 2013).

1.3.6 Sox3 and Septo-optic dysplasia

Septo-optic dysplasia (SOD) is a condition vaguely described by three optic nerve hypoplasia 1) midline neuroradiological abnormalities, like agenesis of corpus callosum, 2) pituitary hypoplasia and 3) hypopituitarism (Kelberman and Dattani, 2008). SOD phenotypes are similar to phenotypes observed within the *Sox3* null mice (Rizzoti et al., 2004). A recent study looked at SoxB1 genes and SOD and

identified *Shh* as a potential effectors (Zhao et al., 2012). Sonic hedgehog (Shh) is a secreted molecule that controls ventral neural patterning throughout development (Dessaud et al., 2007). Shh is important for normal hypothalamic development in chick embryos and after induction the hypothalamus becomes a second site of *Shh* expression (Dale et al., 1997).

It was found that within a *Sox2* heterozygous/ *Sox3* null mouse *Shh* levels were reduced, showing clear enrichment for SOX3 binding to a long range enhancer for *Shh* expression within the developing hypothalamus (Zhao et al., 2012). This indicates that SoxB1 directly promotes *Shh* expression within hypothalamic development. The resultant *Shh* levels have a profound effect on hypothalamic development and downstream phenotypes such as septo-optic dysplasia.

The study also found that *Shh* levels within a *Sox3*-null only mouse were not affected within the hypothalamus. In addition, double compound heterozygous mice (*Sox3* null +/-; *Sox2* mutant +/-) did not show a reduced *Shh* expression within the hypothalamus. However within the *Sox3* null/*Sox2* heterozygous mouse (-/Y;+/-), *Shh* deficiency was observed. These data clearly show that SoxB1 levels have an effect on *Shh* expression within the developing hypothalamus, and provides evidence that the SoxB1 members can functionally compensate for each other for their effects on *Shh* levels within the hypothalamus.

Sox2/*Shh* axis has been widely explored in mouse models, with compound heterozygous knock out *Sox2* mice (*Sox2*^{loxP}/+ x *nestin-cre*/+) showing clear reduction in *Shh* expression (Favaro et al., 2009). In addition the study also identified *Shh* as a direct SOX2 target. *Sox3* expression highly overlaps with *Sox2* and appears to be required in a similar fashion (Zhao et al., 2012). These data suggest

that *Sox3/Shh* interactions would be an interesting field of study when understanding *Sox3* function.

1.3.7 Sox3 and morphogenic signalling within the developing spinal cord

A Recent paper published in PNAS has carefully dissected the relationship between SoxB1 proteins and other critical morphogenic signalling in spinal cord development. Sox3 and Sox2 have been indicated as key co-factors in normal *Shh* signalling for ventral specification of the spinal cord. In addition, ectopic expression of SoxB1 members and constitutive activation of the *Shh* pathway within the developing limb bud, resulted in ectopic expression of ventral spinal cord transcription factors (Oosterveen et al., 2013).

Sox3 over expression alone was enough to ectopically up-regulate *Dbx1*, *Dbx2* and *Pax6* within the developing limb bud. This suggests that *Dbx1*, *Dbx2* and *Pax6* are sensitive to SoxB1 expression levels, in particular Sox3. The paper also shows that endogenous expression of most spinal cord genes requires synergistic cooperation between SoxB1 genes and morphogenic signalling present throughout the spinal cord at this developmental time point.

These new results suggest that Sox3 plays an essential, but passive, role in regulating the expression of spinal cord genes. However, there is a subset of genes involved in dorso-ventral patterning that are susceptible to Sox3 expression. These genes have never been identified as having a functional change within a Sox3 null mouse model. These data suggest the possibility that Sox3 function may have an effect within the developing murine spinal cord (Result Chapter 2: Rogers et al manuscript).

1.4 SOX3 DNA binding and functional effects

1.4.1 SOX3 binding in neural progenitor cells

To help clarify the function of *Sox3*/SOX3 throughout development it is important to identify SOX3 binding sites within cells that it is endogenously expressed. *Sox3* expression has been localised to early neural progenitor cells and recent studies looking at specific SOX binding sites identified SOX2 and SOX3 binding sites (Bergsland et al., 2011) in NP cells. Over four thousand individual SOX3 binding sites were identified within the NPCs, while only a limited number of SOX2 sites were identified. High overlap between the two SoxB1 members binding sites was observed; Sox3 occupied 96% of the Sox2 sites. These data provide more evidence of the ability of SoxB1 members to functionally compensate for each other.

The study focused on sets of neuronal genes that are initially bound by SOX2, SOX3 and subsequently SOX11. 70% of the SOX3 target sites in NP cells overlapped with those identified in mature neurons bound by SOX11. The data suggest that SOX genes work throughout neuronal differentiation to control expression of downstream neuronal targets.

The authors suggest that SOX3 might bind passively to genes to prevent genes being shut down, by epigenetic repression. Although this is not an entirely new concept, the ChIP data was the first evidence that SOX3 has a large number of binding sites in neural progenitor cells surrounding neuronal genes. However it is possible that a large number of the SOX3 binding sites have no SOX3 specific function and function outcomes are dictated by the binding partners at any given location.

1.4.2 SOX3 DNA binding and partner factors

DNA binding of SoxB1 members has been previously studied and provides greater insight into SOX function. As transcription factors, SOX protein function largely relies on their ability to bind DNA. However it has been previously shown that Sox proteins have a relatively low affinity for DNA (Kamachi et al., 2000) and they bind to a relatively small consensus sequence of only 6-8 bases (Bergsland et al., 2011; Remenyi et al., 2003). To increase their binding capacity and their specificity SOX proteins bind DNA as a homodimer or a heterodimer with other partner transcription factors. Data within early chicken lens development has previously confirmed the synergistic requirement between SoxB1 DNA binding and cofactors, showing that PAX6 is an important co-factor of SOX3 (Kondoh et al., 2004).

In addition, ChIP data has previously identified co-enrichment between SOX2 and OCT4 binding within human ES cells (Boyer et al., 2005). However Oct4/Sox2 or Sox3 co-binding sites were rare within the murine NP cells (Bergsland et al., 2011). It was also noted that 88% of the SOX sites observed within the three assays were single sites suggesting that a myriad of different cofactors are the major players in determining SOX function (Kondoh and Kamachi, 2010). Another hypothetical function for SOX DNA binding revolves around its ability to bend DNA (Ferrari et al., 1992).

In summary, as a transcription factor SOX3 function is dictated by its ability to bind DNA. Through these studies it has become evident that SOX3 binding occurs at a large number of distinct sites within neural progenitor cells. Determining which of these binding sites are functionally important to the role of endogenous *Sox3* will

require additional assays looking at potential target genes of *Sox3* (Result Chapter 2: Rogers et al manuscript).

1.5 Studying Sox3/SOX3 within a neural progenitor population

1.5.1 In vivo derived neuroprogenitors

In order to study the expression and function of SOX3 within neural progenitor populations *in vivo/in vitro* tools are commonly used as they provide a controlled system within which to study effects.

One of the most commonly used techniques for studying *in vivo* derived neural progenitors is through a neurosphere assay. Neurospheres are cultured neural progenitor cells in a defined media, derived from a dissected *ex vivo* neural progenitor population. A neurosphere assay could be used to investigate the expression of SOX3 within the adult or embryonic neurogenic populations.

Neurosphere assays are one of the key *in vitro* tools for analysis of neural progenitors within the adult/embryonic mouse brain. Neurospheres ability to be cultured over a number of passages makes them a reasonably easy experimental system for selecting and identifying neural progenitors (Reynolds et al., 1992).

The basic neurosphere assay involves dissection of an area of interest and dissociation of the cells until they are a single cell solution. The cells are then plated in serum-free chemically defined media containing specific growth factors (EGF

&FGF) (Craig et al., 1996; Reynolds et al., 1992). The “progenitors” can then grow indefinitely within the media via serial passages and supply of sufficient media. Neurospheres from the adult SVZ have been shown to culture over extended periods of time (Craig et al., 1996). Hippocampal dissections have also been cultured over extended periods through the addition of extra external factors (Bonaguidi et al., 2008; Walker et al., 2008). Cultured spheres from both SVZ and SGZ can be differentiated into all three neuronal lineages (neurons, astrocytes, oligodendrocytes).

1.5.2 Neural progenitor differentiation from embryonic stem cells

In order to study the functional effects of *Sox3*, the ability to easily culture and experimentally manipulate neural progenitors is required. As discussed above, *Sox3* expression during development is largely restricted to NP population. As such, to study the effects of *Sox3* function and potential downstream targets NP assay could prove advantageous.

A well defined technique used for studying neural progenitor cells is an embryonic stem cell differentiation assay originally established in 2003 (Ying et al., 2003). Using a *Sox1*-Gfp reporter line, it was shown that embryonic stem cells could be differentiated through the addition of specific supplements to a defined, serum free media (N2B27 assay). It was found, after four days of differentiation, the population of cells had enriched for a *Sox1*+ cell population, and after the sixth day of differentiation the population reached maximum enrichment of neural progenitor

cells. This population is a monolayer and can be further differentiated into all three neuronal cell types.

The advantages of this assay revolve around its reproducibility and the homogeneity of the NP cell population. Other assays, such as neurospheres assays, contain cells at many different stages of differentiation. In contrast, the N2B27 assay allows for systematic differentiation of ES so that the neuro-ectodermal cells are at a comparable stage of differentiation. In order to utilize the advantages of the N2B27 assay an appropriate *Sox3* null ES cell line was created to effectively identify functional effects of *Sox3* within NP cells (Results Chapter 2: Rogers et al, manuscript).

1.5.3 Summary of Sox3 function:

As previously discussed, data studying SOX3 function within neural progenitors has identified a role for SOX3 in neural progenitor maintenance within the chick spinal cord. In addition, *Sox3* function is important for normal CNS development within both humans and mice. Recent genome wide data sets looking at SOX3 binding in neural progenitors have identified numerous potential functional sites for SOX3 binding in NP cells. However functional consequences or target genes for *Sox3* within NP cells have not been identified. Potential targets/effects could be cross referenced against the growing SOX3 binding site data sets and help dissect the function of *Sox3* within neural progenitor cells. Elucidating functionality of *Sox3* within neural progenitor cells will help understand the role *Sox3* plays in development and disease (Result Chapter 2: Rogers et al, manuscript).

Aims

Aim1: Study SOX3 expression in the developing and developed CNS

1. Define SOX3 expression within the developing dorsal telencephalic vesicles and ventral diencephalon (Results Chapter 1: Rogers et al 2013)
2. Characterise SOX3 expression within the adult neurogenic zones (Results Chapter 1: Rogers et al 2013)

Aim 2: Elucidate Sox3 function in central nervous system development:

1. Generate a Sox3 null ES cell line for use in studying Sox3 function (Results Chapter 2: Hughes et al 2013)
2. Characterise the effects of the loss of *Sox3* on early gastrulation development in order to fully understand the functional consequences of polyalanine expansion on endogenous Sox3 function (Results Chapter 2: Hughes et al 2013)
3. To identify downstream targets of Sox3 thus elucidating the specific functions of Sox3 in cultured neural progenitor cells (Results Chapter 2: Manuscript Rogers et al 2013)

Methods

Immunohistochemistry

Mouse brains were fixed in 4% PFA overnight (for embryonic samples) or up to 4 days (for adult tissue). Both were then equilibrated in 30% sucrose, set in OCT compound and cryosectioned on a Leica CM1900 at 10-16 μ m. Sections were incubated with primary antibodies (primary antibodies described in table 1 in Rogers et al,2013 and DBX1 primary antibody was provided by Dr. Alessandra Pierani) overnight at 4 degrees Celsius, washed in PBS and then placed into a secondary antibody solution for 1-2 hours. Images were captured on a Zeiss Axioplan 2 upright microscope with Axiovision software, using a 20x (Zeiss Apochromat; NA0.75), 63x (Zeiss Apochromat; NA1.4) or a 100x (Zeiss Apochromat; NA1.3) objective lens. Confocal images were captured on a Leica SP5 spectral scanning confocal microscope with AnalySIS software and a 20x (Leica HCX PL APO Lambda blue; NA0.7), 100x (Leica HCL PL-APO; NA1.4) objective.

Culturing Neurospheres

14.5 dpc dorsal telencephalons were dissected, mechanically triturated and plated/passaged as previously published (Rietze and Reynolds, 2006). For immunostaining neurospheres were fixed with 4% PFA for 20 minutes and cryoprotected in 30% sucrose. Spheres were set in OCT compound and sectioned at 10-16 μ m on a Leica CM1900. Immunofluorescence performed on sections as previously stated.

ES cell culture and Neurodifferentiation

R1 ES cells were maintained on irradiated MEFs in standard conditions and neurodifferentiated as monolayers as previously described (Ying et al., 2003), referred to as (N2B27).

ES cell targeting

Sox3 Null allele: The targeting vector for generation of the null allele was the *Sox3-Neo* floxed vector published previously (Rizzoti et al., 2004). 400 clones were screened using the above probe and homologous recombinants identified by a shift in the *Bgl*II digested fragment from 8.8kb to 7.5kb. Two correctly targeted clones were transiently transfected with a Cre recombinase expression plasmid, seeded at clonal density and screened by PCR assay for the absence of *Sox3* (Rizzoti et al., 2004). Putative *Sox3* null clones were confirmed by Southern blotting (5kb *Bgl*II fragment using the 5' *Sox3* probe (Hughes et al., 2013)).

Chimera generation

129 derived ES clones were injected into c57/Bl6xB6D2F1 2.5dpc morula, cultured overnight and transferred as blastocysts into the uterus of 2.5dpc pseudopregnant recipients. Chimeric embryos were collected around gastrulation (7.5dpc and 8.5dpc) and imaged on a Zeiss Axiophot upright microscope with AnalySIS software, using a 40x (Zeiss Neofluar; NA0.75) or 20x (Zeiss Neofluar; NA0.5) objective lens

qPCR- gene estimation of percentage chimerism

Embryonic chimerism was determined using qPCR comparing against a standard curve of *Sox3* dosage generated from adult gDNA of WT and *Sox3*-null mixed 0:100, 25:75, 50:50, 100:0. Loading was corrected using primers against *Sox1* and *Ngn3*.

In Situ Hybridisations

Embryos were fixed in 4%PFA overnight embryos were then equilibrated in 30% sucrose overnight, set in OCT compound and cryosectioned on a Leica CM1900 at 10µm. ISH on fixed frozen sections was performed using a *Sox3* probe from a fragment cloned with the following primers; 5'-AGCGCCTGGACACGTACAC-3' and 5'-AGCGCCTGGACACGTACAC-3' Images were captured on a Zeiss Axiophot upright microscope with AnalySIS software, using a 2.5x (Zeiss Neofluar; NA0.5) or a 10x (Zeiss Neofluar; NA0.3) objective lens

RNA extraction

RNA was extracted from embryonic tissue/ or samples used for microarray analysis using Qiagen RNeasy RNA extraction kit as per the manufacturer's instructions and eluted in 30-50ul of RNase free H₂O. 3ul of each sample was run on a 1% RNase free agarose gel, Microarray samples checked using the Agilent Bioanalyser, at the Adelaide microarray facility were used to check RNA integrity number to confirm the quality of RNA.

cDNA Production

RNA was converted into cDNA using the Applied Biosystems High Capacity RNA-to-cDNA kit as per manufacturer's instructions for 1 microgram of starting RNA in a final volume of 20ul.

qRT PCR

All qPCR reactions consisted of 2X Fast SYBR green master mix at a final concentration of 1X, approximately 5ng of template cDNA, forward and reverse primers at a final concentration of 250nM each, and H₂O to a final volume of 15ul. Each reaction was set up in triplicate. Reactions were completed using an Applied Biosystems Step One Plus thermo cycler with the following parameters: 95°C for 20 secs, 39 cycles of; 95 °C for 3 secs, 60 °C for 30secs. This was followed by generation of dissociation curve by increasing the sample temperature in 0.5 °C increments and measuring fluorescence levels after each increment

Whole mount *in situ*

Whole mount tissue was rehydrated in MeOH series in the following order. First 75% then 50% and 25% v/v MeOH:PBS (0.1% v/v Triton) for 5minutes each. They were then processed as described in Thomas et al (1998).

Primers

Primer Name	Sequence	Application
mSox3Fwd	GTACCTGCCACCTGGCGGGG	Sox3 null genotyping
mSox3Rev	GTGGGCTGCCTCTGGGTAGG	Sox3 null genotyping
S3_3UTRF2	CGTAACTGTCCGGGTTTTGT	Sox3 null genotyping
S3_3UTRR2	CACGCACACCTGGCTATAAA	Sox3 null genotyping
Dbx1 Ribo F(524)	CAACAGACCCACCACCTTCT	In situ Probe generation
Dbx1 Ribo R(524)	AGGAGCTGGCACTCTGAAAA	In situ Probe generation
Dbx1 F (100)	CGTTCGCCTTTCCATACTTC	qRT pcr
Dbx1 R (100)	GAGAAGGTCCCAGGGATAGG	qRT pcr
Fezf2 F (105)	AGAGCCACAGCAAACCTACCG	qRT pcr
Fezf2 R (105)	GGCGGGTGAGGTTATAGTGA	qRT pcr
Mirg F (115)	CCTCTGTCCCACAGCCGTCG	qRT pcr
Mirg R (115)	GCAGCAACCCTGGAGCTGGAA	qRT pcr
mSox3 5'UTR FWD	GAACGCATCAGGTGAGAGAAG	qRT pcr
mSox3 5'UTR REV	GTCGGAGTGGTGCTCAGG	qRT pcr

Sox1 F	GACTTGCAGGCTATGTACAACATC	qRT pcr/Chimera %
Sox1 R	CCTCTCAGACGGTGGAGTTATATT	qRT pcr/Chimera %
sox2 F	ACCAGCTCGCAGACCTACAT	qRT pcr
sox2 R	TCGGACTTGACCACAGAGC	qRT pcr
Pax6 FWD	GCACATGCAAACACACATGA	qRT pcr
Pax6 REV	ACTTGGACGGGAACTGACAC	qRT pcr
Nestin FWD	CCCTGAAGTCGAGGAGCTG	qRT pcr
Nestin REV	CTGCTGCAGCTCTAAGCGA	qRT pcr
beta-actin (89bp)-FP	CTGCCTGACGGCCAGG	qRT pcr
beta-actin (89bp)-RP	GATTCCATACCCAAGAAGGAAGG	qRT pcr
Ddr2 qRT F (95bp)	CCTGGCGTAACCGGCATGGG	qRT pcr
Ddr2 qRT R (95bp)	TCTGGCGACGATGGGTGGCT	qRT pcr
Slit1 qRT F(90bp)	GCAAGAGTCCGAGTGCCGGG	qRT pcr
Slit1 qRT R(90bp)	GACAGCGGGCGTGTGGTCTG	qRT pcr
Cspg5 F (106bp)	AATTTTGCGACACTCCCCGT	qRT pcr
Cspg5 R (106bp)	TCCTGGATTTTGTGGGGAGC	qRT pcr
Fam174b F (105bp)	CTGCGGAACGTGTGGAAATG	qRT pcr
Fam174b R (105bp)	CTGACAAGCCAGCGAAGTTT	qRT pcr
Efnb3 F (91bp)	TGGAGAGAGACAGAGGGGCAGC	qRT pcr
Efnb3 R (91bp)	GCACCCCGGGAGGTTGCATT	qRT pcr
Fezf2 F RT 108	CCGCCAGGCCAGCACTCTCT	qRT pcr
Fezf2 R RT 108	GTGTTGAGCGTGGAGCTGCGA	qRT pcr
Vgll3 F (101bp)	TGGGCCAAGCCAACACCTTGC	qRT pcr
Vgll3 R (101bp)	CCGCTGGCTCGAAAGAGCTGAG	qRT pcr
Cp F (96bp)	GCCCACCAGGTCGCTCCTCA	qRT pcr
Cp R (96bp)	TGGGTCCCATTTCTTTGGGGACA	qRT pcr
Dbx1 Site1 3' F	CTTTGGTCTCCACAAGCTTTCT	Chip qRT PCR
Dbx1 Site1 3' R	GAATGTGGCCTTTAACAACCTCAC	Chip qRT PCR
Dbx1 Site2 3' F	ACTCTTGAGCCCCACTGCTA	Chip qRT PCR
Dbx1 Site2 3' R	GCTGAATGGAGCCTAGATGG	Chip qRT PCR
Dbx1 Site3 3' F	ATTGGGTGCTACCTGTGAC	Chip qRT PCR
Dbx1 Site3 3' R	TCCTCCGAATTGTTTGTTC	Chip qRT PCR
Dbx1 intron site1 F (Site2 in paper)	AGGGGGAGAGTTTCGCTTTA	Chip qRT PCR
Dbx1 intron site1 R (Site 2 in paper)	ATATGGCTCTCGCTGCTTGT	Chip qRT PCR
Dbx1 intron1 site2 R	TGGACAAGCAGCGAGAGCCAT	Chip qRT PCR
Dbx1 intron1 site2 F	AGCCTCTGCCATCACCGCTC	Chip qRT PCR
Notch1 -ve Jonas F	GTACCCACCTGGCCCAAGA	Chip qRT PCR
Notch1 -ve Jonas R	GGTGACAGCAGCCCTTTGTC	Chip qRT PCR
Notch1 Chip Jonas F	CCCATGCACAACCTACCAA	Chip qRT PCR
Notch1 Chip Jonas R	TCTTCTGCTCAGCACTTGGGA	Chip qRT PCR
Nestin Chip Jonas F	GCCCCAGTCAGTCTTCTGAG	Chip qRT PCR
Nestin Chip Jonas R	GCTGGTGACAGACAAAAGCA	Chip qRT PCR
Ngn3 F	CCCCAGAGACACAACAACCT	qRT PCR/Chimeria %
Ngn3 R	AGTCACCCACTTCTGCTTCCG	qRT PCR/Chimeria %

Animal husbandry:

All animals were approved by the appropriate ethics committees at the University of Adelaide (Adelaide, Australia). All mice were housed and kept in the laboratory Animal services at the University of Adelaide. Mice were kept in a 12 hour day night cycle with sufficient food and water. Mice 2 weeks or older were sacrificed by cervical dislocation or decapitations for mice younger.

Generation of Sox3 null mice and genotyping was conducted as described within (Rizzoti et al 2004).

Mouse embryo generation

Mouse embryos were generated by either of the following crosses

- a) Male *Sox3* null^{-Y} X Female *Sox3* null^{+/-}
- b) Male *Sox3* null^{+Y} X Female *Sox3* null^{-/-}
- c) Male *Sox3* null^{+Y} X Female *Sox3* null^{+/+}

Microarray analysis

Four wild type and four Sox3 null day 4 (N2B27differentiated) samples were used on the GeneChip Mouse Gene 1.0 ST Array, performed by the Adelaide Microarray Facility. 2 way-ANOVA using batch as a factor was used to identify the significantly regulated genes.

In vitro CHIP

ChIP was performed as per (Bergsland et al., 2011) using 5ug of SOX3 or IgG antibodies, with Dynabeads blocked overnight with 200ug BSA, 200ug yeast tRNA and 200ug Glycogen .

In vivo ChIP

10.5 dpc heads were disassociated with scalpel blades before fixation for 10mins as per Bergsland et al 2011. Antibodies and blocking of Dynabeads performed as above.

Results

Result Chapter 1

Expression of the murine transcription factor SOX3 during embryonic and adult neurogenesis

Nicholas Rogers¹, Pike-See Cheah^{1,2}, Eva Szarek¹, Kakoli Banerjee¹, Jeffrey Schwartz³ & Paul Thomas^{1*}

¹Discipline of Biochemistry, School of Molecular and Biomedical Sciences, University of Adelaide, Adelaide, Australia

²Department of Human Anatomy & NeuroBiology & Genetics Group, Genetic Medicine Research Centre, Faculty of Medicine and Health Sciences, University Putra Malaysia.

³School of Medicine, Griffith University, Gold Coast, Queensland, Australia

Statement of Authorship: Appendix 1

Hyperlink to paper:

<http://www.sciencedirect.com/science/article/pii/S1567133X13000422#>

Rogers, N., Cheah, P-S., Szarek, E., Banerjee, K., Schwartz, J. & Thomas, P. (2013) Expression of the murine transcription factor SOX3 during embryonic and adult neurogenesis.
Gene Expression Patterns, v. 13(7), pp. 240-248

NOTE:

This publication is included on pages 43-52 in the print copy of the thesis held in the University of Adelaide Library.

It is also available online to authorised users at:

<http://doi.org/10.1016/j.gep.2013.04.004>

Results Chapter 2

Mechanistic Insight into the Pathology of Polyalanine Expansion Disorders Revealed by a Mouse Model for X Linked Hypopituitarism

James Hughes¹, Sandra Piltz¹, **Nicholas Rogers¹**, Dale McAninch¹, Lynn Rowley² and Paul Thomas^{1*}

¹School of Molecular & Biomedical Science, University of Adelaide, North Terrace, Adelaide, South Australia, Australia 5005

²Pituitary Research Unit, Murdoch Childrens Research Institute, Flemington Road, Melbourne, Victoria, Australia, 3052

Statement of Authorship: Appendix 2

Hyperlink to paper:

<http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.100329>

[0](#)

Mechanistic Insight into the Pathology of Polyalanine Expansion Disorders Revealed by a Mouse Model for X Linked Hypopituitarism

James Hughes¹, Sandra Piltz¹, Nicholas Rogers¹, Dale McAninch¹, Lynn Rowley², Paul Thomas^{1*}

¹ School of Molecular and Biomedical Science, University of Adelaide, Adelaide, South Australia, Australia, ² Pituitary Research Unit, Murdoch Childrens Research Institute, Melbourne, Victoria, Australia

Abstract

Polyalanine expansions in transcription factors have been associated with eight distinct congenital human diseases. It is thought that in each case the polyalanine expansion causes misfolding of the protein that abrogates protein function. Misfolded proteins form aggregates when expressed *in vitro*; however, it is less clear whether aggregation is of relevance to these diseases *in vivo*. To investigate this issue, we used targeted mutagenesis of embryonic stem (ES) cells to generate mice with a polyalanine expansion mutation in *Sox3* (*Sox3-26ala*) that is associated with X-linked Hypopituitarism (XH) in humans. By investigating both ES cells and chimeric mice, we show that endogenous polyalanine expanded SOX3 does not form protein aggregates *in vivo* but rather is present at dramatically reduced levels within the nucleus of mutant cells. Importantly, the residual mutant protein of chimeric embryos is able to rescue a block in gastrulation but is not sufficient for normal development of the hypothalamus, a region that is functionally compromised in *Sox3* null embryos and individuals with XH. Together, these data provide the first definitive example of a disease-relevant PA mutant protein that is both nuclear and functional, thereby manifesting as a partial loss-of-function allele.

Citation: Hughes J, Piltz S, Rogers N, McAninch D, Rowley L, et al. (2013) Mechanistic Insight into the Pathology of Polyalanine Expansion Disorders Revealed by a Mouse Model for X Linked Hypopituitarism. PLoS Genet 9(3): e1003290. doi:10.1371/journal.pgen.1003290

Editor: Sally Camper, University of Michigan, United States of America

Received: July 12, 2012; **Accepted:** December 16, 2012; **Published:** March 7, 2013

Copyright: © 2013 Hughes et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Australian Health and Medical Research Council (<http://www.nhmrc.gov.au/>). PT is a Pfizer Australia Research Fellow. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: PT is supported by a Fellowship from Pfizer Australia. This does not alter the authors' adherence to all the PLOS Genetics policies on sharing data and materials.

* E-mail: paul.thomas@adelaide.edu.au

Introduction

Trinucleotide repeat expansions are a relatively common cause of human disease. The expanded trinucleotide can occur in an untranslated region, for example in Fragile X syndrome in which a CGG repeat adjacent to the *FMRI* promoter causes hypermethylation and gene silencing. Alternatively repeat expansions can occur in exonic regions and result in elongation of homopolymeric amino acid tracts. For example polyglutamine (PQ) diseases such as Huntingtons disease, are associated with long unstable PQ-encoding stretches that lead to the production of a toxic species and late onset disease characterised by the loss of a subset of neurons. In addition to PQ encoding repeats, polyalanine (PA) repeat expansion has recently emerged as a significant cause of human disease. PA expansions have been linked to nine disorders of which eight are congenital and one is late onset. Each is caused by PA expansion in a separate gene, with the eight congenital disorders linked to PA expansions in developmentally-important transcription factors. These, and the associated disorders, are SOX3 (X linked Hypopituitarism), HOXA13 (hand-foot-genital syndrome), ARX (syndromic and non-syndromic X-linked mental retardation), HOXD13 (synpolydactyly type II), PHOX2B (congenital central hypoventilation syndrome), FOXL2 (blepharophimosis, ptosis and epicanthus inversus), ZIC2 (holoprosencephaly) and RUNX2 (cleidocranial dysplasia). The ninth is the ubiquitous

RNA binding protein PABPN1 which is associated with the late onset disease oculopharyngeal muscular dystrophy (OPMD).

Despite considerable functional analyses, the mechanism by which PA expansion mutations cause disease is not completely understood. Phenotype/genotype correlations in humans and mouse models indicate that many PA alleles give rise to disease phenotypes that resemble (*HOXA13*, *FOXL2*, *ZIC2*) or are less severe (*ARX*) than null alleles, consistent with complete or partial loss-of-function (LOF) [1,2,3,4]. In contrast, some PA alleles cause similar but more severe disease phenotypes than null alleles consistent with toxic gain-of-function (GOF) and/or dominant negative activity (*HOXD13*, *PHOX2B*) [5,6]. Despite these differences in mode of inheritance, all PA proteins behave very similarly *in vitro*, such that over-expression in cell culture results in the generation of cytoplasmic and/or nuclear aggregates, which are likely to arise through protein misfolding [7,8,9]. While the relevance of cellular aggregates to PA disease in general is unclear, nuclear inclusions that contain mutant PABPN1 protein are a hallmark of OPMD [10], suggesting that aggregates may also form in patients with PA expansion mutations in developmental transcription factors. Protein aggregation also occurs in the related polyglutamine (PQ) disorders where PQ expansion confers toxic GOF [11,12]. Together, these data suggest that aggregate formation may have a pathogenic role in PA disease alleles with GOF activity. However, the critical question of whether aggre-

Author Summary

Alanine is one of the 20 amino acid building blocks from which proteins are generated. Nearly 500 human proteins contain stretches of consecutive alanine residues ranging from 4 to 20 amino acids in length. Whilst the function of these polyalanine (PA) tracts remains unknown, they are interesting because DNA changes (mutations) that increase their length above a threshold are responsible for nine different human disorders. *In vitro* studies indicate that expanded PA proteins misfold and aggregate, suggesting that there may be a common “gain-of-function” mechanism that underpins this group of disorders. However, these data are difficult to reconcile with genetic studies, which indicate that most PA mutations cause protein loss-of-function. Therefore, to investigate the pathological mechanism of PA disorders we generated a mouse model of X-linked Hypopituitarism (XH), a disease caused by PA expansion in the SOX3 protein. Strikingly, we found that the mouse version of the disease-causing protein was almost completely cleared from cells and that aggregates do not form *in vivo*. These data explain why this type of mutation causes protein loss-of-function and reveals why nature limits the length of PA stretches.

gates form *in vivo* and, if so, how they may be implicated in the pathogenesis of LOF PA diseases remains unresolved.

To investigate these issues, we used targeted mutagenesis of ES cells to generate a 26 alanine PA tract expansion mutation in *Sox3* (*Sox3-26ala*). The analogous human mutation causes X linked Hypopituitarism (XH), a disease in which hemizygous males have GH deficiency (resulting in short stature) and fully penetrant intellectual disability [13]. Importantly, we find no evidence of SOX3-26ala aggregate formation in neural derivatives of targeted ES cells *in vivo* or in neurodifferentiated ES cell cultures. Instead, the *Sox3-26ala* mutation leads to a massive reduction in SOX3 protein within the nucleus. We also present developmental and biochemical evidence that residual mutant protein retains some activity, indicating that *Sox3-26ala* functions as a partial LOF allele.

Results/Discussion

In order to study the effects of disease-associated PA expansion at a cellular level we created R1 ES cells with a targeted mutation of *Sox3*. Homologous recombination was used to generate XY ES cells carrying a 36 bp expansion in the first PA tract of *Sox3*, extending the tract from 14 to 26 alanines (referred to hereafter as *Sox3-26ala* ES cells; Figure 1). Morula injection of mutant ES cells resulted in chimeras with up to 95% mutant cell contribution as assessed by coat colour (Table 1). None of these chimeras displayed any evidence of growth retardation (consistent with the absence of short stature in heterozygous female carriers of the human *SOX3-26ala* mutation who also contain a mixture of WT and mutant cells, due to X inactivation [13]). However, despite extensive breeding, none of these chimeras exhibited germline transmission of the mutant allele. In contrast, germline competence was demonstrated with WT parental ES cells as well as a clone carrying the PGK-*Neo* cassette without the PA tract expansion (referred to hereafter as *Neo* ES cells). These data suggest that the PA mutation in *Sox3* had caused a block in male fertility. This is consistent with a recently published report demonstrating that *Sox3* LOF in the germ cell lineage resulted in

an early defect in spermatogenesis and the lack of any reported transmission of the human *SOX3-26ala* allele from hemizygous males [13,14].

Since we were unable to transmit the *Sox3-26ala* mutation through the male germline, the phenotypic consequences of the mutation were initially investigated in *Sox3-26ala* <-> WT chimeric embryos. Immunostaining with a SOX3-specific antibody revealed a dramatic reduction in SOX3 protein in 13.5 dpc telencephalic ventricular zone cells expressing the *Sox3-26ala* allele (identified by NEO-immunoreactivity) in comparison to neighbouring WT cells (Figure 2A). The ability of the antibody to detect mutant protein was confirmed by staining COS-7 cells expressing exogenous mouse *Sox3-26ala*, in which large peri-nuclear and cytoplasmic aggregates were common (Figure S1). A striking reduction in mutant SOX3 protein was also observed in 7.5 dpc, 9.5 dpc, 10.5 dpc and 11.5 dpc chimeras (Figure 3 and Figure S2) indicating that this phenotype was not stage-dependent. High power microscopy failed to detect any cytoplasmic or nuclear aggregates of mutant protein but did reveal a very low level of protein in the nucleus (Figure 2A). Comparison of *Sox3-26ala* and *Sox3*-null embryonic CNS cells confirmed that the low level of SOX3-26ala protein that we observed was not due to background signal (Figure 2B). Mutant cells appeared morphologically normal and there was no apoptotic induction as assessed by staining for activated Caspase3 (Figure S3). To further characterise the cellular phenotype of *Sox3-26ala* mutant cells, we performed neurodifferentiation of *Sox3-26ala* ES cells (which are XY and therefore lack a WT *Sox3* allele). Immunohistochemical analysis confirmed the overwhelming reduction of nuclear SOX3 protein in mutant cultures (Figure 2C) in which the expression of other neural progenitor markers were unaffected (data not shown). Rare cells with near normal levels of nuclear mutant protein were also detected but in no case was there evidence of protein aggregation (Figure 2C). Western blot analysis further supported the near complete loss of steady-state SOX3 protein levels in mutant cells (Figure 2D). As triplet repeat expansion mutations have been shown to affect mRNA transcription [12], we compared *Sox3* message levels in WT and mutant cells *in vitro* and *in vivo* using qPCR and *in situ* hybridisation, respectively (Figure 2E–2F). No difference in *Sox3* mRNA level was detected, indicating that the massive reduction in SOX3 protein levels in the mutant cells has a post-transcriptional aetiology and presumably occurs via degradation of misfolded protein [8]. This was supported by *in vitro* transcription/translation analysis which indicated that translation of the mutant transcript was unaffected by the PA mutation (Figure S4).

Taken together, these results indicate that *Sox3-26ala* is a LOF allele and argue strongly against a role for aggregation in the pathogenesis of XH. To further investigate LOF as the mechanism of disease, we compared the development of the hypothalamus, infundibulum and anterior pituitary in *Sox3-26ala* chimeric embryos to *Sox3* heterozygous embryos that carry a complete LOF (null) allele. *Sox3* +/- embryos exhibit a defect in infundibular development that results in aberrant induction and bifurcation of the anterior pituitary primordium, Rathke's Pouch (RP) [15] accompanied by expansion of the floor of the ventral diencephalon (Figure 3). Similar defects in infundibular morphology and pituitary localisation have also been identified in XH patients [16]. Importantly, 13.5 dpc and 11.5 dpc high percentage *Sox3-26ala* <-> WT chimeras exhibited dysmorphology of the infundibulum, Rathke's Pouch and ventral diencephalon that was indistinguishable from *Sox3*+/- embryos (Figure 3). These abnormalities were also present in low percentage chimeras, although were less severe consistent with the loss of SOX3

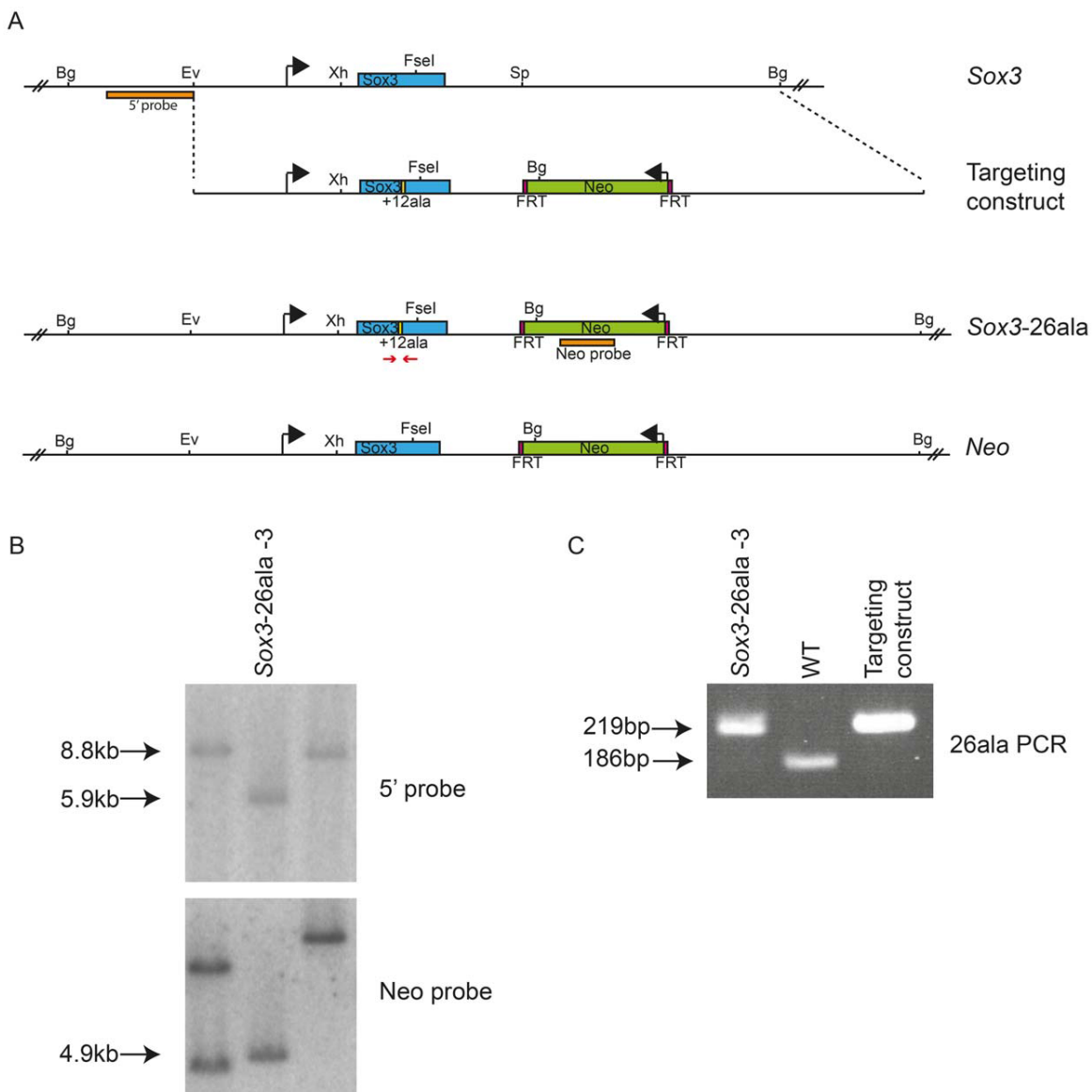


Figure 1. Generation of Sox3-26ala ES cells. Scale representation of the Sox3 locus, targeting vector and recombinant alleles (A). Probing of *BglIII* digested DNA from ES cell clones with the 5' probe yielded an 8.8 kb fragment from the WT locus and a 5.9 kb fragment when the Neo cassette was recombined into the Sox3 locus (Sox3-26ala or Neo). B) Representative Southern blot of 3 clones including a targeted clone (Sox3-26ala-3) is shown. C) PCR using primers spanning the alanine expansion (red arrows in A) was used to distinguish whether targeted clones carried the expansion and gave a 219 bp product instead of 186 bp as seen in WT.
doi:10.1371/journal.pgen.1003290.g001

Table 1. Failure of Sox3-26ala targeted ES clones to transmit through the germline.

Clone	Number of chimeras	Percentage chimerism (range)	Number of pups born	Percentage of ES-derived pups (agouti)
WT	15	5–100	50	18
Neo	4	50–90	35	49
Sox3-26ala-1	6	5–50	127	0
Sox3-26ala-2	4	60–90	81	0
Sox3-26ala-3	3	20–60	53	0
Sox3-26ala-4	4	75–95	156	0

doi:10.1371/journal.pgen.1003290.t001

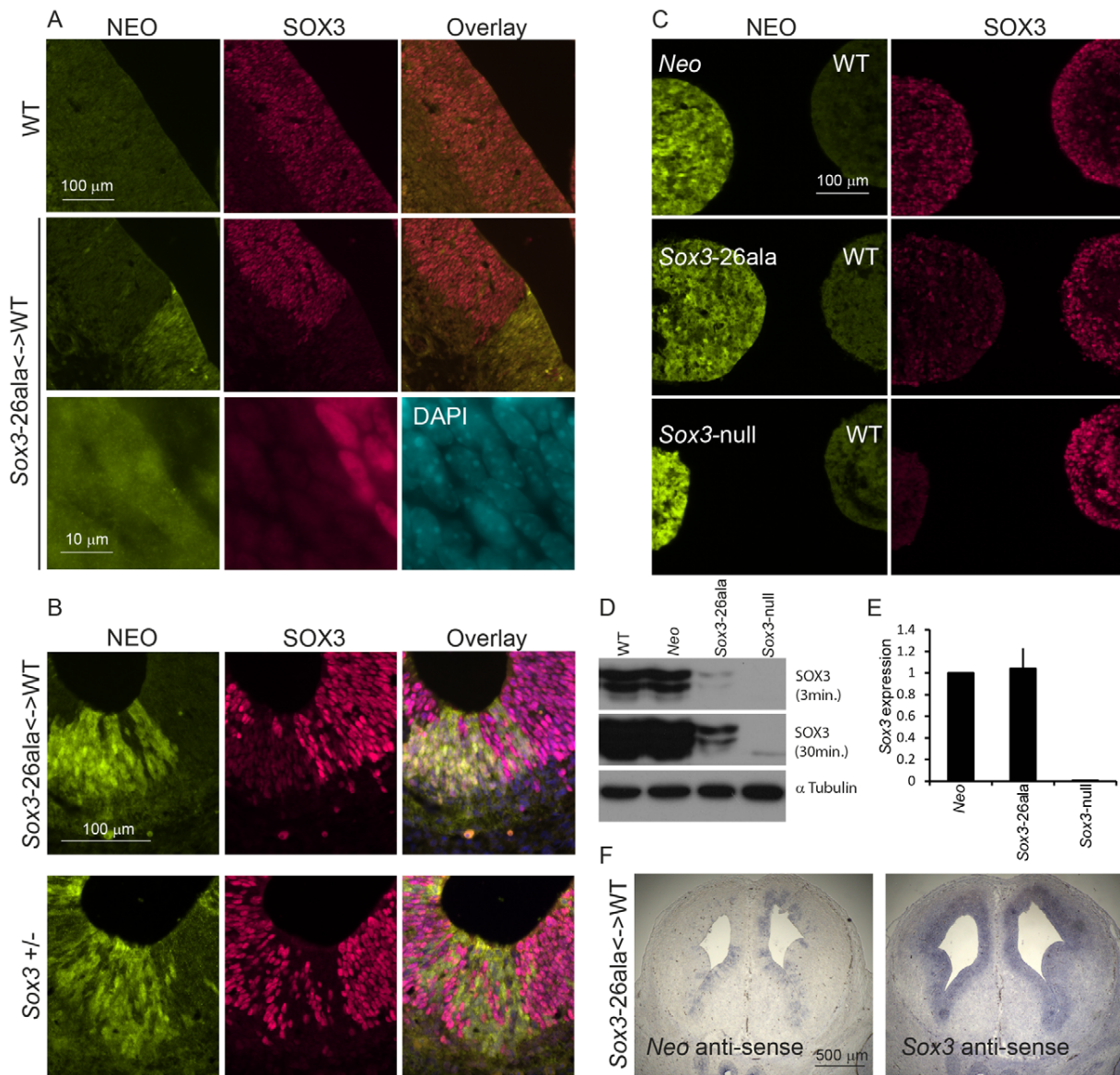


Figure 2. Transcription is unaffected but protein is cleared from mutant cells. A) SOX3 protein is present in every WT cell (NEO⁻) of the 13.5 dpc telencephalic ventricular zone but virtually absent from equivalent tissue derived from *Sox3-26ala* cells (NEO⁺). B) Comparison of SOX3 immunostaining on *Sox3*-null cells (from a 14.5 dpc +/- embryo) and *Sox3-26ala* expressing cells (from a *Sox3-26ala* <-> WT chimera) confirming that the antibody is SOX3-specific and that the *Sox3-26ala* expressing cells exhibit a low level of residual nuclear protein. C) WT, *Neo*, *Sox3-26ala* and *Sox3*-null ES cells were differentiated for 5 days in CDM as multi-cellular bodies. Rare SOX3 positive cells were detected in *Sox3-26ala* CDMs while the majority of cells had low SOX3 protein levels in comparison to neighbouring WT CDM bodies processed on the same slide. D-E) WT, *Neo*, *Sox3-26ala* and *Sox3*-null ES cells were grown in N2B27 for 4 days to form neural progenitors. Western blotting for SOX3 reveals a dramatic reduction of protein in *Sox3-26ala* cells (D); 3 and 30 minute exposures are shown. E) Transcript levels of *Sox3* are not affected in *Sox3-26ala* cells as determined by qPCR. Three experimental replicates are shown. Data was normalised to *Sox3* levels in *Sox3-Neo* control cells and error bars represent SEM. F) ISH confirms that *Sox3* transcript is present at comparable levels in ventricular zone cells at 13.5 dpc derived from both WT (*Neo*⁻) and *Sox3-26ala* (*Neo*⁺) cells. ISH performed on adjacent 10 μm coronal sections of 13.5 dpc chimeric telencephalon. doi:10.1371/journal.pgen.1003290.g002

function being responsible for this phenotype. Notably, there was an obvious reduction in SOX3 protein in mutant (NEO⁺) cells thereby confirming that PA expansion has a functional impact in neural cells that are directly implicated in XH pathology.

Having established that *Sox3-26ala* behaves as a LOF allele, we next considered whether the low level of remnant SOX3 protein in mutant cells was functional. To investigate this, we initially performed transactivation assays in COS-7 cells using wild type

and mutant human and mouse SOX3 expression constructs and a luciferase reporter containing four SOX consensus motif (SOCM) binding sites. Both mouse and human SOX3-26ala proteins showed activity in this assay that was significantly higher than background (Figure 4A). However, consistent with previous reports [7,16], this activity was much lower than WT SOX3 protein. To investigate whether this reduction was caused by lower nuclear protein levels or an inherent defect in mutant protein

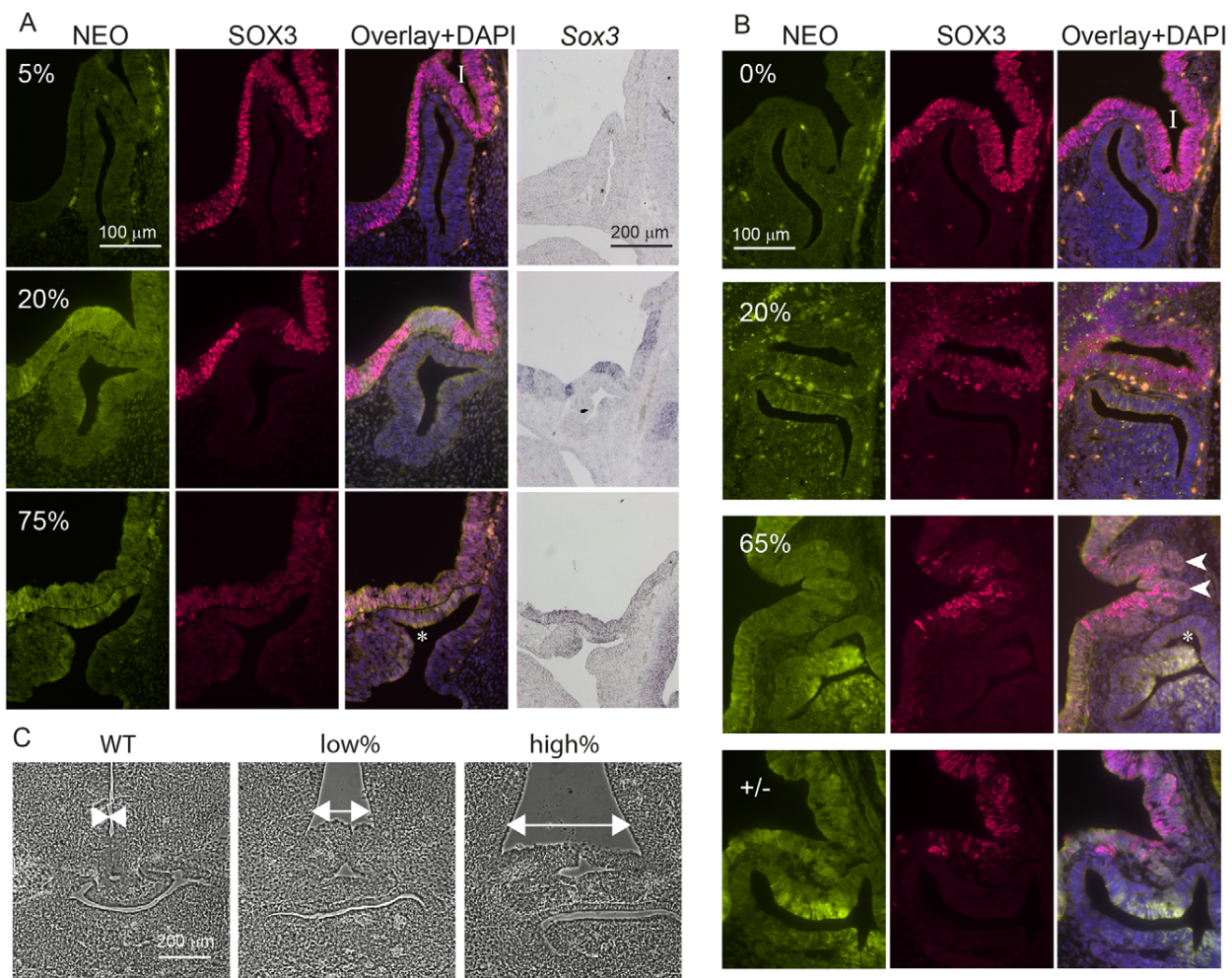


Figure 3. *Sox3-26ala* cells cause pituitary defects indistinguishable from *Sox3*-null cells. WT, *Sox3*^{+/-} or *Sox3-26ala*^{<->}WT chimeras were cut sagittally at 11.5 dpc (A) or 13.5 dpc (B) and immunostained for SOX3 and NEO expression. Percentage chimerism for each embryo in (A) and (B) was determined by qPCR as outlined in the methods. ISH for *Neo* on adjacent sections at 11.5 dpc confirmed the identification of mutant cells within the ventral diencephalon (A). Examples at 11.5 dpc show the infundibulum (I) appears unaffected in a 5% chimera, shallow in a 20% chimera and absent in a 75% chimera that also displayed a Rathke's Pouch (*) that had failed to detach from the oral ectoderm. At 13.5 dpc, heterozygous and high percentage (65%) chimeric embryos displayed a distorted infundibulum (I) with a lobular edge (arrow heads) and a branched Rathke's Pouch (*). Low percentage chimeras (20%) look similar to WT (0%). C) Phase micrographs of 13.5 dpc coronal sections through the developing pituitary show a broadening at the base of the third ventricle in chimeras (arrows). Chimerism for embryos shown in (C) was determined based on immunoreactivity for NEO in adjacent sections (data not shown). doi:10.1371/journal.pgen.1003290.g003

transactivation activity, we measured that relative amount of WT and mutant SOX3 protein in the nucleus of transfected cells by Western Blot (Figure 4B). We observed a reduction of similar magnitude to the reduced luciferase output, suggesting that the mutant protein that is present in the nucleus has similar activity to WT. To investigate whether the residual nuclear SOX3-26ala protein is functional *in vivo* we compared the phenotypes of 7.5 dpc chimeras generated from either *Sox3*-null or *Sox3-26ala* ES cells, as it has been reported that *Sox3*-null ES cell <-> WT embryo chimeras exhibit severe gastrulation defects [15]. As previous analysis of *Sox3*-null ES cell chimeras was not performed with R1 ES cells (the parent line used to generate the 26ala mutant ES cells), we generated R1 *Sox3*-null ES cells for this experiment. Consistent with previous reports, a high proportion (52%) of *Sox3*-null ES cell chimeras generated abnormal gastrulae (Figure 5). In

contrast, only 19% of *Sox3-26ala* ES cell <-> WT embryo 7.5 dpc chimeras were morphologically abnormal which was not significantly different to the proportion of abnormal chimeric embryos generated using *Sox3*-flox control ES cells (Figure 5). Immunostaining of *Sox3-26ala* chimeras revealed a reduction in SOX3 protein levels that was similar to later embryonic stages (Figure S1). These data indicate that nuclear SOX3-26ala protein is functional *in vivo* and is sufficiently abundant to prevent overt defects in gastrulation. We therefore conclude that *Sox3-26ala* is a partial LOF allele.

This study represents the first investigation of the disease-associated *Sox3-26ala* mutation under the control of the endogenous locus in a whole animal setting. The complete lack of aggregates in SOX3 positive CNS zones (including the hypothalamic-pituitary axis) provides strong evidence that aggregation is

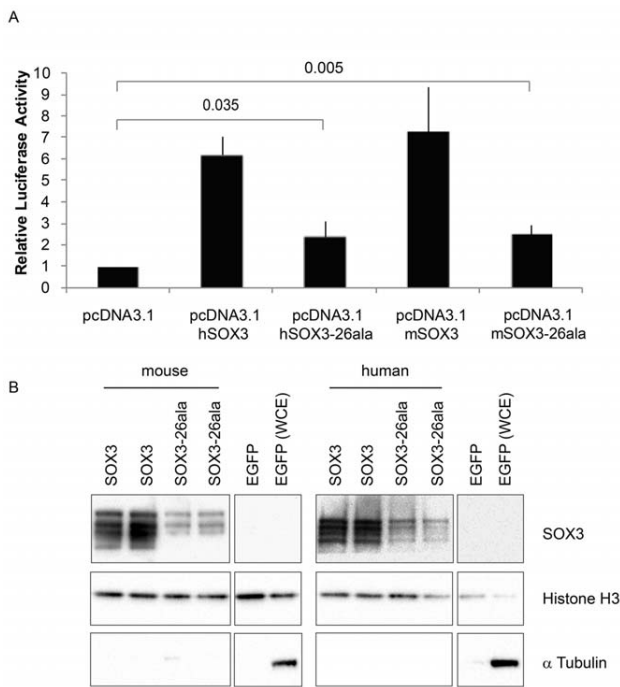


Figure 4. SOX3-26ala from mouse and human retains transactivation activity. A) COS-7 cells were transfected with pcDNA3.1 expression vector containing either mouse *Sox3*, human *SOX3*, mouse *Sox3-26ala*, human *SOX3-26ala* or an empty vector control. Values represent mean normalised luciferase values plus standard deviation of four independent experiments measured 48 hours after transfection. Student's T-tests (two tailed, unequal variance) of *SOX3-26ala* from human or mouse compared to empty vector control show a statistically significant increase in luciferase activity. B) Nuclear protein lysates prepared from duplicate plates 48 hours after transfection show that less *SOX3* is detected in the nucleus of cells expressing both mouse and human *SOX3-26ala*. pcDNA3.1-EGFP transfected cells were used as a control and prepared for both nuclear protein and whole cell extracts (WCE). Blotting for Histone H3, indicates equal loading and blotting for α -Tubulin shows an absence of cytoplasmic contamination in nuclear preparations. Transfection efficiency was determined by co-transfecting EGFP and counting positive cells prior to harvesting and found to be equal for all plasmids. doi:10.1371/journal.pgen.1003290.g004

not a feature of the human disease. In contrast, our data demonstrate that PA expansion results in a dramatic reduction in *SOX3* protein due to a post-translational defect. Given that PA expansion proteins co-localise with chaperones *in vitro* and that aggregation is promoted by pharmacological inhibition of the proteasome [8], it seems likely that mutant *SOX3* protein misfolds and is cleared from the cell. Of note, the small amount of protein that remains translocates to the nucleus and is functionally active. The level of protein that remains is sufficient for some developmental contexts but not others (i.e. gastrulation but not pituitary induction). This context-specific threshold of *SOX3* activity may explain why XH patients with *SOX3* PA mutations, but not null mutations, have been identified.

In direct contrast to the *Hoxd13* (+7ala) *spdh* spontaneous mouse model in which mutant protein is mislocalised to the cytoplasm [8], our data provide the first definitive example of a disease-relevant PA mutant protein that is both nuclear and functional, thereby manifesting as a partial LOF allele. Interestingly, mutant protein levels in the *Hoxa13* and *Arx* PA mouse models, which also

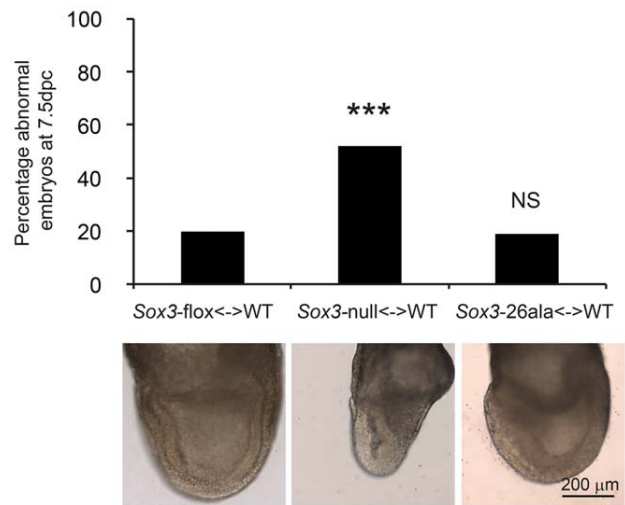


Figure 5. Residual nuclear *SOX3-26ala* protein rescues a gastrulation defect of *Sox3-null* <-> WT chimeric embryos. *Sox3-26ala* <-> WT chimeras are normal at 7.5 dpc (gastrulation) unlike *Sox3-null* <-> WT chimeras. A total of 15 *Sox3-flox* <-> WT ES chimeras, 31 *Sox3-null* <-> WT chimeras and 21 *Sox3-26ala* <-> WT chimeras were blind scored by two independent operators as morphologically normal or abnormal. The average score for each embryo was used to plot the percentage of abnormal embryos for each condition and chi squared analysis was performed with *Sox3-flox* <-> WT embryos used to set expected outcomes. Significantly more *Sox3-null* <-> WT chimeras were abnormal ($p=0.0001$) while *Sox3-26ala* <-> WT chimeras did not deviate from expected ($p=0.95$). An example of normal morphology is shown for *Sox3-flox* <-> WT and *Sox3-26ala* <-> WT chimeras and an abnormal *Sox3-null* <-> WT chimera is also shown that exhibits distortion of the ectodermal layer and apparent expansion of cells at the primitive streak and the adjacent extra-embryonic region. doi:10.1371/journal.pgen.1003290.g005

exhibit LOF inheritance, are diminished but not abolished by western blot and whole mount immunodetection [1,17]. This raises the possibility that nuclear localisation of suboptimal levels of functional protein may be a feature of several PA diseases. In the case of *ARX*, the PA mutant mouse model has been reported to have normal nuclear localisation of mutant protein in some tissues, but to have a reduction in the total number of positive cells [17,18]. Given our data, an alternative interpretation is that cells expressing *ARX* are not lost but are unable to be detected due to misfolding and clearance of the mutant protein. In situ hybridisation of *Arx* transcripts should provide a means of discriminating between these possibilities.

In addition to *SOX3-26ala*, XH is also associated with a 7 alanine expansion mutation (*SOX3-22ala*) in the same polyalanine tract [16,19,20]. Although *SOX3-22ala* is occasionally associated with mild learning difficulties [20], it is interesting to note that the infantile behaviour/severe MR that is fully penetrant in *SOX3-26ala*-carrying males is not found in affected males with the *SOX3-22ala* expansion mutation, suggesting that the shorter expansion is less severe. Given our data showing a massive depletion of *SOX3* protein in *Sox3-26ala* mutant cells *in vivo*, we would predict that *SOX3-22ala* protein would also be reduced (compared to WT) but to a lesser extent than the *SOX3-26ala* protein. This higher level (and therefore activity) of *SOX3-22ala* compared to *SOX3-26ala* would be sufficient for “cognitive” CNS development. Consistent with this idea, the *SOX3-22ala* protein is less prone to aggregate formation in cell culture [7,16], although one must be cautious in

interpreting pathological mechanism from over expression studies in heterologous cell lines. While it would be useful to directly compare SOX3-26ala and SOX3-22ala levels *in vivo*, a mouse model of the SOX3-22ala mutation has not been generated. However, given that it is now possible to generate neural precursors directly from patients via iPS cells, it would be interesting to compare the SOX3 protein levels in neuroprogenitors derived from SOX3-22ala and SOX3-26ala carrying males. This approach might also be informative for PA disorders in general, particularly given that increased disease severity is generally associated with longer PA expansions [21].

Based on these and other published data, we propose that both GOF and LOF PA disease alleles are associated with a primary defect in protein folding but that a critical difference in the capacity of cells to clear mutant protein results in either the accumulation of mislocalised cytoplasmic protein (GOF) or a diminished level of functional nuclear protein (LOF). As indicated by earlier reports [21], the former mechanism provides scope for dominant-negative and toxic GOF through cytoplasmic sequestration of endogenous binding partners. In the latter, we propose that misfolded protein is efficiently cleared by the cell without significant perturbation of other cellular processes, manifesting as LOF or partial LOF, depending on the amount of residual functional protein. Factors that determine whether or not a cell is able to efficiently clear the misfolded protein could include the level of expression, local concentration differences within the cell, the efficiency of the degradation pathway within different affected cell types and the length of the PA expansion. We propose that the aggregates seen when SOX3-26ala and all other PA expansion proteins are expressed *in vitro* reflect the overloading of the cell with mutant protein such that normal degradation pathways are overwhelmed. Resolution of the factors that determine whether a cell is able to clear or tolerate mutant protein will have broader implications for proteinopathies such as Alzheimers Disease and Huntingtons Disease in which only subsets of cells display aggregation. For the future, it will be interesting to determine the behaviour of mutant protein in other PA mouse models with GOF and LOF inheritance and, where possible, in patient-derived induced pluripotent stem (iPS) cell derivatives.

Methods

ES cell targeting

SOX3-26ala allele. The targeting vector was based on a *Sox3-Neo* floxed vector published previously in which GFP and the loxP sites were removed and a 36 bp insertion was introduced to the existing 42 bp alanine encoding tract [15]. This was electroporated into 129 strain-derived (R1) mouse ES cells and 1000 G418 resistant clones were screened by Southern blotting. Integration at the *Sox3* locus was determined by a shift in a *BglII* fragment from 8.8 kb to 5.9 kb when probed with a 5' *Sox3* probe [15]. Clones were subsequently screened for the alanine expansion using PCR primers flanking the alanine tract; 5'-AGACGCTGCTCAA-GAAGGAC-3' and 5'-CTGCACGAGCGAGTAGGC-3'. Clones targeted with the *Neo* cassette but lacking the *Sox3*-26ala expansion were designated (*Neo*).

Null allele. The targeting vector for generation of the null allele was the *Sox3-Neo* floxed vector published previously [15]. 400 clones were screened using the above probe and homologous recombinants identified by a shift in the *BglII* fragment from 8.8 kb to 7.5 kb. Two correctly targeted clones were transiently transfected with a Cre recombinase expression plasmid (kind gift from Duncan Hewett), seeded at clonal density and screened by PCR for the absence of *Sox3*. Putative *Sox3* null clones were confirmed by Southern blotting (5 kb *BglII* fragment using the 5' *Sox3* probe, as above).

Chimera generation

129 derived ES clones were injected into c57/Bl6xB6D2F1 2.5 dpc morula, cultured overnight and transferred as blastocysts into the uterus of 2.5 dpc pseudopregnant recipients.

ES cell culture and neurodifferentiation

R1 ES cells were maintained on irradiated MEFs in standard conditions and neurodifferentiated either in aggregates in chemically defined media as described in [22] referred to as (CDM) or as monolayers as previously described [23], referred to as (N2B27).

qPCR gene expression and estimation of percentage chimerism

RNA was prepared with Trizol (Invitrogen) and reverse transcribed with a High Capacity RNA-to-cDNA kit (ABI). qPCR was performed using Fast SYBR Green Master Mix (ABI) and run on an ABI 7500 StepOnePlus system. Primer sequences and lengths of amplified products were: *Sox3* (117 bp) 5'-GAACGCATCAGGT-GAGAGAAG-3' and 5'-GTCCGAGTGGTGCTCAGG-3', β -*Actin* (89 bp) 5'-CTGCCTGACGGCCAGG-3' and 5'-GATTC-CATACCCAAGAAGGAAGG-3'. *Sox3* expression was normalised to β -*Actin* and expressed as relative quantity (RQ) using ABI software. Embryonic chimerism was determined against a standard curve of *Sox3* dosage generated from adult gDNA of WT and *Sox3*-null mixed 0:100, 25:75, 50:50, 100:0. Loading was corrected using primers against *Sox1* (171 bp) 5'-GACTTGCAGGCTATGTACAACATC-3' and 5'-CCTCTCAGACGGTGGAGTTATATT-3' and *Ngn3* (120 bp) 5'-CCCCAGAGACACAACAACCT-3' and 5'-AGT-CACCCACTTCTGCTTTCG-3'. Chimerism in Figure 2C was estimated by NEO IF and ISH.

ISH and IF

Embryos were fixed in 4% PFA overnight and CDM bodies were fixed for 15 minutes. Both were then equilibrated in 30% sucrose overnight, set in OCT compound and cryosectioned on a Leica CM1900 at 10 μ m. ISH on fixed frozen sections was performed using a *Sox3* probe from a fragment generated with the following primers; 5'-AGCGCCTGGACACGTACAC-3' and 5'-AGCGCCTGGACACGTACAC-3' and a *Neo* probe from a fragment amplified with the following primers; 5'-GATC-GATCCCCCTCAGAAGAAC-3' and 5'-GGCTATTCCGGC-TATGACTGG-3'. Images were captured on a Zeiss Axiophot upright microscope with AnalySIS software, using a 2.5 \times (Zeiss Neofluar; NA0.5) or a 10 \times (Zeiss Neofluar; NA0.3) objective lens. Primary antibodies and dilutions were goat anti-hSOX3 (R&D #AF2569,1:150) and rabbit anti-NTPII (Millipore #06-747, 1:150). Imaging of IF was performed on a Zeiss Axioplan2 upright microscope with Axiovision software, using a 20 \times (Zeiss Apochromat; NA0.75) or a 100 \times (Zeiss Apochromat; NA1.3) objective lens.

Western blotting

Day4 N2B27 neurodifferentiated ES cells were lysed in RIPA buffer supplemented with protease inhibitors. SDS loading buffer was added and the samples were resolved by 10% SDS-PAGE. Duplicate gels were subjected to immunoblot analysis using anti-hSOX3 (1:2500) and anti- α -Tubulin (Sigma #T8203, 1:2500) antibodies since both detect proteins at ~40 kD. Signals were developed using ECL substrate (West Pico, Pierce). For nuclear protein lysates, cells were subjected to hypotonic lysis (10 mM HEPES pH 7.9, 1.5 mM MgCl₂, 10 mM KCl, 0.4% NP-40, 10% Ficoll-400, 1 mM DTT, 1 mM PMSF, 1 \times protease inhibitors), centrifugation to pellet nuclei that were then washed in wash

buffer (10 mM HEPES pH 7.9, 1.5 mM MgCl₂, 150 mM KCl, 10% Ficoll-400, 1 mM DTT, 1 mM PMSF, 1× protease inhibitors) and lysed in nuclear extract buffer (20 mM HEPES pH 7.9, 1.5 mM MgCl₂, 0.5 mM EDTA, 20% glycerol, 0.42 M KCl, 1 mM DTT, 1 mM PMSF, 1× protease inhibitors).

Luciferase transcription assay

Transcriptional activities of mouse and human SOX3 and SOX3-26ala were determined using the Dual-Luciferase Reporter Assay System (Promega). 1.0 µg of plasmid DNA was transfected into COS-7 cells using Lipofectamine according to manufacturer's instructions. All transfections were performed in triplicate and contained luciferase reporter Sox Consensus Motif (SOCM; 4×AACAAAG) [7], *Renilla* luciferase plasmid pRL-CMV and one of pcDNA3.1, pcDNA3.1 hSOX3, pcDNA3.1 hSOX3-26ala, pcDNA3.1 mSOX3 or pcDNA3.1 mSOX3-26ala expression constructs (Promega). The firefly luciferase and *Renilla* luciferase activities were determined after 48 h on a FluorStar Optima (BMG technologies). Relative luciferase activity is the ratio of firefly to *Renilla* normalised to pcDNA3.1. The assay was repeated four times. Statistical analysis was performed using Student's T-test (two tailed, unequal variance).

Ethics statement

Animal experiments were approved by the University of Adelaide Animal Ethics Committee. All studies were conducted in accordance with the principles of animal replacement and reduction and experimental refinement. Animals were monitored daily for evidence of illness and, if distressed, were culled immediately by cervical dislocation by an experienced investigator/animal technician.

Supplementary methods

Cell free transcription/translation. Cell free transcription/translation was performed using the TnT Coupled Reticulocyte Lysate System (Promega, #L4610) as per manufacturer's instructions. Plasmids used were 250 ng of either pcDNA3.1 hSOX3, pcDNA3.1 hSOX3-26ala, pcDNA3.1 mSOX3 or pcDNA3.1 mSOX3-26ala. Reactions were resolved on a Bio-Rad 10% precast gel (#456-1033), transferred to PVDF membrane, exposed to phosphor-screen and scanned.

COS-7 aggregation assay. COS-7 cells were grown on glass coverslips, transfected with one of pcDNA3.1, pcDNA3.1 hSOX3, pcDNA3.1 hSOX3-26ala, pcDNA3.1 mSOX3 or pcDNA3.1 mSOX3-26ala using Eugene6 (Roche) and stained with SOX3 antibody and DAPI two days later. At least 100 cells were scored randomly from each condition. Results are shown as the mean of three independent experiments and error bars represent one standard deviation.

Supporting Information

Figure S1 Mutant SOX3-26ala protein aggregates *in vitro*. COS cells were transfected with mouse or human pcDNA3.1 SOX3 or pcDNA3.1 SOX3-26ala expression plasmids and two days later cells were fixed and stained with a SOX3 antibody (Green) and DAPI (Blue). Cells were scored according to the localisation of the

References

- Innis JW, Mortlock D, Chen Z, Ludwig M, Williams ME, et al. (2004) Polyalanine expansion in HOXA13: three new affected families and the molecular consequences in a mouse model. *Hum Mol Genet* 13: 2841–2851.
- De Baere E, Beysen D, Oley C, Lorenz B, Cocquet J, et al. (2003) FOXL2 and BPES: mutational hotspots, phenotypic variability, and revision of the genotype-phenotype correlation. *Am J Hum Genet* 72: 478–487.

SOX3 staining as nuclear, cytoplasmic or peri-nuclear for which representative examples are shown. At least 100 cells were scored from each condition and the experiment was repeated three times. Error bars show one standard deviation from the mean. Student's T-tests between WT and mutant for each category for both mouse and human revealed statistically significant differences with a P value of less than 0.05 in each case.

(TIF)

Figure S2 SOX3-26ala protein is cleared from neuroprogenitor cells *in vivo*. *Sox3*-26ala<->WT chimeras were cut at 7.5 dpc (transverse section through neural groove), 9.5 dpc (transverse section through neural tube) or 10.5 dpc (sagittal section through ventral diencephalon) and stained with antibodies against SOX3 and NEO. Mutant *Sox3*-26ala cells (NEO positive cells; arrow heads) have low SOX3 staining. Confocal microscopy was performed for 7.5 dpc and 9.5 dpc embryos on a Leica SP5 spectral scanning confocal microscope with AnalySIS software and a 100× (Leica HCL PL-APO; NA1.4) objective. 10.5 dpc embryos were imaged using a Zeiss Axioplan 2 epifluorescence microscopy as described in the methods.

(TIF)

Figure S3 SOX3-26ala expressing cells show no signs of apoptosis. Coronal section of a 13.5 dpc *Sox3*-26ala<->WT chimera stained with antibodies against SOX3 and Activated Caspase3 (BD Pharmingen, #559565, 1:1000). Boxed area in overlay shows a rare Activated Caspase3 positive cell, while the arrowhead indicates a patch of mutant cells (low SOX3) in which no Activated Caspase3 positive cells are seen.

(TIF)

Figure S4 A) Phosphor image of an *in vitro* transcription/translation reaction of mouse and human SOX3 WT and 26ala expression constructs. SOX3 protein of the expected size is indicated by the arrow. Two strong bands were generated by the mouse plasmids, possibly reflecting the inclusion of an upstream ATG in the plasmid, while human plasmids generated only one strong band at the expected size. SOX3-26ala for both mouse and human runs slightly higher than WT protein reflecting the expanded polyalanine tract. In addition, both mouse and human SOX3-26ala have a high molecular weight species that appears to be an aggregated form of SOX3 (*). (-) indicates no template control. B) Immunoblotting using anti-SOX3 antibody (R&D #AF2569) generates a near identical pattern to the phosphor image in (A), thus confirming that the antibody is able to detect mutant and wild type protein with comparable efficiency.

(TIF)

Acknowledgments

We thank Dr. Robin Lovell-Badge for providing *Sox3*-null mice and Alex Jansson and Reuben Buckley for technical assistance.

Author Contributions

Conceived and designed the experiments: PT JH SP NR DM. Performed the experiments: JH SP NR DM LR. Analyzed the data: JH PT SP NR DM LR. Wrote the paper: PT JH.

5. Brison N, Tylzanowski P, Debeer P (2011) Limb skeletal malformations - what the HOX is going on? *Eur J Med Genet* 55: 1–7.
6. Dubreuil V, Ramanantsoa N, Trochet D, Vaubourg V, Amiel J, et al. (2008) A human mutation in *Phox2b* causes lack of CO₂ chemosensitivity, fatal central apnea, and specific loss of parafacial neurons. *Proc Natl Acad Sci U S A* 105: 1067–1072.
7. Wong J, Farlie P, Holbert S, Lockhart P, Thomas PQ (2007) Polyalanine expansion mutations in the X-linked hypopituitarism gene *SOX3* result in aggresome formation and impaired transactivation. *Front Biosci* 12: 2085–2095.
8. Albrecht AN, Kornak U, Boddlich A, Suring K, Robinson PN, et al. (2004) A molecular pathogenesis for transcription factor associated poly-alanine tract expansions. *Hum Mol Genet* 13: 2351–2359.
9. Nasrallah IM, Minarcik JC, Golden JA (2004) A polyalanine tract expansion in *Arx* forms intranuclear inclusions and results in increased cell death. *J Cell Biol* 167: 411–416.
10. Calado A, Tome FM, Brais B, Rouleau GA, Kuhn U, et al. (2000) Nuclear inclusions in oculopharyngeal muscular dystrophy consist of poly(A) binding protein 2 aggregates which sequester poly(A) RNA. *Hum Mol Genet* 9: 2321–2328.
11. Davies SW, Turmaine M, Cozens BA, DiFiglia M, Sharp AH, et al. (1997) Formation of neuronal intranuclear inclusions underlies the neurological dysfunction in mice transgenic for the HD mutation. *Cell* 90: 537–548.
12. La Spada AR, Taylor JP (2010) Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nat Rev Genet* 11: 247–258.
13. Laumonnier F, Ronce N, Hamel BC, Thomas P, Lespinasse J, et al. (2002) Transcription factor *SOX3* is involved in X-linked mental retardation with growth hormone deficiency. *Am J Hum Genet* 71: 1450–1455.
14. Laronda MM, Jameson JL (2011) *Sox3* functions in a cell-autonomous manner to regulate spermatogonial differentiation in mice. *Endocrinology* 152: 1606–1615.
15. Rizzoti K, Brunelli S, Carmignac D, Thomas PQ, Robinson IC, et al. (2004) *SOX3* is required during the formation of the hypothalamo-pituitary axis. *Nat Genet* 36: 247–255.
16. Woods KS, Cundall M, Turton J, Rizzoti K, Mehta A, et al. (2005) Over- and underdosage of *SOX3* is associated with infundibular hypoplasia and hypopituitarism. *Am J Hum Genet* 76: 833–849.
17. Kitamura K, Ito Y, Yanazawa M, Ohsawa M, Suzuki-Migishima R, et al. (2009) Three human *ARX* mutations cause the lissencephaly-like and mental retardation with epilepsy-like pleiotropic phenotypes in mice. *Hum Mol Genet* 18: 3708–3724.
18. Price MG, Yoo JW, Burgess DL, Deng F, Hrachovy RA, et al. (2009) A triplet repeat expansion genetic mouse model of infantile spasms syndrome, *Arx(GCG)₁₀₊₇*, with interneuronopathy, spasms in infancy, persistent seizures, and adult cognitive and behavioral impairment. *J Neurosci* 29: 8752–8763.
19. Burkit Wright EM, Perveen R, Clayton PE, Hall CM, Costa T, et al. (2009) X-linked isolated growth hormone deficiency: expanding the phenotypic spectrum of *SOX3* polyalanine tract expansions. *Clin Dysmorphol* 18: 218–221.
20. Alatzoglou KS, Kelberman D, Cowell CT, Palmer R, Arnhold JJ, et al. (2011) Increased transactivation associated with *SOX3* polyalanine tract deletion in a patient with hypopituitarism. *J Clin Endocrinol Metab* 96: E685–690.
21. Albrecht A, Mundlos S (2005) The other trinucleotide repeat: polyalanine expansion disorders. *Curr Opin Genet Dev* 15: 285–293.
22. Wataya T, Ando S, Mugaruma K, Ikeda H, Watanabe K, et al. (2008) Minimization of exogenous signals in ES cell culture induces rostral hypothalamic differentiation. *Proc Natl Acad Sci U S A* 105: 11796–11801.
23. Ying QL, Stavridis M, Griffiths D, Li M, Smith A (2003) Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat Biotechnol* 21: 183–186.

***Dbx1* is a direct target of SOX3 in the spinal cord**

(Rogers et al, Manuscript)

Nicholas Rogers¹, Dale McAninch¹, Paul Thomas¹

¹Discipline of Biochemistry, School of Molecular and Biomedical Sciences, University of Adelaide, Adelaide, Australia

Statement of Authorship: Appendix 3

Abstract:

SoxB1 genes have been comprehensively studied within central nervous system development, with numerous studies suggesting core overlapping functions between members. Using genome wide expression analysis examining a Sox3 null neural progenitor population, we identified a number of putative Sox3 targets. Our data also identified Dbx1 as a robust Sox3 target. Dbx1 down-regulation, at both the mRNA and protein level, within Sox3 null mice at early stages of CNS development was observed. We also independently confirm a number of SOX3 binding sites surrounding Dbx1, one of which showed clear enrichment in vivo. Correlation between our putative targets and that of a previously published SOX3-ChIP set show a clear enrichment for SOXB1 binding sites near our Sox3 targets. Our data provides evidence of the individual specificity of the SoxB1 subgroup members, while concurrently supporting the idea of functional redundancy between members.

Introduction

SOX3 is a member of the SOX (Sry-related HMG box) family of transcription factors, of which 20 members have been identified in mammals. *SOX* genes generally have developmentally-regulated expression and play important roles in cell specification, self-renewal and differentiation in a broad range of embryonic contexts (Lefebvre et al., 2007). Within the developing central nervous system (CNS), *Sox3*, and the closely related genes *Sox1* and *Sox2* (which together make up the SoxB1 subgroup), are expressed in neuroprogenitor cells throughout the neuroaxis and are down regulated upon differentiation (Bylund et al., 2003; Rogers et al., 2013). Overexpression studies in chick embryonic spinal cords and cultured murine neuroprogenitor (NP) cells indicate that SOXB1 proteins function as inhibitors of neurodifferentiation and that they have overlapping roles in this process (Bergsland et al., 2011; Bylund et al., 2003). SOXB1 group functional redundancy is also supported by loss-of-function studies in mammals. *Sox3* null mice exhibit specific CNS defects within the hippocampus, corpus callosum and hypothalamus despite the widespread *Sox3* expression in NP cells throughout the developing brain (Rizzoti et al., 2004; Rizzoti and Lovell-Badge, 2007). In addition, humans with polyalanine tract expansion mutations in *SOX3* have a relatively mild phenotype that includes infundibular hypoplasia, hypothalamic-pituitary axis dysfunction and incompletely penetrance of intellectual disability (Laumonnier et al., 2002; Woods et al., 2005). CNS-specific deletion of *Sox2* or *Sox1* in mice also results in regionally-restricted defects as opposed to a general NP phenotype (Ekonomou et al., 2005; Favaro et al., 2009; Ferri et al., 2004; Malas et al., 2003; Nishiguchi et al., 1998).

While these studies provide strong support for SOXB1 functional redundancy, the existence of congenital CNS defects in *Sox3*, *Sox2* and *Sox1* single gene mutant mice also indicates that there is a specific requirement for each SOXB1 factor in a (relatively small) subpopulation of NP cells. This phenomenon can be explained by at least three possibilities. Firstly, given that subtle differences in *Sox1*, *Sox2* and *Sox3* gene expression have been identified in the developing CNS (Rogers et al., 2013; Uchikawa et al., 2011; Wood and Episkopou, 1999), it is possible that some NP cells express only one SOXB1 gene. Deletion of the single expressing SoxB1 gene in these cells would likely cause a specific developmental defect due to the complete absence of SOXB1 activity. However, while this is an attractive hypothesis,

CNS regions expressing only one SOXB1 gene have not been definitively identified. Furthermore, contrary to this possibility, structures that express more than one SOXB1 gene can be defective in single gene mutants. For example, development of the infundibulum is abnormal in *Sox3* null embryos despite expression of *Sox2* in this structure (Rizzoti et al., 2004). A second possibility is that SOXB1 proteins can bind to and regulate the same set of target genes but that each factor has a unique preference for specific targets over others due to their inherent sequence-specific DNA binding activity. In this scenario, deletion of a single SoxB1 gene would only affect the expression of SOXB1 targets for which it has a uniquely high affinity. Consistent with this idea, recent ChIP-seq analysis has revealed extensive overlap between SOX2 and SOX3 binding sites in NP cells (Bergsland et al., 2011). A third possibility is that the SOXB1 proteins are functionally identical and that changes in the dosage of SOXB1 protein, as a whole, in an NP cell will alter the expression of some target genes. A prediction of this hypothesis is that mutation of individual SoxB1 genes would result in near identical or at least similar phenotypes. However, on the contrary, the phenotypes of single gene mutations are quite distinct and have few overlapping features.

To explore the mechanism of SOXB1 functional redundancy, and to identify genes that are most sensitive to loss of the *Sox3* gene, we performed genome wide expression profiling of *Sox3* null NP cells. Nineteen genes with abnormal/delayed expression were identified that included the homeobox gene *Dbx1*. *In vivo* analysis of *Sox3* null embryos revealed that *Dbx1* was significantly reduced in the neural tube and developing brain and that SOX3 bound directly to conserved elements associated with this gene in cultured NP cells and *in vivo*. These data define *Dbx1* as a direct SOX3 target gene whose expression, intriguingly, is not fully rescued by other SOXB1 transcription factors, suggesting that there are inherent differences in SOXB1 protein activity.

Methods:

ES cell generation

Sox3 null embryonic stem cells were previously targeted as described in (Hughes et al 2013). *Sox3* null mice have been generated (Rizzoti et al., 2004).

ES cell culture and Neurodifferentiation

R1 ES cells were maintained on irradiated MEFs in standard conditions and neurodifferentiated as monolayers as previously described (Ying et al., 2003), referred to as (N2B27).

Microarray and qRT-PCR

RNA, from differentiated cells, was extracted using RNeasy mini kit (Qiagen). cDNA was reverse transcribed with High Capacity RNA-to-cDNA kit (ABI). Four wild type and four Sox3 null day 4 samples were used on **GeneChip Mouse Gene 1.0 ST Arrays**, performed by the Adelaide Microarray Facility. 2 way-ANOVA using batch as a factor was used to identify the significantly regulated genes. qRT-PCR conditions as previously described (Hughes et al., 2013)

Immunohistochemistry

IHC was performed as described previously (Rogers et al., 2013).

Chromatin Immune Precipitation (ChIP)

In vitro ChIP was performed as per (Bergsland et al., 2011) using 5 ug of SOX3 or IgG antibodies, with Dynabeads blocked over night with 200ug BSA, 200ug yeast tRNA and 200ug Glycogen. In Vivo ChIP was performed with wild type dissected 10.5dpc heads, disassociated with scalpel blades and fixed for 10 minutes using the above protocol.

Ethics Statement

Animal experiments were approved by the University of Adelaide Animal Ethics Committee. All studies were conducted in accordance with the principles of animal replacement and reduction and experimental refinement. Animals were monitored daily for evidence of illness and, if distressed, were culled immediately by cervical dislocation by an experienced investigator/animal technician.

Results:

Sox3 expression during in vitro neural progenitor differentiation

To begin to identify SOX3 target genes in NP cells, we utilised the N2B27 culture system, which has previously been shown to generate NP cells from ESC with high efficiency (Ying et al., 2003). To characterise *Sox3* expression during N2B27 neuroinduction, we performed qRT-PCR (Fig. 1A) and immunohistochemistry analysis (Fig. 1B;(Rogers et al., 2013)). At the mRNA level, *Sox3* expression was virtually undetectable in ESC (day 0). *Sox3* expression was upregulated at day 3, peaked at day 4, and remained relatively high until day 7. Consistent with these data, weak expression of SOX3 protein was evident in day 2 cultures. By day 4, robust expression of SOX3 was detected in most cells. At day 6, robust expression of SOX3 could still be detected in most cells, although some cells (particularly those at the periphery of colonies) had begun to downregulate SOX3.

To investigate the impact of *Sox3* deletion on NP differentiation, we performed N2B27 neuroinduction using our previously-generated *Sox3* null ES cells that contain a GFP reporter knock in allele (Hughes et al., 2013). FACS analysis of Day 4 NP cultures revealed that the majority (over 60%) of cells were GFP+, which is comparable to the proportion of cells that are SOX3+ in the WT NP cultures (Figure 1B,C). To determine whether the loss of *Sox3* has a major impact on induction of neural progenitors (including other SoxB1 genes), expression of neural progenitor markers was analysed by qRT-PCR. No significant difference in *Sox1*, *Sox2* or *Pax6* expression was detected at day 4 (Fig. 1D). No change in expression of the other SoxB1 subgroup members was also observed on day 6 (data not shown). As expected, *Sox3* expression was not detected in the mutant samples (Fig. 1E). Together these data indicate that *Sox3* is rapidly upregulated during N2B27 neurodifferentiation and that loss of *Sox3* does not overtly affect NP induction across this timeframe, probably due to functional redundancy with other SOXB1 proteins, although *Sox1* and *Sox2* have not been upregulated at the mRNA level.

Identification of SOX3 target genes

To identify putative SOX3 target genes, we compared the global gene expression profile of day 4 NP cell cultures derived from *Sox3* null cells and a control cell line containing a conditional *Sox3* allele (Hughes et al., 2013). Microarray analysis was performed in quadruplicate using 4 biological replicates from two independent experiments. A total of 19 genes had significantly ($p \leq 0.05$) altered expression with a

fold change of 1.4 or more, as determined by batch corrected two-way ANOVA analysis (Table 1). Fifteen and 4 genes were upregulated and downregulated, respectively. *Sox3* provided an internal control for this experiment and, as expected, this gene was significantly downregulated in the *Sox3* null samples. Further validation of a subset of the misregulated genes was conducted using qRT-PCR on independent biological samples. *Dbx1* ($p=0.049$) and *Fezf2* ($p=0.0228$) showed significant down regulation in mutant cells (Fig 2A). Significantly elevated expression of *Cp* ($p=0.0079$), *Efnb3* ($p=0.0229$), *Cspg5* ($p=0.0251$), *Fam174b* ($p=0.0065$), *Ddr2* ($p=0.0386$) and *Slit1* ($p=0.0456$) was also confirmed in the independent samples. This data identifies a number of potential SOX3 targets in cultured NP cells.

Correlation between SOX3 binding sites and *Sox3* targets

Differentially expressed genes in *Sox3* null NPs potentially represent genes that are direct SOX3 targets. To investigate this possibility, we examined the differentially expressed genes for SOX3 binding sites using a previously published SOX3 ChIP-seq data set generated using a comparable NP differentiation system (Bergsland et al., 2011). A significant enrichment of SOX3 binding sites was present in genes with altered expression in *Sox3* null NP cells. Over 83% of the differentially expressed genes (Table 1) have at proximal (closest gene) SOX3 binding site (Fig 2C). In contrast, SOX3 binding sites were detected in only 25% of 20 randomly selected genes that do not display altered expression in *Sox3* null NPs. The enrichment of SOX3 binding sites near genes with significantly altered expression genes suggests that many of these genes are direct targets of SOX3.

In vivo validation of SOX3 targets

Next, we investigated whether the differentially-expressed NP culture genes were affected by the loss of *Sox3* in vivo. This analysis was performed using 9.5 dpc embryos (somite stages 12-15) as embryos at this stage of development have an abundance of NP cells and a paucity of differentiating neurons. In addition, like day 4 NP cell cultures, *Sox1* and *Sox2* expression in not significantly altered in *Sox3* null embryos at this stage (Fig. 3C). qRT-PCR analysis of *Slit1* and *Fezf2* showed no significant difference in expression between WT and *Sox3* null embryos at this developmental time point (Fig 3C). In contrast, *Dbx1* expression was significantly lower in *Sox3* null embryos at 9.5 dpc (Fig. 3A), consistent with its downregulation in

Sox3 null NP cells. A significant reduction in *Dbx1* expression was also evident in 10.5 dpc embryonic heads (Fig 3).

Differences in gene expression between WT and mutant cells could be due to altered expression levels in NP cells. An alternate explanation for the gene expression differences could be due to a reduction (or increase) in the number of cells expressing the differentially expressed gene. To investigate this issue, we compared the number of DBX1+ cells in day 4 NP cultures using immunohistochemistry (Figure 4 A). DBX1+ cells were readily identified in control cultures and were located in a subpopulation of SOX3+ cells located at the periphery of cell clusters. In contrast, the number of DBX1 cells was significantly and dramatically reduced in SOX3 KO cultures (Figure 4 B). To determine whether there is a similar reduction in DBX1+ cells in *Sox3* mutant embryos, we compared DBX1 expression in 9.5 dpc transverse sections through the trunk region. Again, a clear reduction in the number of DBX1+ cells within the spinal cord of *Sox3* null mice was evident (Fig 4C). Together these data suggest that *Dbx1* is regulated by SOX3 *in vivo* and that other SoxB1 factors cannot fully compensate for the loss of SOX3.

SOX3 binds to the *Dbx1* locus *in vivo*

To investigate whether *Dbx1* is directly targeted by SOX3 in day 4 NP cells, we performed ChIP-PCR on five SOX3 binding sites located in or near the *Dbx1* locus that were previously identified by ChIP-seq analysis ((Bergsland et al., 2011); Fig 5A). Significant ($p \leq 0.05$) enrichment of all five sites was detected by SOX3 ChIP-PCR, when compared to the IgG isotype control (Fig 5B). To determine if SOX3 bound these sites *in vivo*, we analysed SOX3 binding to one of the most enriched sites (Intronic site 2) using chromatin extracted from 10.5 dpc embryonic heads. Significant enrichment ($p \leq 0.001$) of this SOX3 binding this site was observed in comparison to the IgG negative control (Fig 5C). Together, these data suggest that *Dbx1* is a direct target of SOX3 *in vivo*.

DISCUSSION

Despite the widespread expression of *Sox3* in progenitor cells throughout the neuroaxis, the CNS phenotypes of mice and humans with *Sox3* mutations are relatively mild. It is generally believed that other SOXB1 proteins (SOX1 and SOX2) can function redundantly with SOX3 to “rescue” the SOX3 deficiency. However, the lack of a complete phenotypic rescue suggests that, at least in some NP cells, the *level* of SOXB1 protein may be important for neurodevelopment. Alternatively, each of the SOXB1 factors could bind preferentially to a unique subset of SOXB1 targets. As a result, target gene expression resulting from the loss of a single SOXB1 gene would only be partially rescued by the other two SOXB1 factors. To explore this issue in more detail, we attempted to identify genes that were sensitive to the loss of SOX3 in NP cells. To do this we used a previously targeted *Sox3* null ES cell line in which the single SOX3 exon was replaced with GFP (Hughes et al., 2013). N2B27-induced generation of conditionally targeted and *Sox3* null NP cells resulted in rapid upregulation of *Sox3/GFP* at the mRNA and protein level. The NP markers *Sox1* and *Pax6* were also rapidly induced and showed comparable kinetics in the control and mutant cultures. In addition, expression of *Sox2* was comparable between the two populations. Notably, no overt defects in the morphology or global gene expression (see below) were identified in *Sox3* null NP cells. This is likely due to the expression of *Sox1* and *Sox2* in N2B27-induced NP cells and is consistent with the mild CNS phenotype of *Sox3* null embryos.

To maximise the likelihood of identifying genes that are directly regulated by SOX3, we performed gene expression profiling at day 4, the earliest time point at which high levels of *Sox3* were expressed in the majority of NP cells. Of over 24000 genes assayed by microarray, only 19 had significantly different expression levels at day 4. Expression differences (when tested by qRT-PCR) were validated on independent samples, for a majority of the targets, thereby confirming the authenticity of the microarray data. Although, of genes tested, not all differentially expressed genes identified by the microarray validated on independent samples (*Tmem163*, *Ednrb*, *Slc44a5* and *Gstm6*, data not shown). The relatively small number of differentially expressed genes is interesting to consider in the context of a recently-published ChIP-seq analysis of a similar NP cell type, which showed that SOX3 binds at thousands of genomic sites, many of which are likely to have

regulatory functions (Bergsland et al., 2011). Preliminary ChIP-seq analysis from our laboratory also indicates that SOX3 binds many of these sites in day 4 N2B27-induced NP cells (D.M. and P.T. unpublished data). The most parsimonious explanation for the small number of differentially expressed genes is that SOX1 and/or SOX2 (both of which are expressed by day 4 NP cells) can bind and functionally compensate for loss of SOX3. Consistent with this hypothesis, almost all SOX2 binding sites in NPs also bind SOX3 (Bergsland et al., 2011). However, while SOXB1 redundancy provides an explanation for the unaltered expression of most genes associated with SOX3 binding sites, it also suggests that differentially expressed genes cannot be fully compensated by SOX1 and SOX2. Given the HMG domain of SOXB1 proteins is not identical, it is possible that they have subtly different preferences for sequence-specific binding in vivo, as has been suggested by in vitro DNA binding experiments (Collignon et al., 1996). In addition, amino acid differences within the HMG box (or outside of it) may influence interactions with DNA binding partners such as the POU-class transcription factors (Kamachi et al., 2000; Kondoh and Kamachi, 2010). Regardless of the mechanism, it is important to note that the differentially expressed genes are highly enriched for SOX3 binding sites (over randomly selected genes without significantly different expression levels), supporting the idea that they represent direct SOX3 targets. As a number of these putative direct target genes are upregulated, it appears that that SOX3 could function as a repressor in NP cells (Acloque et al., 2011) in addition to its more established role as an activator (Bylund et al., 2003).

Interestingly, many of the genes identified by the microarray analysis have established roles in neurodevelopment. Putative target *Cspg5* has been associated with intellectual disability disorders (Zhang et al., 2013), and ectopic *Slit1* expression has been published in mice with abnormal corpus callosum development (Amaniti et al., 2013). Intellectual disability is present in a subset of individuals with *SOX3* mutations (Laumonnier et al., 2002), while *Sox3* null mice have abnormal corpus callosum development (Rizzoti et al., 2004) providing a strong link between these putative targets and published *Sox3* function.

Another putative SOX3 target was the downregulated transcription factor gene *Fezf2*. *Fezf2* has been shown to be important for the early development of the posterior hypothalamus/thalamus in zebrafish/mouse (Hirata et al., 2006; Wolf and Ryu, 2013) *Fezf2* expression has been previously published to be downstream of the SOXC group genes *Sox4* and *Sox11*, double deletion of which leads to a loss of *Fezf2* expression. Regulation of *Fezf2* by SOX4/11 gene is mediated by a *cis* acting enhancer binding site which, interestingly, was also identified as a SOX3 binding site in NP cells by ChIP-seq and independently within our lab (Data not shown (Bergsland et al., 2011; Shim et al., 2012)). These data suggest that, in addition to acting as a pioneer factor (Bergsland et al., 2011), *Sox3* may directly regulate *Fezf2* expression in NP cells.

***Dbx1* is a direct SOX3 target in vivo**

To investigate whether NP targets were dysregulated *in vivo*, we compared their expression by qRT-PCR using whole 9.5 dpc embryos. Of the 7 putative targets analysed (Fig 3C, data not shown), only one (*Dbx1*) had significantly different expression matching the microarray result. The lack of *in vivo* validation for other targets may reflect a phenotypic or stage-specific difference in the NP cells induced by N2B27 differentiation compared with their 9.5 dpc embryonic counterparts. Consistent with this idea, expression analysis of *Sox3* targets in day 6 NP cells revealed that many genes no longer had significantly different expression levels or had greatly reduced expression fold changes (Supp. Fig. 1). *Dbx1* expression, however, was reduced in whole 9.5 dpc embryos lacking *Sox3*. In addition, DBX1+ cells were not detected in the spinal cord of *Sox3* null embryos at 9.5 dpc and were massively reduced in day 4 null NP cells. Further, SOX3 bound to evolutionarily conserved elements at the *Dbx1* locus in cultured NP cell and embryos. Together these data indicate that *Dbx1* is directly activated by SOX3 in NP cells.

Dbx1 encodes a homeoprotein that is transiently expressed in Vo interneuron precursors and required for Vo interneuron differentiation (Pierani et al., 2001). The lack of DBX1 expression in the spinal cord of *Sox3* null embryos is the first clear evidence that there is a phenotype within the developing spinal cord of *Sox3* null mice and raises the possibility that generation of Vo interneurons is abnormal in *Sox3* mutants (Rizzoti et al., 2004). The significantly reduced DBX1 expression in

Sox3 null spinal cord NPs also provides some molecular insights into SOXB1 functional redundancy. Given that the spinal cord NPs have virtually identical SOXB1 expression (Bylund et al., 2003; Wood and Episkopou, 1999), it appears that SOX1 and SOX2 are unable to substitute for SOX3 in the context of *Dbx1* regulation. As discussed above, this might reflect preferential binding of SOX3 at the *Dbx1* locus. Alternatively, *Dbx1* expression could be sensitive to the overall dosage of SOXB1 protein. At least two experimental approaches could be used to test these possibilities. Firstly, it would be interesting to examine *Dbx1* expression in *Sox1* and *Sox2* single gene KO NP cells ie. under conditions where the overall SOXB1 dosage is lower but SOX3 is present. Secondly, a targeted gene replacement strategy (eg. replacing *Sox3* with *Sox2*) could be used to generate NP cells that have an equivalent SOXB1 dosage but which lack SOX3. In addition to the spinal cord, we identified a reduction in *Dbx1* expression in the head of 10.5 dpc *Sox3* null embryos. However, we were unable to determine whether this reflected a lack of *Dbx1* expression in a specific region or general reduction in the *Dbx1* expression level (data not shown).

A recent report by Oosterveen et al also indicated that *Dbx1* is a direct target of *Sox3* (Oosterveen et al., 2013). However, the SOX3 binding sites reported by Oosterveen et al are contained within a putative cis-regulatory module (CRM) that does not overlap with the SOX3 binding sites identified in this study. The Oosterveen et al data raises the possibility that (the lack of) SOX3 binding at the CRM sites may contribute to the reduction of *Dbx1* expression in *Sox3* null NPs. The data also suggest that the regulation of *Dbx1* by SOX3 is complex and subtle differences between early or later expression or spinal cord/brain expression might be regulated any number of these sites. Further studies are needed to address this issue. Given the increasing ease with which the genome can be manipulated, it would be interesting to perform systematic mutagenesis of all known SOXB1 sites at the *Dbx1* locus and assess their impact on *Dbx1* expression *in vitro* and *in vivo*.

Acknowledgements: We thank Dr. Alessandra Pierani and Prof. Robin Lovell-Badge for provision of the DBX1 antibody and *Sox3* mutant mice, respectively. PT is a Pfizer Australia Research Fellow.

References:

- Acloque, H., Ocana, O. H., Matheu, A., Rizzoti, K., Wise, C., Lovell-Badge, R. and Nieto, M. A.** (2011). Reciprocal repression between Sox3 and snail transcription factors defines embryonic territories at gastrulation. *Dev Cell* **21**, 546-58.
- Amaniti, E. M., Hasenpusch-Theil, K., Li, Z., Magnani, D., Kessar, N., Mason, J. O. and Theil, T.** (2013). Gli3 is required in Emx1+ progenitors for the development of the corpus callosum. *Dev Biol* **376**, 113-24.
- Bergsland, M., Ramskold, D., Zaouter, C., Klum, S., Sandberg, R. and Muhr, J.** (2011). Sequentially acting Sox transcription factors in neural lineage development. *Genes Dev* **25**, 2453-64.
- Bylund, M., Andersson, E., Novitsch, B. G. and Muhr, J.** (2003). Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nat Neurosci* **6**, 1162-8.
- Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N. and Lovell-Badge, R.** (1996). A comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development* **122**, 509-20.
- Ekonomou, A., Kazanis, I., Malas, S., Wood, H., Alifragis, P., Denaxa, M., Karagogeos, D., Constanti, A., Lovell-Badge, R. and Episkopou, V.** (2005). Neuronal migration and ventral subtype identity in the telencephalon depend on SOX1. *PLoS Biol* **3**, e186.
- Favaro, R., Valotta, M., Ferri, A. L., Latorre, E., Mariani, J., Giachino, C., Lancini, C., Tosetti, V., Ottolenghi, S., Taylor, V. et al.** (2009). Hippocampal development and neural stem cell maintenance require Sox2-dependent regulation of Shh. *Nat Neurosci* **12**, 1248-56.
- Ferri, A. L., Cavallaro, M., Braida, D., Di Cristofano, A., Canta, A., Vezzani, A., Ottolenghi, S., Pandolfi, P. P., Sala, M., DeBiasi, S. et al.** (2004). Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. *Development* **131**, 3805-19.
- Hirata, T., Nakazawa, M., Muraoka, O., Nakayama, R., Suda, Y. and Hibi, M.** (2006). Zinc-finger genes Fez and Fez-like function in the establishment of diencephalon subdivisions. *Development* **133**, 3993-4004.
- Hughes, J., Piltz, S., Rogers, N., McAninch, D., Rowley, L. and Thomas, P.** (2013). Mechanistic insight into the pathology of polyalanine expansion disorders revealed by a mouse model for X linked hypopituitarism. *PLoS Genet* **9**, e1003290.
- Kamachi, Y., Uchikawa, M. and Kondoh, H.** (2000). Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet* **16**, 182-7.
- Kondoh, H. and Kamachi, Y.** (2010). SOX-partner code for cell specification: Regulatory target selection and underlying molecular mechanisms. *Int J Biochem Cell Biol* **42**, 391-9.
- Laumonier, F., Ronce, N., Hamel, B. C., Thomas, P., Lespinasse, J., Raynaud, M., Paringaux, C., Van Bokhoven, H., Kalscheuer, V., Fryns, J. P. et al.** (2002). Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. *Am J Hum Genet* **71**, 1450-5.
- Lefebvre, V., Dumitriu, B., Penzo-Mendez, A., Han, Y. and Pallavi, B.** (2007). Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. *Int J Biochem Cell Biol* **39**, 2195-214.
- Malas, S., Postlethwaite, M., Ekonomou, A., Whalley, B., Nishiguchi, S., Wood, H., Meldrum, B., Constanti, A. and Episkopou, V.** (2003). Sox1-deficient mice suffer from

epilepsy associated with abnormal ventral forebrain development and olfactory cortex hyperexcitability. *Neuroscience* **119**, 421-32.

Nishiguchi, S., Wood, H., Kondoh, H., Lovell-Badge, R. and Episkopou, V. (1998). Sox1 directly regulates the gamma-crystallin genes and is essential for lens development in mice. *Genes Dev* **12**, 776-81.

Oosterveen, T., Kurdija, S., Enstero, M., Uhde, C. W., Bergsland, M., Sandberg, M., Sandberg, R., Muhr, J. and Ericson, J. (2013). SoxB1-driven transcriptional network underlies neural-specific interpretation of morphogen signals. *Proc Natl Acad Sci U S A* **110**, 7330-5.

Pierani, A., Moran-Rivard, L., Sunshine, M. J., Littman, D. R., Goulding, M. and Jessell, T. M. (2001). Control of interneuron fate in the developing spinal cord by the progenitor homeodomain protein Dbx1. *Neuron* **29**, 367-84.

Rizzoti, K., Brunelli, S., Carmignac, D., Thomas, P. Q., Robinson, I. C. and Lovell-Badge, R. (2004). SOX3 is required during the formation of the hypothalamo-pituitary axis. *Nat Genet* **36**, 247-55.

Rizzoti, K. and Lovell-Badge, R. (2007). SOX3 activity during pharyngeal segmentation is required for craniofacial morphogenesis. *Development* **134**, 3437-48.

Rogers, N., Cheah, P. S., Szarek, E., Banerjee, K., Schwartz, J. and Thomas, P. (2013). Expression of the murine transcription factor SOX3 during embryonic and adult neurogenesis. *Gene Expr Patterns* **13**, 240-8.

Shim, S., Kwan, K. Y., Li, M., Lefebvre, V. and Sestan, N. (2012). Cis-regulatory control of corticospinal system development and evolution. *Nature* **486**, 74-9.

Uchikawa, M., Yoshida, M., Iwafuchi-Doi, M., Matsuda, K., Ishida, Y., Takemoto, T. and Kondoh, H. (2011). B1 and B2 Sox gene expression during neural plate development in chicken and mouse embryos: universal versus species-dependent features. *Dev Growth Differ* **53**, 761-71.

Wolf, A. and Ryu, S. (2013). Specification of posterior hypothalamic neurons requires coordinated activities of Fezf2, Otp, Sim1a and Foxb1.2. *Development* **140**, 1762-73.

Wood, H. B. and Episkopou, V. (1999). Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev* **86**, 197-201.

Woods, K. S., Cundall, M., Turton, J., Rizzotti, K., Mehta, A., Palmer, R., Wong, J., Chong, W. K., Al-Zyoud, M., El-Ali, M. et al. (2005). Over- and underdosage of SOX3 is associated with infundibular hypoplasia and hypopituitarism. *Am J Hum Genet* **76**, 833-49.

Ying, Q. L., Stavridis, M., Griffiths, D., Li, M. and Smith, A. (2003). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat Biotechnol* **21**, 183-6.

Zhang, C., Mejia, L. A., Huang, J., Valnegri, P., Bennett, E. J., Anckar, J., Jahani-Asl, A., Gallardo, G., Ikeuchi, Y., Yamada, T. et al. (2013). The X-linked intellectual disability protein PHF6 associates with the PAF1 complex and regulates neuronal migration in the mammalian brain. *Neuron* **78**, 986-93.

Figure legends:

Table 1: Table of significantly regulated genes with a fold change equal to or greater than 1.4 in Day 4 N2B27 differentiated ES cells.

Figure 1: Sox3 expression in neural progenitor differentiation. A) Time course of Sox3 gene expression levels by qRT-PCR throughout N2B27 differentiation (normalised to day 4 Sox3 expression, n=3). B) IHC of SOX3 showing expression at days 2, 4 and 6 of neuronal differentiation. C) GFP expressing cells sorted by FACS within Sox3 null N2B27 differentiated ES cells at days 2, 4 and 7. D) qRT-PCR of Sox1, Sox2, Sox2 and Pax6 expression at day 4 of N2B27 differentiation. RQ: relative quantification normalised to β actin.

Figure 2: Validation of putative Sox3 targets. A) qRT-PCR validation of gene expression levels of down regulated genes, *Dbx1* and *Fezf2*. B) qRT-PCR validation of gene expression levels of up regulated gene, *Ddr2*, *Slit1*, *Cspg5*, *Efnb3* and *Fam174b*. C) Percentage of significantly regulated genes with SOX3 binding sites, as closest gene. D) Percentage of significantly regulated genes with an overlapping SOX2 and SOX3 binding site. RQ: relative quantification normalised to β actin.

Figure 3: Validation of putative Sox3 targets *in vivo*. A) *Dbx1* gene expression within whole 9.5dpc embryos. B) *Dbx1* gene expression within 10.5dpc dissected heads. C) qRT-PCR analysis of mRNA levels of *Sox1*, *Sox2*, *Slit1* and *Fezf2* showing no significant difference. RQ: relative quantification normalised to β actin.

Figure 4: DBX1 levels within neural progenitor cells and the developing spinal cord. A) DBX1 expression at day 4 of N2B27 differentiation. B) Number of DBX1+ cell per field of view at day 4 of N2B27 differentiation. C) DBX1 expression within the developing 9.5dpc spinal cord of wild type and Sox3 null mice.

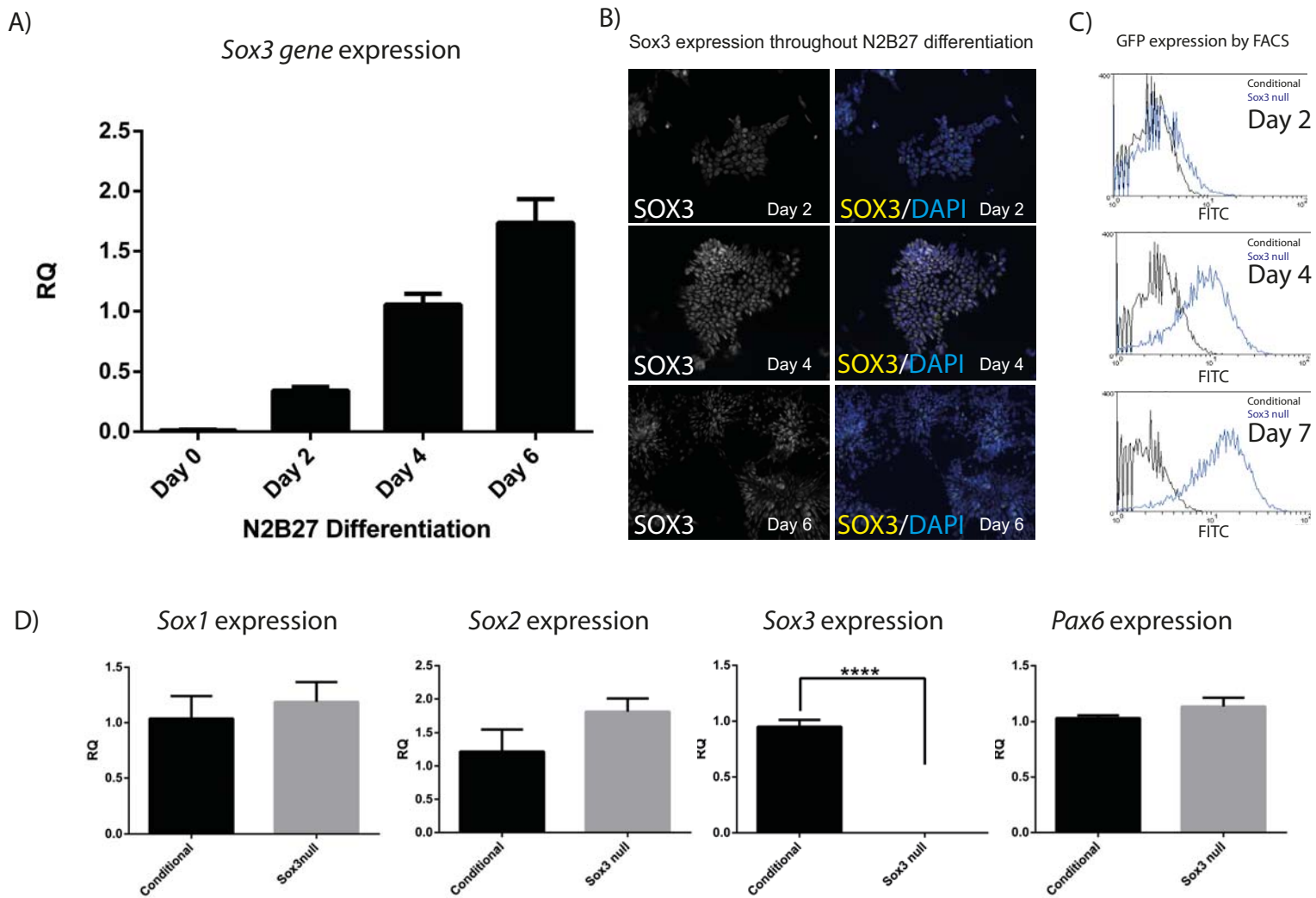
Figure 5: SOX3 binding sites surrounding DBX1. A) Schematic diagram of previously published SOX3 binding sites. B) *In vitro* validation of the five surrounding SOX3 binding sites (n=3). C) *In vivo* validation of SOX3 binding to intronic site 2 (n=3).

Supplementary Figure 1: Expression of putative Sox3 targets at day 6 of neural progenitor differentiation. A) Putative targets *Dbx1* and *Slit1* relative gene

expression levels by qRT-PCR at day 6 of N2B27 differentiation, (n=3 normalised to day 4). B) Putative targets Cp, Fezf2 and Efnb3 showing reduced and/or insignificant fold change at day 6 of N2B27 differentiation (n=3, normalised to day 4). RQ: relative quantification normalised to β actin.

Manuscript Table 1

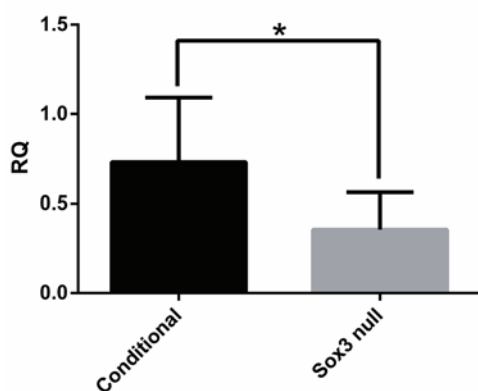
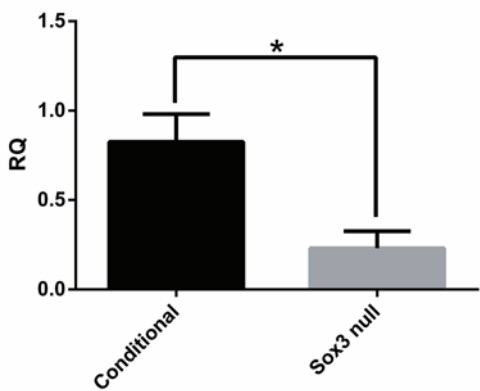
Gene Symbol	RefSeq	stepup p-value	Fold-Change	Fold-Change (Description)
Sox3	NM_009237	0.01	-4.10	KO down vs WT
Cspg5	NM_001166273	0.02	1.49	KO up vs WT
Ddr2	NM_022563	0.02	1.83	KO up vs WT
Slit1	NM_015748	0.04	1.87	KO up vs WT
Tagln3	NM_019754	0.04	1.54	KO up vs WT
120009022Rik	NM_025817	0.04	1.41	KO up vs WT
Tmem163	NM_028135	0.04	-1.45	KO down vs WT
Slc44a5	NM_001081263	0.05	1.44	KO up vs WT
Fezf2	NM_080433	0.05	-1.61	KO down vs WT
Fam174b	NM_001162532	0.05	1.64	KO up vs WT
Fgfr3	NM_008010	0.05	1.65	KO up vs WT
Flrt2	NM_201518	0.05	1.46	KO up vs WT
Dbx1	NM_001005232	0.05	-2.35	KO down vs WT
Gpr56	NM_018882	0.05	1.56	KO up vs WT
Efnb3	NM_007911	0.05	1.59	KO up vs WT
Nexn	NM_199465	0.05	1.72	KO up vs WT
Ctgf	NM_010217	0.05	1.45	KO up vs WT
Cp	NM_001042611	0.05	2.04	KO up vs WT
Ednrb	NM_007904	0.05	1.53	KO up vs WT



A)

Dbx1 expression (n=3)

Fezf2 expression (n=3)

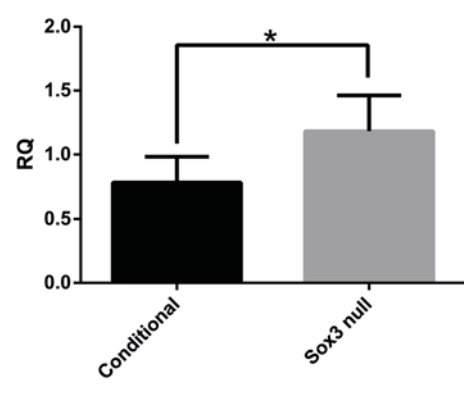
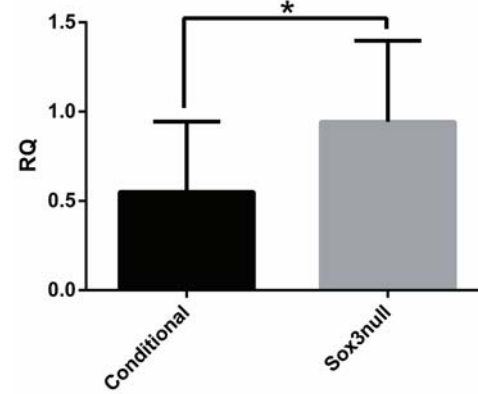
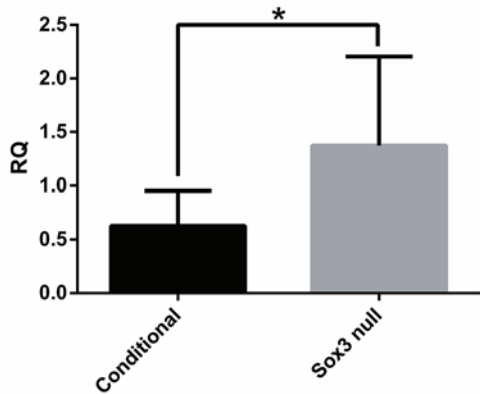


B)

Ddr2 expression (n=3)

Slit1 expression (n=3)

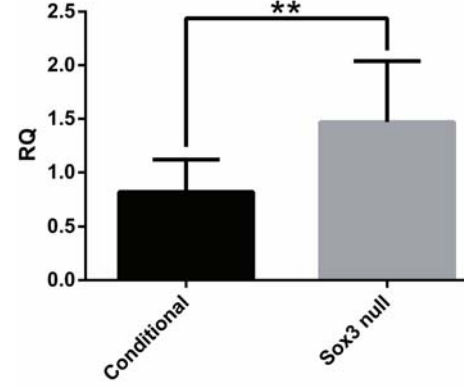
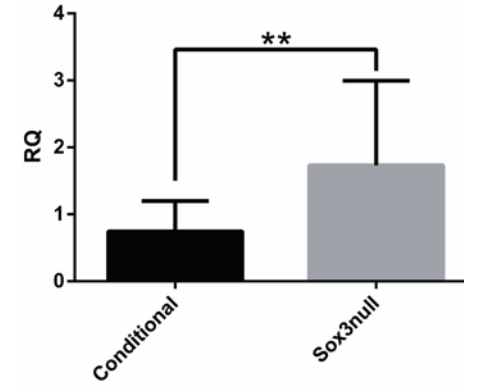
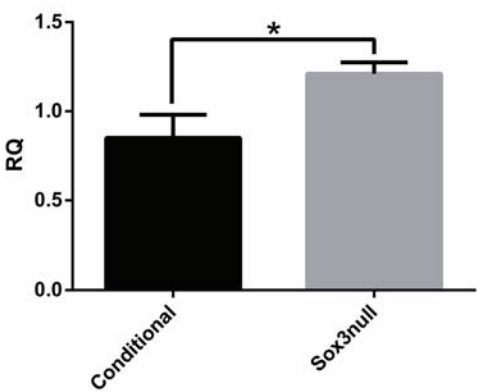
Cspg5 expression (n=3)



Efnb3 expression (n=3)

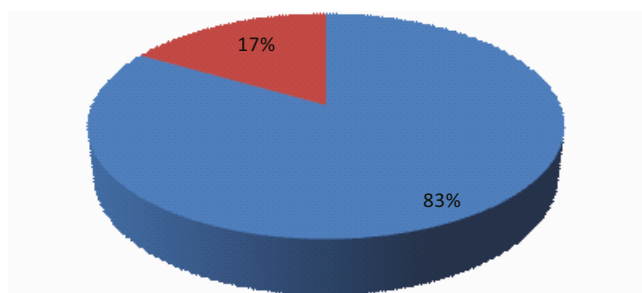
Cp expression (n=5)

Fam174b expression (n=3)



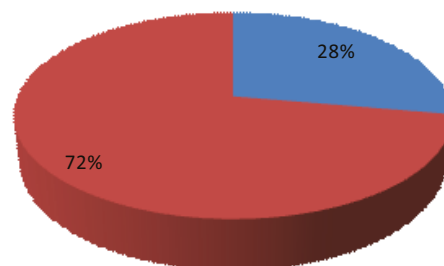
C)

■ Significantly regulated genes with SOX3 binding site(s) as the closest gene
 ■ Significantly regulated genes with no SOX3 binding site



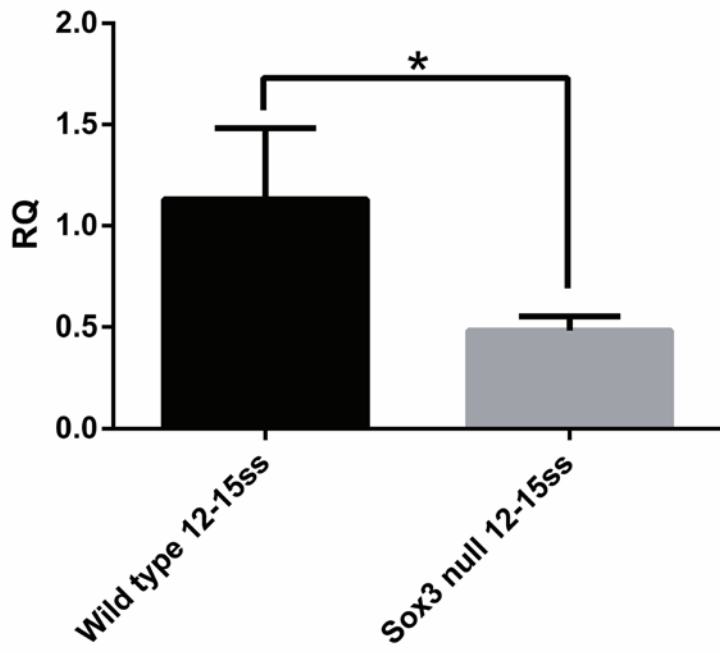
D)

■ Significantly regulated genes with an overlapping SOX2/3 binding site(s) at the closest gene
 ■ Significantly regulated genes with no SOX2/3 binding site



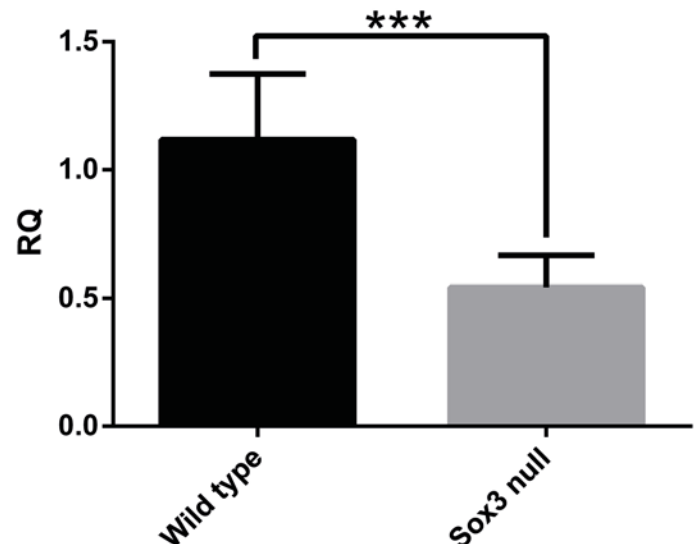
A)

Dbx1 expression (n≥3)



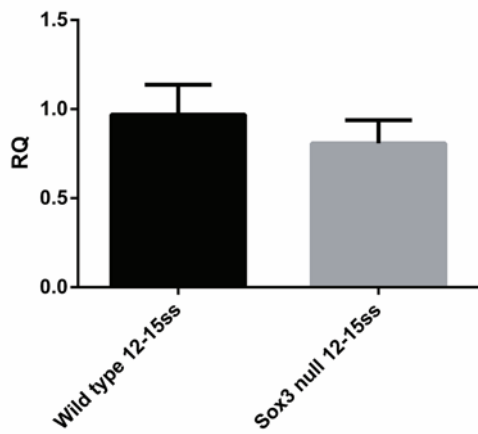
B)

Dbx1 expression (n≥5, 10.5dpc heads)

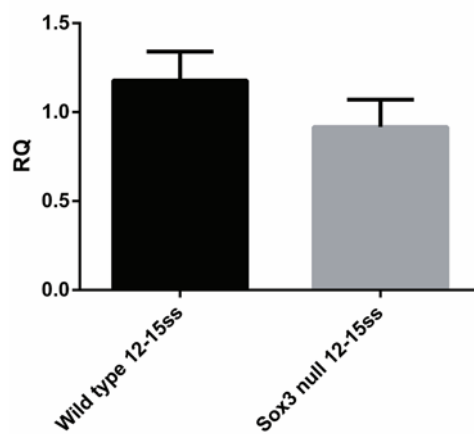


C)

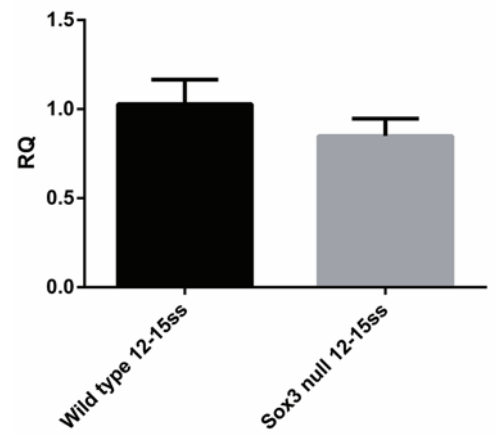
Sox1 expression (n≥3)



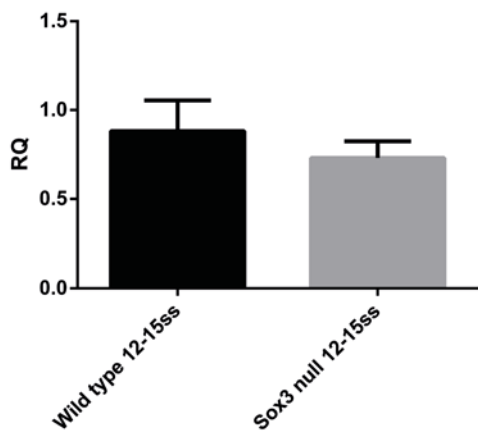
Slit1 expression (n≥3)



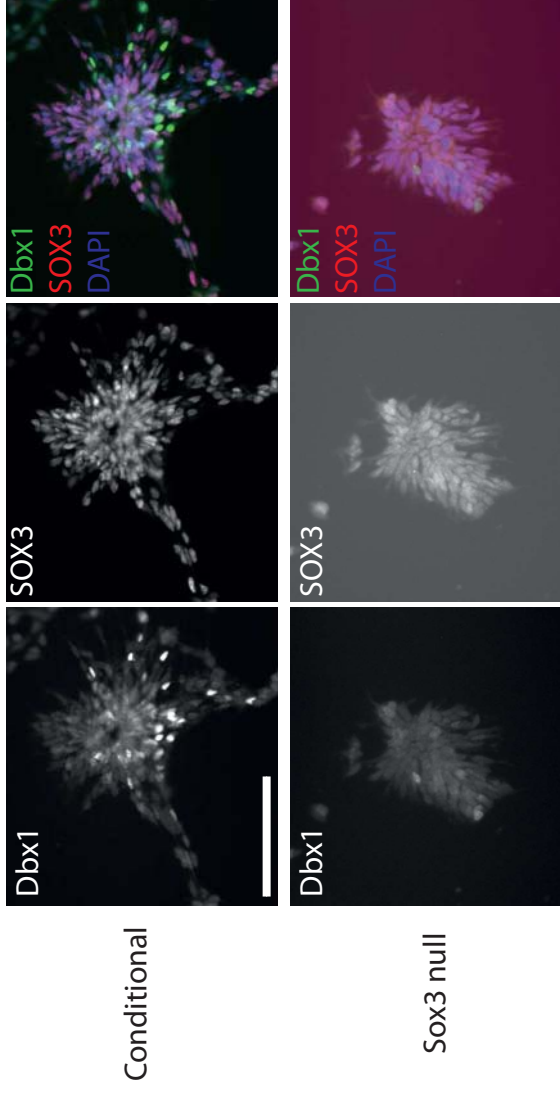
Fezf2 expression (n≥3)



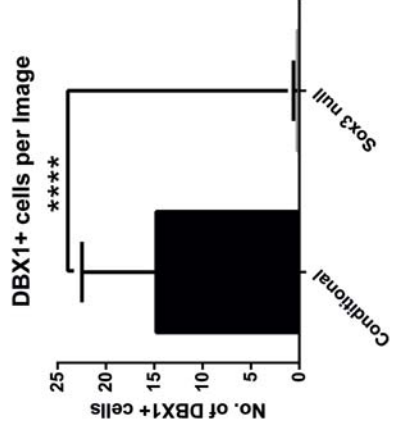
Sox2 expression (n≥3)



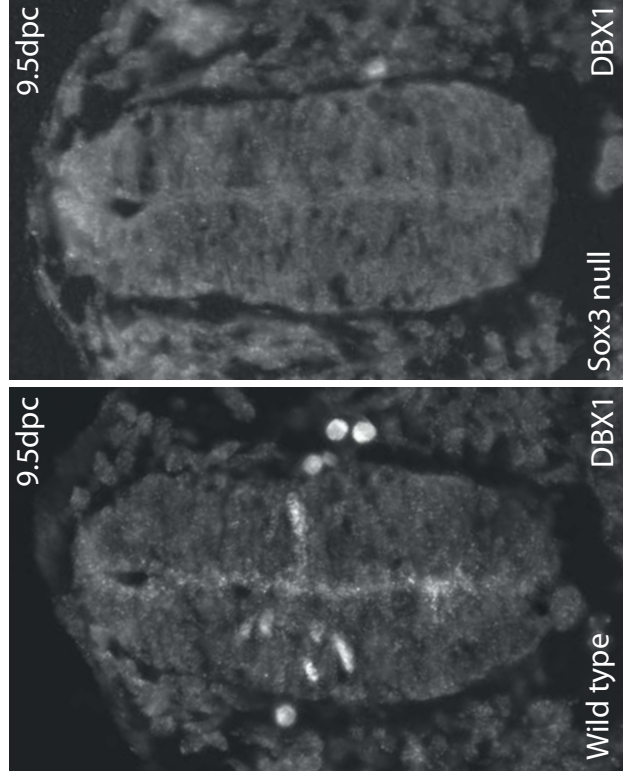
A)



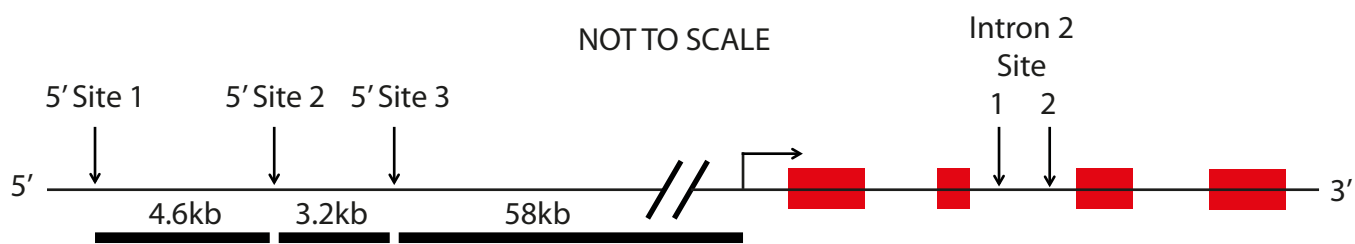
B)



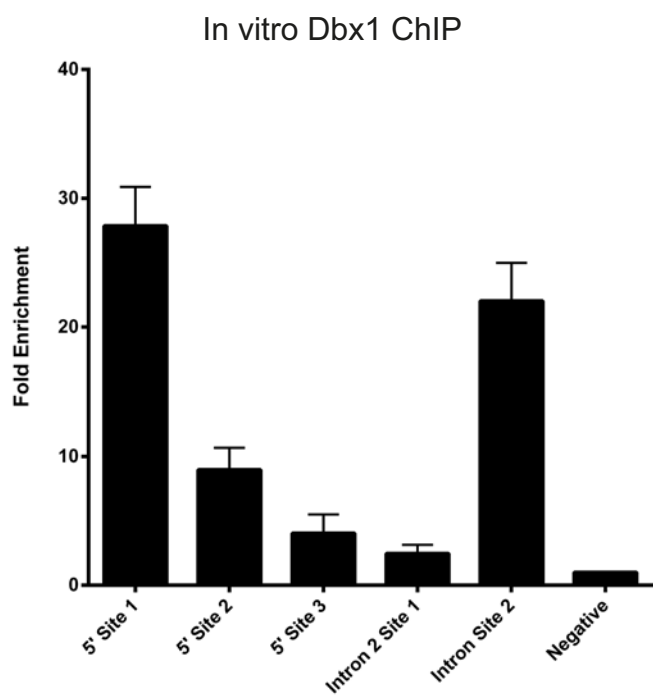
C)



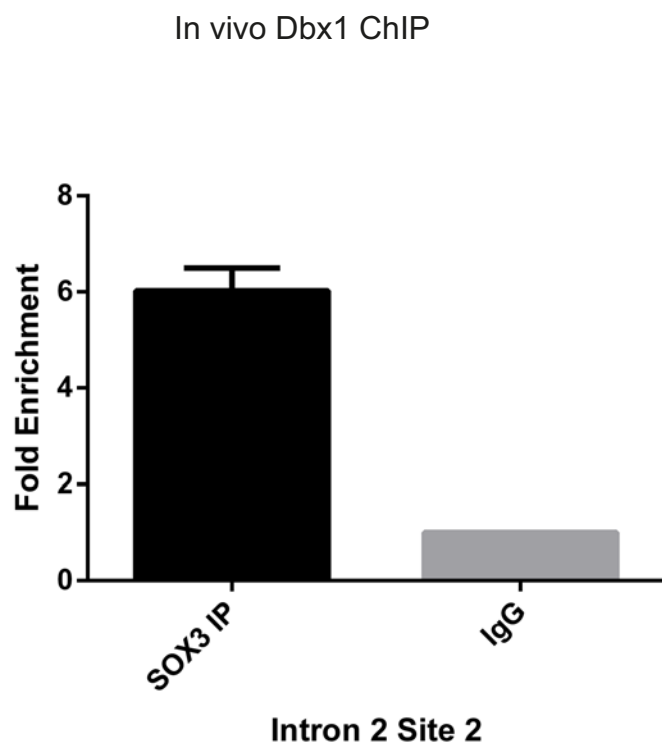
A)



B)

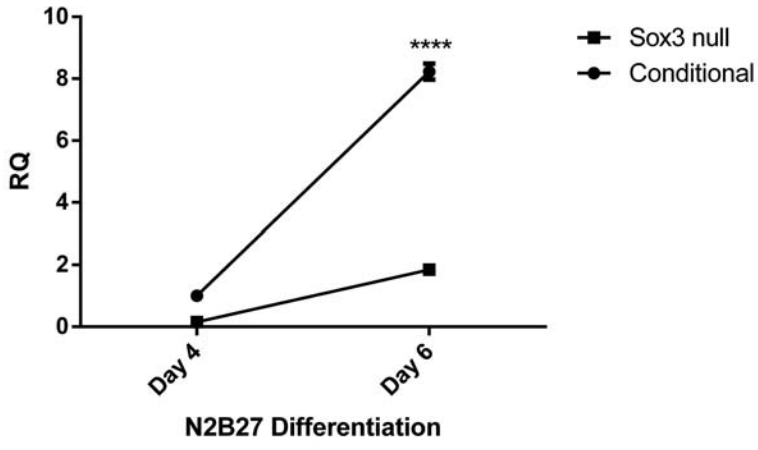


C)

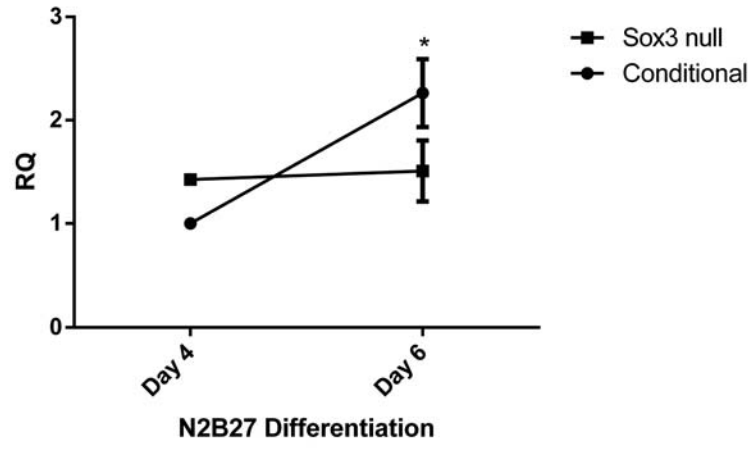


A)

Dbx1 expression

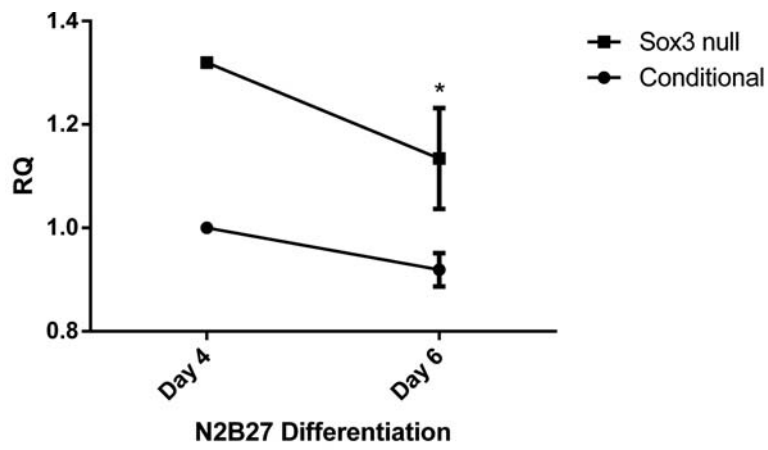


Slit1 expression

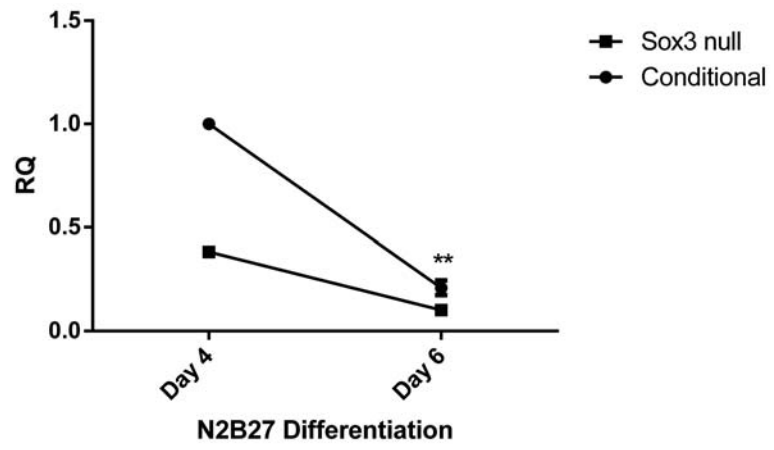


B)

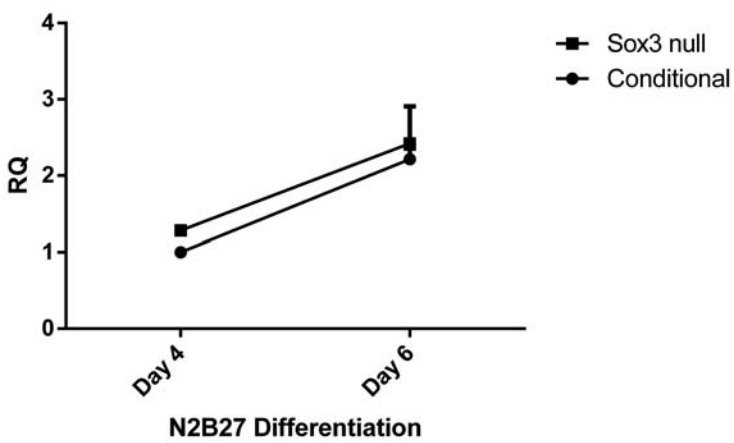
Cp expression



Fezf2 expression



Efnb3 expression



Discussion

Sox3 expression throughout dorsal telencephalic development

Sox3 is known to be expressed within nascent neuroepithelium of post-gastula stage vertebrate embryos (Collignon et al., 1996; Uchikawa et al., 2011; Wood and Episkopou, 1999), however little is known about its expression in the brain during embryonic or adult neurogenesis. Using a SOX3-specific antibody, I have shown that neural stem/progenitor cells located in the ventricular zone of the presumptive cortex exhibit robust levels of SOX3. SOX3 expression overlaps extensively with the neural progenitor cells (Hartfuss et al., 2001; Quinn et al., 2007). Conversely, SOX3 is not co-expressed with immature neurons indicating that it is rapidly downregulated during neuronal differentiation (Rogers et al 2013). Together, these data suggest a role for SOX3 in neural progenitor maintenance, consistent with published functional data in chick spinal cord (Bylund et al., 2003).

Deletion of *Sox3* in mice results in a comparatively mild CNS phenotype that does not include any overt defect in the cortex (Rizzoti et al., 2004). Although the developmental mechanism that underpins these defects is not known, it seems likely that the defect may arise from a population of SOX3+ neuroprogenitor cells that are susceptible to SoxB1 levels.

SOX3 expression within the adult neurogenic niches

Sox3 expression in adult neurogenic zones is robust with SOX3 expression detected in both the SGZ and SVZ. These data suggest that SOX3 is expressed within the slowly dividing neural stem (type1) cells (Seri et al., 2001) that are thought to be the most primitive neural stem/progenitor cells in the dentate gyrus. The expression of SOX3 within dividing neural progenitors, combined with the extensive overlap with SOX2 expression, indicates that SOX3+ cells are also found within the actively dividing (but not committed) NS/PC population (type 2a cells (Suh et al., 2007)). Interestingly, SOX3 expression is rapidly down regulated when the newly born NS/PC become committed to the neuronal fate (the type2a to type 2b transition). SOX3 expression within the SGZ is therefore reminiscent of its regulation during embryonic neurogenesis.

Robust SOX3 expression was also detected within the SVZ of the lateral ventricles. However, unlike the SGZ, SOX3+ cells in the SVZ constitute a small subpopulation of the SOX2+ cell pool. My data suggest that SOX3 is likely expressed within the “type B” cell population.

If so, this would establish SOX3 as one of the earliest and most restricted markers in this pathway. These data also suggest that adult neurogenesis may be defective in *Sox3* null mice. However, the extensive overlap with SOX2 and most likely SOX1 (Venere et al., 2012) would also suggest that any defect, if present, would be subtle within the developed brain.

Dynamic expression of SOX3 in the developing ventral diencephalon

Sox3 expression within the ventral diencephalon is of particular interest due to the published phenotypes within the *Sox3* null mice and humans with SOX3 mutations, which have implicated SOX3 in the generation of the hypothalamic-pituitary axis (Lagerstrom-Fermer et al., 1997; Rizzoti et al., 2004; Solomon et al., 2002; Wood and Episkopou, 1999; Woods et al., 2005). The hypothalamus originates from the third ventricular neuroepithelium, a region which I have shown has clear SOX3 expression within the periluminal progenitors. This data matches the telencephalic data showing SOX3 expression within the progenitor population. However, as hypothalamic neurogenesis proceeds, *Sox3* is robustly expressed in a subpopulation of β -III-tubulin⁺ immature neurons. Furthermore, magnocellular (Calbindin-D28k⁺) neurones also express SOX3. This is the first definitive example of a terminally differentiated neuron that expresses SOX3. Given that there is widespread *Sox3* expression in embryonic hypothalamic nuclei that continues through to at least P21, it seems likely that *Sox3* is expressed by a wide variety of hypothalamic neurons. These data provide a useful framework for deciphering the cellular basis of the hypothalamic-pituitary axis dysfunction of *Sox3* null mice.

Although hypothalamic function is compromised in *Sox3* null mice, the cellular mechanism of this defect is not fully understood. One possibility is that SOX3⁺ neurons in the mature hypothalamus that we have identified are functionally compromised. For example, mutation of *Sox3* in neurons of the arcuate nucleus may impact on GHRH-mediated release of GH from the anterior pituitary (Szarek et al., 2010). An alternative possibility is that *Sox3* loss-of-function compromises neurogenesis in the ventral diencephalon. In this respect, it is notable that there is

little overlap between SOX3 and SOX2 expression in the ventral diencephalon during the neurogenesis. As far as we are aware, this is the only known region of the CNS in which SOX3+ cells do not also express Sox2, thereby providing scope for phenotypic manifestation of *Sox3* LOF mutations. Notably *Sox2* mutations also lead to hypothalamic–pituitary axis dysfunction that appears to have a hypothalamic component (Jayakody et al., 2012).

SOX3 expression outside of the neurogenic niches within the adult brain

As published within Rogers et al 2013 (Results Chapter 1) SOX3 expression within the adult mouse brain is not only restricted to the adult neurogenic zones. Sporadic expression of SOX3 can be observed throughout numerous regions of the adult brain (Table 2 of Rogers et al 2013). This data is intriguing and suggests that SOX3 may be functionally active in post mitotic cells throughout the developed mouse brain.

Generation of Sox3 null embryonic stem cells

As discussed in Hughes et al 2013 (Result Chapter 2), during my PhD studies I targeted ES cells with the construct used by Rizzotti et al 2004 to generate a *Sox3* conditional ES cells. Using electroporation of a transient circular plasmid expressing Cre recombinase, *Sox3* null ES cells were generated. These ES cells were generated for two purposes. First, the cells could be used to study the impact of *Sox3* deletion on gene expression in neural progenitor cells. Second, the *Sox3* null ES cells could be used to study a previously published gastrulation phenotype (Introduction 1.3.5), whereby *Sox3* null chimeric embryos failed to progress through gastrulation (Acloque et al., 2011).

Sox3 null chimera generation and implications for Sox3 disease models

Sox3 null chimeras failure to progress through gastrulation, thereby providing a functional assay for the effects of the loss of *Sox3* in early development. Using two independently generated *Sox3* null ES cell clones the number of abnormal 7.5 dpc chimeric embryos, generated through morula injection, were scored. Both *Sox3* conditional and *Sox3*-26ala chimeric embryos had comparable levels of abnormal embryos. As expected, a significantly larger number of abnormal embryos were observed within *Sox3* null ES-derived chimeras. Comparison between chimeric embryos generated with conditional or polyalanine expansion ES cells determined that SOX3 protein with a +12 polyalanine expansion had residual SOX3 function. In addition, my independently derived *Sox3* null chimeras observed abnormal gastrula development as expected from previously published data (Acloque et al., 2011;

Rizzoti et al., 2004). As such, the *Sox3* null chimeria were used as an important functional assay to elucidate the effects of the polyalanine expansions on SOX3 function with potential implications for the mechanism underpinning polyalanine diseases (Results Chapter 2: Hughes et al 2013).

SOX3 expression in cultured neural progenitor differentiation

Recently a number of studies have established a role for SOX3 DNA binding within neural progenitor maintenance and differentiation. However genes that functionally require *Sox3* for their expression within neural progenitors remain unidentified. As such, it has proven difficult to identify the specific functions of *Sox3* within the developing CNS and its relationship with published phenotypes. Through the use of *Sox3* conditional/null ES cells and a well established NP differentiation assay (N2B27, Introduction 1.5.2) identification of putative *Sox3* targets is possible.

Before identification of putative *Sox3* target genes could be performed, characterisation of *Sox3*/SOX3 expression needed to be defined within the N2B27 differentiation assay. *Sox3*/SOX3, as expected, was robustly expressed within the NP cells in the N2B27 assay with *Sox3*/SOX3 expression initially observed on the second day of differentiation.

SOX3 protein expression observed at the fourth day of differentiation indicated that the majority of cells express SOX3 which was maintained within the NP population until at least the sixth day of differentiation.

Markers for neural progenitors *Sox1*, *Sox2* and *Pax6* expression levels were comparable, at day four of differentiation, between the conditional and *Sox3* null NP

cells. My data show that *Sox3*/SOX3 is expressed at the initiation of differentiation of ES cells into neural progenitors, however is not required for ES to NP differentiation. This data is not surprising due to the viability of *Sox3* null mice and potential functional redundancy between SoxB1 members, as both *Sox1* and *Sox2* are expressed within the N2B27 differentiation series.

SOX3 function within neural progenitors

A relatively small number of significantly differentially expressed genes appear to be responsive to the loss of *Sox3* within the neural progenitors (Rogers et al manuscript ,Table 1).The number of significant differentially regulated genes suggested that we may only be identifying a specific subset of *Sox3* targets, those most susceptible to SoxB1 levels. Importantly, both *Sox1* and *Sox2* are expressed at the fourth day of N2B27 differentiation. Therefore, putative targets identified by the microarray analysis are of particular interest as they are “less” susceptible to the effects of functional redundancy on the microarray results. Ergo, the targets identified are specifically susceptible to SOX3 levels in neural progenitor cells.

A majority of the putative targets validated on samples independent of the microarray suggest that the microarray was an accurate representation of the mRNA levels within the NP populations. As discussed previously in results chapter 2:(Rogers et al 2013 manuscript) most of the potential downstream targets of SOX3 are neural genes with close associations with central nervous system development and/or mental retardation. In addition, a number of the putative *Sox3* target genes have been demonstrated to be responsive to *Sox* genes (Oosterveen et al., 2013; Shim et al., 2012).

In vitro validation of putative *Sox3* targets indicated that most identified genes were sensitive to *Sox3* expression. However validation of only *Dbx1* could be observed within an *in vivo* context. Further studies across a number of different developmental time points may confirm a number of these genes to be targets (direct or indirect) of *Sox3*. Correlation between the microarray and published ChIP data suggests that almost all of the microarray identified target genes are direct and susceptible to *SoxB1* levels, subtle differences potentially leading to identification of populations critical to *Sox3* function in CNS development. However the data are confounded by the ever present prospect of functional redundancy within the developing mouse. Functional redundancy would limit the observed effects of the loss of *Sox3* on the neural progenitor population and within the developing mouse embryo.

Sox3 putative targets and potential function

Of all the putative *Sox3* targets identified within the NP microarray, the most significant (by P value, excluding *Sox3* itself) differentially expressed gene was chondroitin sulfate proteoglycan 5 (*Cspg5*). *Cspg5* was 1.49 fold higher in the *Sox3* null NP population. *Cspg5* has recently been shown to be involved in neuronal migration and is required for normal cognitive function in humans (Zhang et al., 2013a). *Cspg5* was found to be downstream of *PHF6*, mutations in which cause the intellectual disability disorder Borjeson-Forssman-Lehmann syndrome (Zhang et al., 2013a). Dysregulation of CSPG5 may therefore contribute to the intellectual disability that is present in some individuals with *SOX3* mutations (Helle et al., 2013; Laumonier et al., 2002; Stevanovic et al., 1993; Vencesla et al., 2007).

Another intriguing target identified by the microarray and validated *in vitro* was Discoidin domain receptor family member 2 (*Ddr2*). *Ddr2* knockout mice suffer from dwarfism, however *Ddr2* is upregulated within the *Sox3* null NP population (Labrador et al., 2001). Recent studies have shown that *Ddr2* can trigger signalling cascades that result in an up-regulation of *Snail1* (Zhang et al., 2013b). *Sox3* has previously been shown to directly repress *Snail1* with aberrant expression of *Snail1* resulting in embryonic lethality within *Sox3* null chimeras at gastrulation (Acloque et al., 2011). Although the *Snail1/Sox3* interaction has already been published to be direct, an increase in *Ddr2* as observed within the *Sox3* null NP population could result in aberrant upregulation of *Snail1* at later stages. This suggests a separate but consistent mode of regulation between *Sox3* and *Snail1* to that of the previously published role within gastrulation development (Acloque et al., 2011). The increased *Ddr2* expression within the *Sox3* null neural progenitor culture potentially identifies an independent pathway by which *Sox3* could functionally repress *Snail1* activity within neuro-ectodermal lineages.

As briefly discussed earlier (Results Chapter 2: Rogers et al, manuscript) another putative SOX3 target was the down-regulated transcription factor gene Fez family zinc finger 2 (*Fezf2*). *Fezf2* has been shown to be important for the early development of the posterior hypothalamus/thalamus in zebrafish and mouse (Hirata et al., 2006; Wolf and Ryu, 2013). In addition, *Fezf2* null mutant mice have abnormal formation of the corpus callosum and loss of the fimbria and fornix, hippocampal structures closely associated with the dorsal hippocampal commissure (Molyneaux et al., 2005; Shimizu and Hibi, 2009).

Independent studies have found that *Fezf2* expression is downstream of the SOXC group genes *Sox4* and *Sox11*, double deletion of which leads to a loss of *Fezf2*

expression, resulting in a Reeler like inversion of the neocortical layers (Shim et al., 2012). Regulation of *Fezf2* by SOX4/11 gene is mediated by a *cis* acting enhancer binding site which, interestingly, was also identified as a SOX3 binding site in NP cells by ChIP-seq (Bergsland et al., 2011; Shim et al., 2012). These data suggest that, in addition to acting as a pioneer factor (Bergsland et al., 2011), *Sox3* may directly regulate *Fezf2* expression in NP cells.

Another target identified by the microarray was developing brain homeobox 1 gene (*Dbx1*). *Dbx1* levels were shown to be sensitive to *Sox3* levels at the mRNA and protein level in vivo within the developing mouse CNS. *Dbx1* expressing cells have been previously implicated in craniofacial development, an area which is known to require *Sox3* function in *Sox3* null mice (Causeret et al., 2011; Rizzoti and Lovell-Badge, 2007). A recent study showed that ectopically expressing *Sox3* within the developing limb bud resulted in ectopic expression of *Dbx1*, *Dbx2* and *Pax6* (Oosterveen et al., 2013). These data together with our data strongly implicate *Dbx1* as a downstream target of endogenous *Sox3* function.

***Dbx1* as a downstream target of *Sox3* function in the developing spinal cord**

My results have indicated a potential role for *Sox3* in *Dbx1* regulation in murine CNS development. Endogenous *Dbx1* is required for V_0 interneuron specification within the developing spinal cord and it therefore becomes interesting to look at spinal cord development within *Sox3* null murine development (Pierani et al., 2001). An obvious reduction in *Dbx1* expression was observed within somite matched *Sox3* null 9.5 dpc whole embryos. In addition, a reduction in the number of cells expressing DBX1

within somite matched *Sox3* null 9.5pc spinal cords was also be observed. These data are indicative of a potentially unidentified function of *Sox3* within the developing murine spinal cord.

Previously published data have suggested that *Sox3* function within the developing spinal cord requires external morphogenic cofactors (Shh, RA and BMP) to determine its target gene specificity. *Shh*, RA and BMP form signalling gradients responsible for D-V patterning of the developing spinal cord. In addition, published data independently show that each of these major signalling gradients has an effect on *Dbx1* expression within the developing spinal cord (Gribble et al., 2009; Litingtung and Chiang, 2000; Novitch et al., 2003; Pierani et al., 1999; Timmer et al., 2002).

BMP has been shown to repress the *Dbx1*⁺ population, while retinoic acid and Wnt signalling has been shown to be required for maintenance of the *Dbx1* positive population (Gribble et al., 2009; Novitch et al., 2003; Pierani et al., 1999; Timmer et al., 2002). Recent data observing genetically conserved elements containing *Sox3* binding sites have clearly identified synergistic binding of SOX3 and Shh, RA and BMP within the developing spinal cord (Oosterveen et al., 2013).

These data, in combination with my recent studies, clearly denote regulation of *Dbx1* by *Sox3*, although partner factors to this function have yet to be specified.

Recovery of DBX1 expression within *Sox3* null embryos

qRT-PCR analysis of *Dbx1* expression within the developing mouse brain shows a reduction at 10.5 dpc although immunohistochemistry appear to demonstrated that the DBX1⁺ population within the hypothalamus appear comparable between wild

type and *Sox3* null mice. Expression of DBX1 at 9.5 dpc by IHC within the anterior CNS also appeared comparable between wild type and *Sox3* null tissue. In addition, a number of cells begin expressing DBX1 within the *Sox3* null spinal cord at 10.5 dpc. These data combined suggest that the lack *Sox3* expression delays the initiation of *Dbx1* expression. DBX1 expression within the developing brain appears to be less susceptible to the loss of *Sox3*, particularly within the hypothalamus.

Dbx1 misregulation within the *Sox3* null spinal cord may be indicative of a wider developmental delay phenotype within the *Sox3* null spinal cord. However such defects could be recovered at latter stages, providing an explanation for the lack of morphological defects observed within the *Sox3* null spinal cord. Alternatively the data may suggest a previously unidentified spinal cord phenotype within *Sox3* null mice.

My results suggest that the regulation of *Dbx1* by *Sox3* is a temporal event, although these small delays/reductions in *Dbx1* expression may have profound effects within the spinal cord and unstudied regions where *Dbx1* function may play a role.

DBX1 as a direct target of SOX3

SOX3 binds conserved elements near *Dbx1* (As briefly discussed in Results Chapter 2: Rogers et al, manuscript) suggesting that it directly regulates *Dbx1* in cultured neural progenitors. The most highly enriched and proximal site was validated *in vivo* using tissue from wild type 10.5 dpc heads. This data is the first *in vivo* validation of Bergsland et al 2011 published SOX3 element. Although these sites have not been directly linked to *Dbx1* regulation, it is the first evidence that SOX3 is binding in close proximity to the *Dbx1* promoter within anterior CNS tissue (Bergsland et al., 2011).

My data, in combination with recent published evidence within spinal cord strongly support that SOX3 directly regulates *Dbx1/DBX1* expression (Oosterveen et al., 2013). However, as SOX3 is bound to a number of different sites during the developmentally important time point, determining the exact functional effects of each site would be the focus of future studies (see below in future directions).

Future Directions

Sox3 expression and role in adult neurogenic zones

During my PhD studies I identified SOX3 expression within both neurogenic zones of the adult mouse brain. However, due to the limitations of marker studies, SOX3 expression within exact neural progenitor subpopulations was not precisely defined. My SOX3 expression study suggests that SOX3 is expressed within the stem cell population but may or may not be expressed within the quiescent population until they are actively dividing. Such distinctions could help identify the functional role SOX3 plays in each neurogenic zone.

To truly elucidate the exact expression profile of SOX3 in both neurogenic zones, long term EdU chase experiments could be conducted. By marking the dividing cells with EdU 4-6 weeks before the mice are sacrificed it would be possible to co-localise SOX3 expression with the quiescent neural progenitor population located in the SGZ and SVZ.

In addition, it would be interesting to generate and utilise an eGFP reporter mouse line for *Sox3* positive cells. Through dissection and disassociation of both of the adult neurogenic niches and FAC sorting of the GFP+ cells it would be possible to create SOX3+ enriched NP population from which to create neurosphere cultures. These neurosphere cultures could be used to assay the long term progenitor potential and

differentiation qualities of the *Sox3* NP cells. These experiments would help identify the role and function of *Sox3* within the adult neurogenic niches. The experiments could be conducted concurrently with *Sox3* null mice to help identify functional consequences of the loss of *Sox3* on adult neurogenesis.

SOX3 expression in non neurogenic adult brain zones

Data produced during my PhD also identified SOX3+ cells within the adult mouse brain that are not located within the neurogenic niches. The expression of SOX3 in terminally differentiated zones of the adult mouse brain suggests additional roles for SOX3 function outside of its role in neural progenitors. However the exact identities of these SOX3+ cells remains unknown.

Marker studies co-localising neuronal (NeuN), oligodendrocytes (O4) or glial (GFAP) markers could identify the general characteristics of SOX3+ cells. Further studies of the discrete sub populations within each distinct zone of SOX3+ expression, similar to those studied in the developed hypothalamus (Rogers et al, 2013) could then be performed. These data will help with the identification of the SOX3+ cells within the non-neurogenic zones of the adult mouse brain and potentially identify additional roles for *Sox3* function.

Comprehensive SOXB1 expression profiling

Another future experiment that would provide essential insight of the role of *SoxB1* in the developing CNS is a comprehensive overview of all SOXB1 members expression

patterns throughout development. Similar to marker studies conducted within Rogers et al, 2013, immunohistochemistry using antibodies for SOX1,2 and 3 could be used across a large span of development time points(8-18 dpc). Co-localisation of the SOXB1 expression profiles could help to identify populations of cells that express individual SOXB1 members, cells that may not be subject to functional redundancy. Ideally whole mount immunohistochemistry could be used to broadly view expression patterns, thus systematically observing the intact CNS. Sectioning stained whole mount embryos could then be performed to identify new zones of interest at each time point. This type of study would help with the identification of phenotypes associated with each of the individual SOXB1 null mouse lines.

Functional redundancy between SoxB1 genes

SoxB1 functional redundancy is a common idea promoted by researchers within the *SoxB1* field. A study on zebrafish embryo development, looking at compound quadruple *SoxB1* knockout embryos found that individual knock-downs of *Sox2/3* resulted in no gross abnormalities within the developing embryos, however compound quadruple null embryos had severe phenotypes, thus providing evidence for functional redundancy (Okuda et al., 2010). Such experiments have not been conducted within the mammalian systems due to the difficulties associated with generation of null tissues.

However, the current explosion of new technology for rapid generation of null mouse lines through the use of CRISPRs, provides the capability to study compound *SoxB1* null mice to truly elucidate the affects of functional redundancy between *SoxB1* genes.

Sox1 and *Sox3* null mice have been previously published to have relatively minor phenotypes however *Sox2* null mice are embryonic lethal as *Sox2* is required for normal ES cell maintenance (Ferri et al., 2004; Malas et al., 2003; Rizzoti et al., 2004). CRISPRs targeting *Sox1,2* and *3* open reading frames could be used to create compound deletions (or heterozygous in the case of *Sox2*) and combinations thereof within the developing murine embryo. Studying the phenotypes of these embryos it becomes possible to observe the effects of *SoxB1* functional redundancy in mice. In addition, gross morphological studies of these embryos may identify new roles for *SoxB1* function in embryonic development.

Sox3 function and D-V patterning within the developing spinal cord

As previously discussed, *Sox3* function within the developing spinal cord has been associated with the major morphogenic signals. It is, therefore possible that a previously unidentified phenotype may exist within the *Sox3* null murine spinal cord and experiments are required to elucidate the exact mechanism of action of *Sox3*.

Studies examining at all the potential sub populations of interneurons/motorneuron populations would give a clear indication of any fate diversion within the *Sox3* null spinal cord (eg expansion of dorsal populations at the expense of the V_0 interneurons).

cDNA could be generated from RNA dissected out of wild type and *Sox3* null somite matched spinal cords and qRT-PCR analysis of markers for each interneuron/motorneuron sub population could be performed. If significant differences are observed at the mRNA level within a given sub-population, marker

studies with previously published antibodies could be used to gain spatial/temporal conformation of changes. If evidence of a potential spinal cord phenotype exists within the *Sox3* null mice, behavioural studies looking specifically at motor neurons could be conducted. This data will characterise the developing spinal cord as a region that is potential susceptible to the loss of *Sox3*.

SOX3 Binding and function within spinal cord development

Recent studies have identified a large number of SOX3 binding sites within cultured neural progenitors however identifying the functional consequences of individual binding sites is of great importance. Data produced during my PhD has suggested a functional role for *Sox3* regulation of *Dbx1* within the developing spinal cord. As such, prioritisation of SOX3 binding sites targeting developmentally important spinal cord genes could help identify critical sites.

Studying SOX3 binding sites identified in close proximity to spinal cord class 1 and class 2 transcription factors, like *Dbx1* may prove informative (Bergsland et al., 2011; Oosterveen et al., 2013). Ideally, performing a SOX3 ChIP on dissected spinal cord tissue, using the *Dbx1* sites validated in Results Chapter 2 as positive controls might provide a set of *in vivo* validated SOX3 binding sites. In addition, ChIP-seq could be performed to identify all of the SOX3 binding sites within the developing spinal cord and be compared to sites identified within cultured neural progenitor cells. Initially these targets could be cross referenced against previously published ChIP data for spinal cord transcription factors (*Wnt/Tcf3*, *BMP/smad1/5* and *Shh/Gli3*), looking for both overlapping binding sites and enrichment of paired consensus binding sites.

Through using only spinal cord tissue in these ChIP experiments a possible reduction in overall number of SOX3 binding sites may be observed while increasing the frequency of spinal cord specific binding partners. Differences observed between SOX3 ChIP data sets could help identified spatial functions for *Sox3*.

This data will allow for a greater understanding of SOX3 binding sites and, ultimately helping to identify SOX3 function within the developing spinal cord. In addition, these experiments could be conducted with all SoxB1 members in a hope to identify subtle differences in binding and function between members.

SOX3 binding sites and functional effects

CRISPR technologies have also opened the door to a number of previously unfeasible experiments looking directly at functional impacts of individual SOX3 binding sites. These experiments would be faster than traditional targeting technology and have direct *in vivo* implications for both *Sox3* function in CNS development and the analysis of all future ChIP data sets. Through systematic deletion of the 5 sites surrounding a *Dbx1* and a whole mount *in situ* assay for *Dbx1* expression at 9.5 dpc, we can analyse individual effects of SOX3 binding sites on *Dbx1* expression.

Through *in vivo* analysis of both our validated sites and other independent published sites we could identify functionally important SOX3 binding sites for *Dbx1* expression. This data will help solidify SOX3 binding as an important functional output for spinal cord development and identify the number of SOX3 sites required for *Dbx1* expression within the early embryo. In addition, whole mount *in situ*

experiments will help identify spatial differences between each binding site, e.g those required for anterior expression.

Such experiments could be conducted concurrently with deletion of published surrounding SOX2 binding sites (Bergsland et al., 2011). However, of the five *Dbx1* SOX3 binding sites, SOX2 binding overlaps at three, studying the deletion effects at the SOX3 only binding sites may give novel insight into *Dbx1* regulation and functional redundancy of SoxB1 members.

Dbx1 and hypothalamic development

While conducting my studies of DBX1 expression throughout the developing murine embryo it became apparent that DBX1 is expressed within the developing hypothalamus (posterior hypothalamus). *Dbx1*/DBX1 expression has not been fully characterised within the developing brain and the identity of hypothalamic *Dbx1*⁺ population of cells is unknown. A published paper using a LacZ reporter mouse that marks *Dbx1*⁺ derived cells from 9.5-12.5 dpc has shown that the *Dbx1*⁺ derived cells form part of the mature amygdala and hypothalamus (Hirata et al., 2009). The hypothalamic derived cells, although not the focus of the paper, are of potential interest due to phenotypes identified within *Sox3* null patients and mice. Cre recombinase fused to the tamoxifen sensitive estrogen receptor (ERT2) is expressed within *Dbx1*⁺ cells and is translocated to the nucleus on tamoxifen induction (*Dbx1*^{+/creERT2}). R26R lacZ mice express LacZ on Cre recombination, as such crossing these two mouse lines together will result in reporting *Dbx1*⁺ cells on tamoxifen induction (*Dbx1*^{+/creERT2};R26RLacZ). It would be interesting to cross the *Dbx1*⁺ inducible reporter mice with the *Sox3* null mice (*Dbx1*^{+/creERT2};R26RLacZ;*Sox3*

null^{-Y}). This cross would allow the observation of the Dbx1 derived hypothalamic population (particularly those derived from the preoptic area) within the *Sox3* null mice, the identification of the cell types and if they are subtly affected by the loss of *Sox3* function.

This data will link the hypothalamic population to *Sox3* expression and help explain the reduction of *Dbx1* within the developing brain. In addition, the data will provide potential evidence that *Dbx1* plays an important role in hypothalamic development, a previously unstudied area.

Concluding remarks

Data produced during my PhD have identified the expression pattern of SOX3 within the developing mouse brain and adult neurogenic zones. In addition, I have identified putative *Sox3* targets within neural progenitor cells and characterised *Dbx1/DBX1* as a target of *Sox3* within the developing spinal cord. Together these data have increased the understanding of the role of *Sox3* within the developing CNS and provide a new framework for future experiments to elucidate the mechanisms underpinning the role of *Sox3* in development.

Bibliography

- Acloque, H., Ocana, O. H., Matheu, A., Rizzoti, K., Wise, C., Lovell-Badge, R. and Nieto, M. A.** (2011). Reciprocal repression between Sox3 and snail transcription factors defines embryonic territories at gastrulation. *Dev Cell* **21**, 546-58.
- Ambrosetti, D. C., Basilico, C. and Dailey, L.** (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol Cell Biol* **17**, 6321-9.
- Anthony, T. E., Klein, C., Fishell, G. and Heintz, N.** (2004). Radial glia serve as neuronal progenitors in all regions of the central nervous system. *Neuron* **41**, 881-90.
- Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. and Lovell-Badge, R.** (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **17**, 126-40.
- Bergsland, M., Ramskold, D., Zaouter, C., Klum, S., Sandberg, R. and Muhr, J.** (2011). Sequentially acting Sox transcription factors in neural lineage development. *Genes Dev* **25**, 2453-64.
- Bonaguidi, M. A., Peng, C. Y., McGuire, T., Falciglia, G., Gobeske, K. T., Czeisler, C. and Kessler, J. A.** (2008). Noggin expands neural stem cells in the adult hippocampus. *J Neurosci* **28**, 9194-204.
- Botquin, V., Hess, H., Fuhrmann, G., Anastassiadis, C., Gross, M. K., Vriend, G. and Scholer, H. R.** (1998). New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by Oct-4 and Sox-2. *Genes Dev* **12**, 2073-90.
- Bowles, J., Schepers, G. and Koopman, P.** (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol* **227**, 239-55.
- Boyer, L. A., Lee, T. I., Cole, M. F., Johnstone, S. E., Levine, S. S., Zucker, J. P., Guenther, M. G., Kumar, R. M., Murray, H. L., Jenner, R. G. et al.** (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **122**, 947-56.
- Brunelli, S., Silva Casey, E., Bell, D., Harland, R. and Lovell-Badge, R.** (2003). Expression of Sox3 throughout the developing central nervous system is dependent on the combined action of discrete, evolutionarily conserved regulatory elements. *Genesis* **36**, 12-24.
- Bulfone, A., Martinez, S., Marigo, V., Campanella, M., Basile, A., Quaderi, N., Gattuso, C., Rubenstein, J. L. and Ballabio, A.** (1999). Expression pattern of the Tbr2 (Eomesodermin) gene during mouse and chick brain development. *Mech Dev* **84**, 133-8.
- Bylund, M., Andersson, E., Novitch, B. G. and Muhr, J.** (2003). Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nat Neurosci* **6**, 1162-8.
- Causeret, F., Ensini, M., Teissier, A., Kessar, N., Richardson, W. D., Lucas de Couville, T. and Pierani, A.** (2011). Dbx1-expressing cells are necessary for the survival of the mammalian anterior neural and craniofacial structures. *PLoS One* **6**, e19367.
- Chanas-Sacre, G., Rogister, B., Moonen, G. and Leprince, P.** (2000). Radial glia phenotype: origin, regulation, and transdifferentiation. *J Neurosci Res* **61**, 357-63.
- Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N. and Lovell-Badge, R.** (1996). A

comparison of the properties of Sox-3 with Sry and two related genes, Sox-1 and Sox-2. *Development* **122**, 509-20.

Craig, C. G., Tropepe, V., Morshead, C. M., Reynolds, B. A., Weiss, S. and van der Kooy, D. (1996). In vivo growth factor expansion of endogenous subependymal neural precursor cell populations in the adult mouse brain. *J Neurosci* **16**, 2649-58.

Dale, J. K., Vesque, C., Lints, T. J., Sampath, T. K., Furley, A., Dodd, J. and Placzek, M. (1997). Cooperation of BMP7 and SHH in the induction of forebrain ventral midline cells by prechordal mesoderm. *Cell* **90**, 257-69.

Dessaud, E., Yang, L. L., Hill, K., Cox, B., Ulloa, F., Ribeiro, A., Mynett, A., Novitsch, B. G. and Briscoe, J. (2007). Interpretation of the sonic hedgehog morphogen gradient by a temporal adaptation mechanism. *Nature* **450**, 717-20.

Esposito, M. S., Piatti, V. C., Laplagne, D. A., Morgenstern, N. A., Ferrari, C. C., Pitossi, F. J. and Schinder, A. F. (2005). Neuronal differentiation in the adult hippocampus recapitulates embryonic development. *J Neurosci* **25**, 10074-86.

Farmer, J., Zhao, X., van Praag, H., Wodtke, K., Gage, F. H. and Christie, B. R. (2004). Effects of voluntary exercise on synaptic plasticity and gene expression in the dentate gyrus of adult male Sprague-Dawley rats in vivo. *Neuroscience* **124**, 71-9.

Favaro, R., Valotta, M., Ferri, A. L., Latorre, E., Mariani, J., Giachino, C., Lancini, C., Tosetti, V., Ottolenghi, S., Taylor, V. et al. (2009). Hippocampal development and neural stem cell maintenance require Sox2-dependent regulation of Shh. *Nat Neurosci* **12**, 1248-56.

Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R. and Bianchi, M. E. (1992). SRY, like HMG1, recognizes sharp angles in DNA. *Embo J* **11**, 4497-506.

Ferri, A. L., Cavallaro, M., Braidà, D., Di Cristofano, A., Canta, A., Vezzani, A., Ottolenghi, S., Pandolfi, P. P., Sala, M., DeBiasi, S. et al. (2004). Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. *Development* **131**, 3805-19.

Fukuda, S., Kato, F., Tozuka, Y., Yamaguchi, M., Miyamoto, Y. and Hisatsune, T. (2003). Two distinct subpopulations of nestin-positive cells in adult mouse dentate gyrus. *J Neurosci* **23**, 9357-66.

Garcia, A. D., Doan, N. B., Imura, T., Bush, T. G. and Sofroniew, M. V. (2004). GFAP-expressing progenitors are the principal source of constitutive neurogenesis in adult mouse forebrain. *Nat Neurosci* **7**, 1233-41.

Gotz, M. and Huttner, W. B. (2005). The cell biology of neurogenesis. *Nat Rev Mol Cell Biol* **6**, 777-88.

Gribble, S. L., Kim, H. S., Bonner, J., Wang, X. and Dorsky, R. I. (2009). Tcf3 inhibits spinal cord neurogenesis by regulating sox4a expression. *Development* **136**, 781-9.

Guillemot, F. (2005). Cellular and molecular control of neurogenesis in the mammalian telencephalon. *Curr Opin Cell Biol* **17**, 639-47.

Hartfuss, E., Galli, R., Heins, N. and Gotz, M. (2001). Characterization of CNS precursor subtypes and radial glia. *Dev Biol* **229**, 15-30.

Haubensak, W., Attardo, A., Denk, W. and Huttner, W. B. (2004). Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: a major site of neurogenesis. *Proc Natl Acad Sci U S A* **101**, 3196-201.

Helle, J. R., Baroy, T., Misceo, D., Braaten, O., Fannemel, M. and Frengen, E. (2013). Hyperphagia, mild developmental delay but apparently no structural brain anomalies in a boy without SOX3 expression. *Am J Med Genet A* **161A**, 1137-42.

Hirata, T., Li, P., Lanuza, G. M., Cocas, L. A., Huntsman, M. M. and Corbin, J. G. (2009). Identification of distinct telencephalic progenitor pools for neuronal diversity in the amygdala. *Nat Neurosci* **12**, 141-9.

Hirata, T., Nakazawa, M., Muraoka, O., Nakayama, R., Suda, Y. and Hibi, M. (2006). Zinc-finger genes Fez and Fez-like function in the establishment of diencephalon subdivisions. *Development* **133**, 3993-4004.

Hodge, R. D., Kowalczyk, T. D., Wolf, S. A., Encinas, J. M., Rippey, C., Enikolopov, G., Kempermann, G. and Hevner, R. F. (2008). Intermediate progenitors in adult hippocampal neurogenesis: Tbr2 expression and coordinate regulation of neuronal output. *J Neurosci* **28**, 3707-17.

Holmberg, J., Hansson, E., Malewicz, M., Sandberg, M., Perlmann, T., Lendahl, U. and Muhr, J. (2008). SoxB1 transcription factors and Notch signaling use distinct mechanisms to regulate proneural gene function and neural progenitor differentiation. *Development* **135**, 1843-51.

Hughes, J., Piltz, S., Rogers, N., McAninch, D., Rowley, L. and Thomas, P. (2013). Mechanistic insight into the pathology of polyalanine expansion disorders revealed by a mouse model for X linked hypopituitarism. *PLoS Genet* **9**, e1003290.

Hughes, J. and Thomas, P. (2013). Molecular pathology of polyalanine expansion disorders: new perspectives from mouse models. *Methods in Molecular Biology* **1017**, 135-151.

Huttner, W. B. and Brand, M. (1997). Asymmetric division and polarity of neuroepithelial cells. *Curr Opin Neurobiol* **7**, 29-39.

Jayakody, S. A., Andoniadou, C. L., Gaston-Massuet, C., Signore, M., Cariboni, A., Bouloux, P. M., Le Tissier, P., Pevny, L. H., Dattani, M. T. and Martinez-Barbera, J. P. (2012). SOX2 regulates the hypothalamic-pituitary axis at multiple levels. *J Clin Invest* **122**, 3635-46.

Kamachi, Y., Uchikawa, M. and Kondoh, H. (2000). Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet* **16**, 182-7.

Kelberman, D. and Dattani, M. T. (2008). Septo-optic dysplasia - novel insights into the aetiology. *Horm Res* **69**, 257-65.

Kempermann, G. and Gage, F. H. (2002). Genetic determinants of adult hippocampal neurogenesis correlate with acquisition, but not probe trial performance, in the water maze task. *Eur J Neurosci* **16**, 129-36.

Kondoh, H. and Kamachi, Y. (2010). SOX-partner code for cell specification: Regulatory target selection and underlying molecular mechanisms. *Int J Biochem Cell Biol* **42**, 391-9.

Kondoh, H., Uchikawa, M. and Kamachi, Y. (2004). Interplay of Pax6 and SOX2 in lens development as a paradigm of genetic switch mechanisms for cell differentiation. *Int J Dev Biol* **48**, 819-27.

Kriegstein, A. R. and Gotz, M. (2003). Radial glia diversity: a matter of cell fate. *Glia* **43**, 37-43.

Labrador, J. P., Azcoitia, V., Tuckermann, J., Lin, C., Olaso, E., Manes, S., Bruckner, K., Goergen, J. L., Lemke, G., Yancopoulos, G. et al. (2001). The collagen receptor DDR2 regulates proliferation and its elimination leads to dwarfism. *EMBO Rep* **2**, 446-52.

Lagerstrom-Fermer, M., Sundvall, M., Johnsen, E., Warne, G. L., Forrest, S. M., Zajac, J. D., Rickards, A., Ravine, D., Landegren, U. and Pettersson, U. (1997). X-linked recessive panhypopituitarism associated with a regional duplication in Xq25-q26. *Am J Hum Genet* **60**, 910-6.

Laumonier, F., Ronce, N., Hamel, B. C., Thomas, P., Lespinasse, J., Raynaud, M., Paringaux, C., Van Bokhoven, H., Kalscheuer, V., Fryns, J. P. et al. (2002). Transcription factor SOX3 is involved in X-linked mental retardation with growth hormone deficiency. *Am J Hum Genet* **71**, 1450-5.

Litingtung, Y. and Chiang, C. (2000). Specification of ventral neuron types is mediated by an antagonistic interaction between Shh and Gli3. *Nat Neurosci* **3**, 979-85.

Lois, C. and Alvarez-Buylla, A. (1994). Long-distance neuronal migration in the adult mammalian brain. *Science* **264**, 1145-8.

Malas, S., Postlethwaite, M., Ekonomou, A., Whalley, B., Nishiguchi, S., Wood, H., Meldrum, B., Constanti, A. and Episkopou, V. (2003). Sox1-deficient mice suffer from epilepsy associated with abnormal ventral forebrain development and olfactory cortex hyperexcitability. *Neuroscience* **119**, 421-32.

Malatesta, P., Hack, M. A., Hartfuss, E., Kettenmann, H., Klinkert, W., Kirchhoff, F. and Gotz, M. (2003). Neuronal or glial progeny: regional differences in radial glia fate. *Neuron* **37**, 751-64.

Molyneaux, B. J., Arlotta, P., Hirata, T., Hibi, M. and Macklis, J. D. (2005). Fezl is required for the birth and specification of corticospinal motor neurons. *Neuron* **47**, 817-31.

Monuki, E. S. and Walsh, C. A. (2001). Mechanisms of cerebral cortical patterning in mice and humans. *Nat Neurosci* **4 Suppl**, 1199-206.

Morshead, C. M., Reynolds, B. A., Craig, C. G., McBurney, M. W., Staines, W. A., Morassutti, D., Weiss, S. and van der Kooy, D. (1994). Neural stem cells in the adult mammalian forebrain: a relatively quiescent subpopulation of subependymal cells. *Neuron* **13**, 1071-82.

Novitsch, B. G., Wichterle, H., Jessell, T. M. and Sockanathan, S. (2003). A requirement for retinoic acid-mediated transcriptional activation in ventral neural patterning and motor neuron specification. *Neuron* **40**, 81-95.

Okuda, Y., Ogura, E., Kondoh, H. and Kamachi, Y. (2010). B1 SOX coordinate cell specification with patterning and morphogenesis in the early zebrafish embryo. *PLoS Genet* **6**, e1000936.

Oosterveen, T., Kurdija, S., Enstero, M., Uhde, C. W., Bergsland, M., Sandberg, M., Sandberg, R., Muhr, J. and Ericson, J. (2013). SoxB1-driven transcriptional network underlies neural-specific interpretation of morphogen signals. *Proc Natl Acad Sci U S A* **110**, 7330-5.

Petreatu, L. and Alvarez-Buylla, A. (2002). Maturation and death of adult-born olfactory bulb granule neurons: role of olfaction. *J Neurosci* **22**, 6106-13.

Pevny, L. H. and Lovell-Badge, R. (1997). Sox genes find their feet. *Curr Opin Genet Dev* **7**, 338-44.

Pevny, L. H., Sockanathan, S., Placzek, M. and Lovell-Badge, R. (1998). A role for SOX1 in neural determination. *Development* **125**, 1967-78.

Pierani, A., Brenner-Morton, S., Chiang, C. and Jessell, T. M. (1999). A sonic hedgehog-independent, retinoid-activated pathway of neurogenesis in the ventral spinal cord. *Cell* **97**, 903-15.

Pierani, A., Moran-Rivard, L., Sunshine, M. J., Littman, D. R., Goulding, M. and Jessell, T. M. (2001). Control of interneuron fate in the developing spinal cord by the progenitor homeodomain protein Dbx1. *Neuron* **29**, 367-84.

Quinn, J. C., Molinek, M., Martynoga, B. S., Zaki, P. A., Faedo, A., Bulfone, A., Hevner, R. F., West, J. D. and Price, D. J. (2007). Pax6 controls cerebral cortical

cell number by regulating exit from the cell cycle and specifies cortical cell identity by a cell autonomous mechanism. *Dev Biol* **302**, 50-65.

Remenyi, A., Lins, K., Nissen, L. J., Reinbold, R., Scholer, H. R. and Wilmanns, M. (2003). Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev* **17**, 2048-59.

Reynolds, B. A., Tetzlaff, W. and Weiss, S. (1992). A multipotent EGF-responsive striatal embryonic progenitor cell produces neurons and astrocytes. *J Neurosci* **12**, 4565-74.

Reynolds, B. A. and Weiss, S. (1992). Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science* **255**, 1707-10.

Rietze, R. L. and Reynolds, B. A. (2006). Neural stem cell isolation and characterization. *Methods Enzymol* **419**, 3-23.

Rizzoti, K., Brunelli, S., Carmignac, D., Thomas, P. Q., Robinson, I. C. and Lovell-Badge, R. (2004). SOX3 is required during the formation of the hypothalamo-pituitary axis. *Nat Genet* **36**, 247-55.

Rizzoti, K. and Lovell-Badge, R. (2007). SOX3 activity during pharyngeal segmentation is required for craniofacial morphogenesis. *Development* **134**, 3437-48.

Rogers, N., Cheah, P. S., Szarek, E., Banerjee, K., Schwartz, J. and Thomas, P. (2013). Expression of the murine transcription factor SOX3 during embryonic and adult neurogenesis. *Gene Expr Patterns* **13**, 240-8.

Seri, B., Garcia-Verdugo, J. M., McEwen, B. S. and Alvarez-Buylla, A. (2001). Astrocytes give rise to new neurons in the adult mammalian hippocampus. *J Neurosci* **21**, 7153-60.

Shim, S., Kwan, K. Y., Li, M., Lefebvre, V. and Sestan, N. (2012). Cis-regulatory control of corticospinal system development and evolution. *Nature* **486**, 74-9.

Shimizu, T. and Hibi, M. (2009). Formation and patterning of the forebrain and olfactory system by zinc-finger genes *Fezf1* and *Fezf2*. *Dev Growth Differ* **51**, 221-31.

Solomon, N. M., Nouri, S., Warne, G. L., Lagerstrom-Fermer, M., Forrest, S. M. and Thomas, P. Q. (2002). Increased gene dosage at Xq26-q27 is associated with X-linked hypopituitarism. *Genomics* **79**, 553-9.

Solomon, N. M., Ross, S. A., Morgan, T., Belsky, J. L., Hol, F. A., Karnes, P. S., Hopwood, N. J., Myers, S. E., Tan, A. S., Warne, G. L. et al. (2004). Array comparative genomic hybridisation analysis of boys with X linked hypopituitarism identifies a 3.9 Mb duplicated critical region at Xq27 containing SOX3. *J Med Genet* **41**, 669-78.

Stevanovic, M., Lovell-Badge, R., Collignon, J. and Goodfellow, P. N. (1993). SOX3 is an X-linked gene related to SRY. *Hum Mol Genet* **2**, 2013-8.

Suh, H., Consiglio, A., Ray, J., Sawai, T., D'Amour, K. A. and Gage, F. H. (2007). In vivo fate analysis reveals the multipotent and self-renewal capacities of Sox2+ neural stem cells in the adult hippocampus. *Cell Stem Cell* **1**, 515-28.

Szarek, E., Cheah, P. S., Schwartz, J. and Thomas, P. (2010). Molecular genetics of the developing neuroendocrine hypothalamus. *Mol Cell Endocrinol* **323**, 115-23.

Timmer, J. R., Wang, C. and Niswander, L. (2002). BMP signaling patterns the dorsal and intermediate neural tube via regulation of homeobox and helix-loop-helix transcription factors. *Development* **129**, 2459-72.

Uchikawa, M., Kamachi, Y. and Kondoh, H. (1999). Two distinct subgroups of Group B Sox genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken. *Mech Dev* **84**, 103-20.

Uchikawa, M., Yoshida, M., Iwafuchi-Doi, M., Matsuda, K., Ishida, Y., Takemoto, T. and Kondoh, H. (2011). B1 and B2 Sox gene expression during neural plate development in chicken and mouse embryos: universal versus species-dependent features. *Dev Growth Differ* **53**, 761-71.

van Praag, H., Christie, B. R., Sejnowski, T. J. and Gage, F. H. (1999). Running enhances neurogenesis, learning, and long-term potentiation in mice. *Proc Natl Acad Sci U S A* **96**, 13427-31.

Vencesla, A., Barcelo, M. J., Baena, M., Quintana, M., Baiget, M. and Tizzano, E. F. (2007). Marker and real-time quantitative analyses to confirm hemophilia B carrier diagnosis of a complete deletion of the F9 gene. *Haematologica* **92**, 1583-4.

Venere, M., Han, Y. G., Bell, R., Song, J. S., Alvarez-Buylla, A. and Blulloch, R. (2012). Sox1 marks an activated neural stem/progenitor cell in the hippocampus. *Development* **139**, 3938-49.

Walker, T. L., White, A., Black, D. M., Wallace, R. H., Sah, P. and Bartlett, P. F. (2008). Latent stem and progenitor cells in the hippocampus are activated by neural excitation. *J Neurosci* **28**, 5240-7.

Wang, T. W., Stromberg, G. P., Whitney, J. T., Brower, N. W., Klymkowsky, M. W. and Parent, J. M. (2006). Sox3 expression identifies neural progenitors in persistent neonatal and adult mouse forebrain germinative zones. *J Comp Neurol* **497**, 88-100.

Weiss, S., Reynolds, B. A., Vescovi, A. L., Morshead, C., Craig, C. G. and van der Kooy, D. (1996). Is there a neural stem cell in the mammalian forebrain? *Trends Neurosci* **19**, 387-93.

Wilson, S. W. and Rubenstein, J. L. (2000). Induction and dorsoventral patterning of the telencephalon. *Neuron* **28**, 641-51.

Wolf, A. and Ryu, S. (2013). Specification of posterior hypothalamic neurons requires coordinated activities of Fezf2, Otp, Sim1a and Foxb1.2. *Development* **140**, 1762-73.

Wood, H. B. and Episkopou, V. (1999). Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev* **86**, 197-201.

Woods, K. S., Cundall, M., Turton, J., Rizotti, K., Mehta, A., Palmer, R., Wong, J., Chong, W. K., Al-Zyoud, M., El-Ali, M. et al. (2005). Over- and underdosage of SOX3 is associated with infundibular hypoplasia and hypopituitarism. *Am J Hum Genet* **76**, 833-49.

Wright, E. M., Snopce, B. and Koopman, P. (1993). Seven new members of the Sox gene family expressed during mouse development. *Nucleic Acids Res* **21**, 744.

Ying, Q. L., Stavridis, M., Griffiths, D., Li, M. and Smith, A. (2003). Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. *Nat Biotechnol* **21**, 183-6.

Zaki, P. A., Quinn, J. C. and Price, D. J. (2003). Mouse models of telencephalic development. *Curr Opin Genet Dev* **13**, 423-37.

Zhang, C., Mejia, L. A., Huang, J., Valnegri, P., Bennett, E. J., Anckar, J., Jahani-Asl, A., Gallardo, G., Ikeuchi, Y., Yamada, T. et al. (2013a). The X-linked intellectual disability protein PHF6 associates with the PAF1 complex and regulates neuronal migration in the mammalian brain. *Neuron* **78**, 986-93.

Zhang, K., Corsa, C. A., Ponik, S. M., Prior, J. L., Piwnica-Worms, D., Eliceiri, K. W., Keely, P. J. and Longmore, G. D. (2013b). The collagen receptor discoidin domain receptor 2 stabilizes SNAIL1 to facilitate breast cancer metastasis. *Nat Cell Biol* **15**, 677-87.

- Zhao, C., Deng, W. and Gage, F. H.** (2008). Mechanisms and functional implications of adult neurogenesis. *Cell* **132**, 645-60.
- Zhao, C., Teng, E. M., Summers, R. G., Jr., Ming, G. L. and Gage, F. H.** (2006). Distinct morphological stages of dentate granule neuron maturation in the adult mouse hippocampus. *J Neurosci* **26**, 3-11.
- Zhao, L., Zevallos, S. E., Rizzoti, K., Jeong, Y., Lovell-Badge, R. and Epstein, D. J.** (2012). Disruption of SoxB1-dependent Sonic hedgehog expression in the hypothalamus causes septo-optic dysplasia. *Dev Cell* **22**, 585-96.
- Zimmer, C., Tiveron, M. C., Bodmer, R. and Cremer, H.** (2004). Dynamics of Cux2 expression suggests that an early pool of SVZ precursors is fated to become upper cortical layer neurons. *Cereb Cortex* **14**, 1408-20.

Appendix 1:

Statement of Authorship

Expression of the murine transcription factor SOX3 during embryonic and adult neurogenesis

Publication Status: Published

Nicholas Rogers (Candidate)

Contribution to the Paper:

Performed experiments leading to figures 1, 2 and 3, supplementary figures 1 and 4 and Tables 1 and 2, analysed data, interpreted data and wrote the draft of the manuscript

I hereby certify that the statement of contribution is accurate

Signature:

Date

2/12/13

Pike-See Cheah

Contribution to the Paper:

Performed experiments leading to figures 4 and 5, analysed data, interpreted data, and wrote part of the manuscript

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:.....

.....Date: *20th November 2013*.....

Eva Szarek

Contribution to the Paper:

Performed experiments leading to supplementary figure 2, assisted with data interpretation and manuscript preparation

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis.

Signature:

.....Date: 30th November 2013

Kakoli Banerjee

Contribution to the Paper:

Performed experiments leading to supplementary figure 3, assisted with data interpretation and manuscript preparation

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:.....Date: 27th November 2013

Jeffrey Schwartz

Contribution to the Paper:

Supervised experiments and contributed to manuscript evaluation

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:..

.....Date: 20/11/2013.....

Paul Thomas

Contribution to the Paper:

Supervised experiments, assisted with data interpretation and manuscript preparation, acted as corresponding author

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:.....

.....Date: 2 / 12 / 13

Appendix 2:

Statement of Authorship

Mechanistic Insight into the Pathology of Polyalanine Expansion Disorders Revealed by a Mouse Model for X Linked Hypopituitarism

Publication Status: Published

Nicholas Rogers (Candidate)

Contribution to the Paper:

Generated Sox3 null embryonic stem cells. Performed experiments leading to figure 5, analysed data, interpreted data and helped evaluate the manuscript

I hereby certify that the statement of contribution is accurate

Signature:..

.....Date: 2/12/13.....

James Hughes

Contribution to the Paper:

Performed experiments leading to figures 1, 2 and 3, Supplementary figure 1, Table 1 analysed data, interpreted data, and co-wrote the manuscript

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:

Date: 2/12/13

Sandra Piltz

Contribution to the Paper:

Performed morula injections leading to the generation of chimeric mice used in figure 2a, 2b, 3, 5 and assisted with interpretation and manuscript preparation.

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:..... Date: 2/12/13

Dale McAninch

Contribution to the Paper:

Performed experiments leading to figure 4b and assisted with data interpretation and manuscript preparation

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:

.....Date: 2/12/13.....

Lynn Rowley

Contribution to the Paper:

Performed experiments leading to figure 4a and assisted with manuscript evaluation

hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:.....

.....Date: 20 | 11 | 2013

Paul Thomas

Contribution to the Paper:

Supervised experiments, assisted with data interpretation, co-wrote the manuscript and acted as corresponding author

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature:.....

.....Date: 2/12/13

Appendix 3:

Statement of Authorship

***Dbx1* is a direct target of SOX3 in the spinal cord**

Publication Status: Manuscript

Nicholas Rogers (Candidate)

Contribution to the Paper:

Performed experiments leading to figure 1,2,3,4,5a 5c, supplementary figure 1, analysed data, interpreted data and wrote the manuscript

I hereby certify that the statement of contribution is accurate

Signature: ..

.....Date:

2/12/13

Dale McAninch

Contribution to the Paper:

Performed experiments leading to figure 5b, helped in data interpretation and manuscript preparation

I hereby certify that the statement of contribution is accurate and I give permission for the inclusion of the paper in the thesis

Signature

.....Date: 2 / 12 / 13.....

Paul Thomas

Contribution to the Paper:

Supervised development of work, helped data interpretation and manuscript evaluation,
acted as corresponding author

I hereby certify that the statement of contribution is accurate and I give permission for the
inclusion of the paper in the thesis

Signature:.....

.....Date:..... 2/12/13