From Ludwig Institute for Cancer Research Ltd and the Department of Cell and Molecular Biology, Karolinska Institute, Stockholm, Sweden

FROM STEM CELL TO NEURON – A SOX PERSPECTIVE

Maria Bergsland

Karolinska Institutet

Stockholm 2010

Previously published paper was reproduced with permission from the publisher.

Published by Karolinska Institutet. Printed by Larserics Digital Print AB.

ISBN 978-91-7409-902-7

I am among those who think that science has great beauty. A scientist in his laboratory is not only a technician; He is also a child placed before natural phenomena, which impresses him like a fairy tale.

Marie Curie (1867-1934)

Abstract

The differentiation of stem cells into the more than 100 billion neurons that compose the central nervous system (CNS) is one of the most remarkable transformations during vertebrate development. A central problem of this process is how neural lineage selection initially is specified in pluripotent stem cells and maintained during the course of neurogenesis. This thesis focuses on how different Sox transcription factors of the HMG-box family control neural cell-type restricted gene expression from the early lineage specification stages to later stages of neuronal maturation.

In the two papers on which this thesis is based, we present a molecular pathway where diverse Sox proteins can act to coordinate neural gene expression as development proceeds, from the early neural lineage specification in pluripotent stem cells to the gene regulatory control operating in maturing post-mitotic neurons. We show that, already in pluripotent stem cells, Sox2 binds to inactive neural genes and that, in neural progenitors, Sox3 "takes over" and binds to activated neural genes, but also to silent genes destined to become active in maturing post-mitotic neurons. Finally, in post-mitotic neurons these genes are bound by Sox11. The binding of Sox3 to silent genes is associated with bivalently marked chromatin domains, containing both trimethylated H3K4 and H3K27, which is resolved to an active state upon Sox11 binding.

Proneural basic helix-loop-helix (bHLH) transcription factors have key roles in promoting NPCs to commit to a differentiation program leading to the generation of post-mitotic neurons. It is shown that Sox4 and Sox11 act downstream of proneural bHLH proteins and are of critical importance for the activation of neuronal protein expression but not for cell cycle withdrawal.

Together these data reveal a regulatory logic whereby sequentially acting Sox transcription factors preselect transcriptional programs that are destined to be activated at later stages of neural differentiation. Thus, a single family of transcription factors acts to coordinate neural gene expression from the early lineage specification to later stages of neural development.

PUBLICATIONS

I. **Maria Bergsland**, Martin Werme, Michael Malewicz, Thomas Perlmann, Jonas Muhr (2006)
The establishment of neuronal properties is controlled by Sox4 and Sox11
Genes & Development 20, 3475–3486

II. **Maria Bergsland***, Daniel Ramsköld*, Cécile Zaouter, Susanne Klum, Rickard Sandberg, Jonas Muhr (2010)
Neural lineage commitment is maintained through sequentially expressed Sox transcription factors
Manuscript

* These authors contributed equally

Other publications (not included in this thesis)

Elisabet Hermanson, Lotta Borgius, **Maria Bergsland**, Eliza Joodmardi, Thomas Perlmann
Neuropilin1 is a direct downstream target of Nurr1 in the developing brain stem.
J Neurochem (2006) 97(5): 1403-11.

Stina Friling, **Maria Bergsland**., Susanna Kjellander
Activation of Retinoid X Receptor increases dopamine cell survival in models for Parkinson's disease.
BMC Neurosci (2009) 10: 146.

CONTENTS

LIST OF SELECTED ABBREVIATIONS

AP	Anterior-Posterior
bHLH	Basic Helix-loop-helix
BMP	Bone morpgogenic protein
ChIP	Chromatin immunoprecipitation
ChIP-Seq	ChIP followed by DNA sequencing
CNS	Central nervous system
DV	Dorsal-Ventral
EC cells	Embryonic carcinoma cells
ES cells	Embryonic stem cells
HD	Homeodomain
HMG	High mobility group
ICM	Inner cell mass
Id	Inhibitor of differentiation
NICD	Notch intracellular domain
NPC	Neural progenitor cell
NRSF	Neuron restrictive silencing factor
PcG	Polycomb-group
PNS	Peripheral nervous system
POU	Pit-Oct-Unc
Shh	Sonic hedgehog
Sox	Sry related HMG box
TSS	Transcription start site

INTRODUCTION

The vertebrate central nervous system (CNS) consists of more than 100 billion cells that work together in a complex and strictly regulated network. Most of the neurons within the adult CNS are born during early stages of embryogenesis and are preserved throughout adulthood. This thesis focuses on the generation of neurons during embryonic stages, specifically the molecular mechanisms controlling the differentiation from neural stem or progenitor cells to mature neurons.

EARLY DEVELOPMENT OF THE VERTEBRATE EMBRYO

Shortly after fertilization, the vertebrate zygote starts to divide rapidly and a number of events will lead to the formation of a blastula. The inner cell mass (ICM) of the vertebrate blastula will give rise to the epiblast cells that will form the embryo, while cells from the trophoblast later form the extraembryonic tissue.

An important event of the early life is when cells become commited to one of the three different germ layers of the embryo. This phenomena, referred to as gastrulation, gives rise to the endodermal, ectodermal and mesodermal germ layers and starts with re-arrangement of the cells from the ICM of the blastula. The three germ layers are formed through a complicated series of events, were the primitive streak and primitive node (Hensen´s node in birds, Spemann´s organizer in amphibians) migrates anteriorly in the epiblast layer and set the anterior-posterior (AP), dorsal-ventral (DV) and left/right axis of the embryo. Epiblast cells migrating through the primitive streak and into the blastocoel form the endodermal and mesodermal layers which later give rise to bone, muscles, intestines and other inner organs. The outer epiblast cells give rise to the ectoderm that form the skin, the peripheral nervous system (PNS) and the central nervous system (CNS). (Gilbert, 2003)

When the primitive node later starts to regress posteriorly, the notochord is formed from the condensation of mesodermal (chordamesoderm) tissue, starting at the level of the future midbrain and extends caudally to form the tail of the embryo. The node and the notochord releases Noggin, Chordin and Follastin signaling molecules to block ectodermal Bone Morphogenetic Protein (BMP)- 4 and BMP-7 signaling in the

overlying surface ectoderm (Stern, 2005). The block of BMP reinforces the neural fate and thereby helps to establish the neural plate along the AP axis (Wilson *et al.*, 1997). Initial experiments to determine the inductive properties of the node were performed by C. Waddington in 1933 (Waddington and Waterman, 1933). In this study, Hensen's node was grafted from one chick embryo to another, producing an early embryo with two Hensen's nodes. The two nodes formed two separate notochords and Waddington found that the resulting embryo contained a second neural tube.

NEURAL TUBE FORMATION

During neurulation, surrounding cells direct the cells of the neural plate to proliferate and fold in upon itself to form a hollow tube (Smith and Schoenwolf, 1997). This process also divides the ectodermal cells into three sets of cells, the inner cells of the neural tube that form the spinal cord and brain, the outer epidermis that forms the skin and finally, the neural crest cells that eventually will form for example peripheral neurons and glial cells (Gilbert, 2003). Neurulation occurs in several steps including shaping and bending of the neural plate and closing of the neural tube.

Closure of the neural tube does not occur simultaneously throughout the AP axis. In birds for example, the closure is initiated at the level of future midbrain and spread in both directions. However, in mammals, the neural tube closure initiates at several places along the AP axis (Golden and Chernoff, 1993; Nakatsu *et al.*, 2000). Different neural tube defects are caused when various parts of the neural tube fail to close. In humans for instance, failure to close the posterior regions results in a condition known as *spina bifida*, whereas failure to close the anterior regions results in a lethal condition, *anencephaly*, where the forebrain remains in contact with the amnionic fluid and degenerates. Human neural tube closure depends on both genetic and environmental factors (Botto *et al.*, 1999). Genes such as *Pax3* and *Sonic hedgehog* (*Shh*) are essential for closure, but also dietary factors such as cholesterol and folic acid (vitamin B_{12}) appears to be critically important (Gilbert, 2003).

Apart from the massive proliferation of cells in the neural tube, the developing neural tube is also patterned along the anterior-posterior (AP) axis to form the brain and the spinal cord. The developing brain initially has the form of three vesicles which will

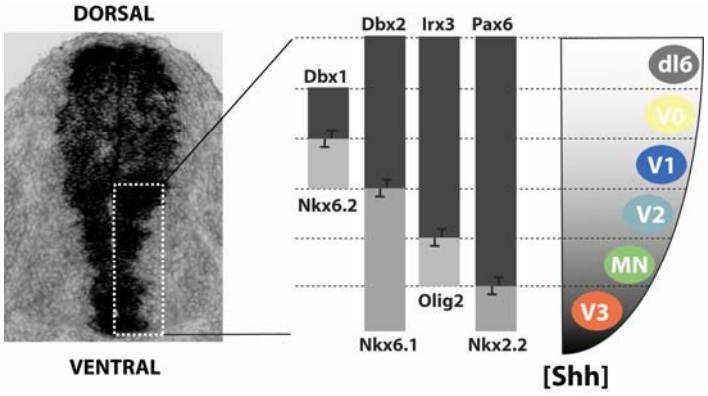
give rise to the forebrain (Prosencephalon), midbrain (Mesencephalon) and hindbrain (Rhombencephalon). These are futher subdivided as the development proceeds into 5 secondary vesicles, from anterior to posterior; telencephalon, diencephalon, mesencephalon, metencephalon and myelencephalon. AP patterning is regulated by the spatial and temporal distribution of signaling molecules that induces molecular differences in neural progenitor cells. The cells respond to these signals by activating the expression of transcription factors along the AP axis of the neural tube. For inctance, AP patterning results in the nested and combinatorial expression of Hox genes, which act cell autonomously to give cells an identity along the rostro-caudal axis (Hunt and Krumlauf, 1991).

DORSAL-VENTRAL PATTERNING

The dorsal-ventral (DV) patterning of the neural tube is important in order to produce correct subtypes of neurons at the correct location. Ventralization is initiated by secretion of the morphogen Sonic hedgehog (Shh) from the notochord, which induces the floor plate of the neural tube. The floor plate cells also secrete Shh and a protein gradient is formed with the highest concentration of Shh in the most ventral region. At the same time in the dorsal part of the neural tube, the epidermis establishes a second signaling center, the roof plate, which secretes members of the TGF-β family; BMP-4 and BMP-7. The roof plate is, in the same way as the floor plate, building up a signaling gradient, but this time from the dorsal direction of the neural tube and with BMP-4 and BMP-7. The combinations of different concentrations of the ventrally expressed Shh and the dorsally expressed BMPs result in a specific homeodomain (HD) gene expression pattern in progenitor cells along the DV axis. (Hollyday, 2001) (Glibert, 2003)

Homeodomain transcription factors are divided into class I and class II proteins based on their response to Shh. Class II proteins are induced by Shh and are expressed in the ventral parts of the neural tube, close to the Shh secreting floor plate. These proteins include homeodomain containing proteins such as Nkx, but also bHLH proteins such as Olig2. Class I proteins, for example Dbx1 and Dbx2, are repressed by the presence of Shh (Pachikara *et al.*, 2007) and are thus expressed by progenitor cells surrounded by low levels of Shh. Class I and Class II proteins subsequently form cross-repressive

pairs of transcription factors (Figure 3), which further refine the Shh induced gene expression pattern (Briscoe *et al.*, 2000). This establishes a transcription factor code in distinct areas of the progenitor population (Figure 1). Hence, the specification of neuronal subtypes is determined already at the progenitor state.



Figur 1. Section of a mouse spinal cord at embryonic day (E) 11.5 marked for a gene expressed in the proliferative ventricular zone (VZ). The VZ is divided in patterned regions, expressing different HD proteins. These domains will later give rise to the different neuronal subtypes of the vental spinal cord (V3, V2, V1 and V0 interneurons, and motorneurons. Dl6 are classified as dorsal neurons).

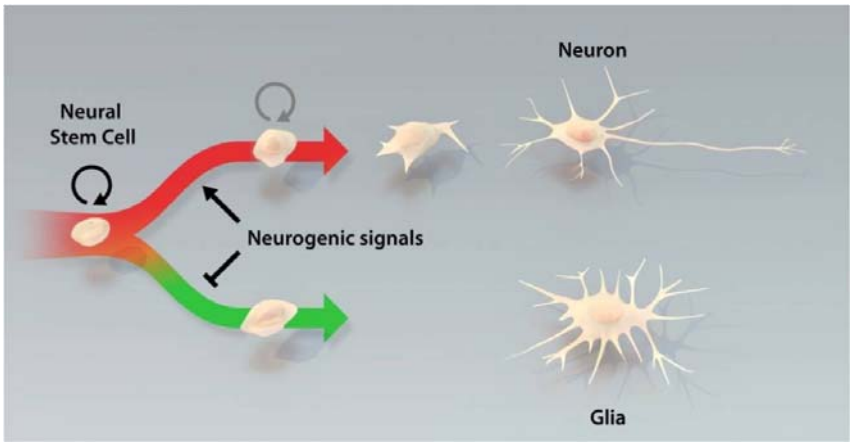
In the ventral half of the spinal cord, five distinct progenitor domains are formed (the p3, pMN, p2, p1 and p0 domains). These will later give rise to at least five neuronal subtypes; motor neurons and interneurons of class V3, V2, V1 and V0 (Figure 1) (Lupo *et al.*, 2006). Certain key determinants among the HD proteins have been identified through gain-and-loss of function experiments (Jessell, 2000). For example, mis-expression of Olig2 outside its normal expression domain in the spinal cord will induce the generation of ectopic motor neurons along the DV axis (Novitch *et al.*, 2001).

NEUROGENESIS

The pool of cells that constitute the early neural tube are the precursors of most mature cells of the adult CNS. Extrinsic and intrinsic signals will provide the neural progenitor cells the information that will influence the decision to proliferate and remain as progenitor cells or to exit the cell cycle and start to differentiate. The molecular mechanisms that regulate neural progenitor cells to differentiate into distinct cell types at the correct time and position, have been a major issue for developmental neurobiologists in the past years. This has resulted in the identification of complex signaling pathways and gene regulatory networks that underlie many of the early steps

of neurogenesis. The course of neurogenesis comprises several cellular stages, all with a specific and unique gene expression pattern. This section focuses on the neural progenitor or stem cell state and some regulatory signaling pathways important for this specific cell stage.

Neural progenitor cells are located within the ventricular zone of the developing CNS and differentiate in a temporal and spatially conserved manner. Using the developing neural tube as a model for CNS differentiation, this means that un-differentiated cells are located most medially and migrate marginally during the process of differentiation. This process takes place at most places of the neural tube, from anterior to posterior as well as from dorsal to ventral. When the progenitor cells receive signals to differentiate, they exit the cell cycle, down-regulate progenitor specific genes and sub-sequentially start to express neuronal features as differentiation proceeds. Thus, the differentiating cell undergoes distinct cellular stages before it has formed a mature neuron (Figure 2).



Figur 2. Model figure showing the differentiation process from neural progenitor cell to neuron or glia. Cells go through several intermediate steps before they reach their final identity. (Illustration by M. Karlén)

Proneural Proteins

Genetic studies in fruit flies (*Drosophila*) have shown that a number of genes, known as proneural genes, play essential roles during neurogenesis (Bertrand *et al.*, 2002). In vertebrates, there are two groups of proneural proteins with *Drosophila* homologous *achaete-scute* (*asc*) and *atonal* (*ato*) (Campuzano and Modolell, 1992). *Asc* related proteins include asc-like (*ascl*) in vertebrates (*Mash* in mouse), whereas *ato* related proteins are for instance *Neurogenins* (*Ngn*) and *NeuroD* in vertebrates. In mammals, Mash1 and Neurogenins are the main proneural proteins and have been shown to

regulate several aspects of neurogenesis, including promotion of cell cycle exit, commitment of neural progenitor cells to a neuronal fate and subtype identity specification (Bertrand *et al.*, 2002; Nakada *et al.*, 2004). In addition to the proneural function of these proteins, there is also an inhibitory function of glial differentiation. Studies made in cortical primary cultures from rat revealed that Ngn1 inhibits glial differentiation at several molecular levels (Sun *et al.*, 2001).

Tissue-specific basic-helix-loop-helix (bHLH) transcription factors, such as some proneural proteins, bind DNA as heterodimers with ubiquitously expressed bHLH proteins such as the E proteins E12 and E47 (Massari and Murre, 2000). Since heterodimerization is required for DNA binding of proneural proteins, the factors that interfere with this heterodimerization act as passive repressors of proneural gene activity. Example of such proteins is *Inhibitor of differentiation* (*Id*) in vertebrates that will, thus, have the opposite activity compared to proneural proteins and maintain cells undifferentiated.

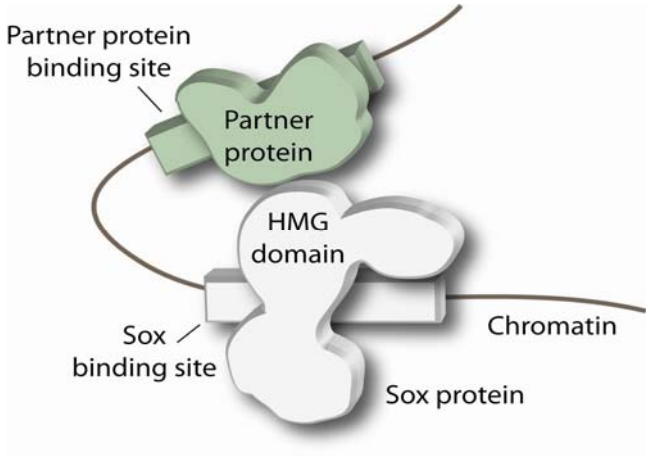
Notch signaling

The Notch signaling pathway plays a fundamental role in the decision whether progenitor cells enter a neurogenic program or remain as self-renewing progenitor cells. It has been shown, both in vertebrates and *Drosophila*, that disruption of the Notch pathway results in an increased generation of neurons (Hitoshi *et al.*, 2002; Yoon and Gaiano, 2005; Louvi and Artavanis-Tsakonas, 2006) indicating that active Notch signaling prevents cells from differentiating. In mammals, the Notch family of signaling proteins includes four receptors (Notch1-4) and several ligands (Delta-like1, 3, 4 and Jagged1, 2) compared to the single Notch receptor and two ligands (Delta and Serrate) in *Drosophila* (Yoon and Gaiano, 2005; Bray, 2006). Ligand activation results in intra-membranous cleavage of the receptor by a complex called γ-secretase that releases the intracellular domain of Notch (NICD). The NICD translocates into the nucleus and interacts with DNA to activate downstream target genes such as *Hes* and *Hey*. The bHLH transcription factors Hes and Hey (Enhancer of split in *Drosophila*) repress transcription of proneural genes (Corbin *et al.*, 2008) and thereby keep the Notch receptor expressing cells in an undifferentiated state. In contrast, the neighboring ligand-expressing cells are able to respond to proneural proteins and can, thus, differentiate. These series of events are known as lateral inhibition and induce and

stabilize differences between neighboring cells and define cells in a population that commit to differentiation and cells that remain as progenitors (Artavanis-Tsakonas *et al.*, 1999).

Sox in Neurogenesis

The Sox (Sry related HMG box) family of transcription factors plays key roles in CNS development. Members of this family have high sequence similarities within their DNA-binding HMG domain. About 20 different vertebrate Sox proteins are known and have been divided into subgroups (A-H) based on their amino acid sequence similarity within the HMG domain (Table 1) (Kamachi *et al.*, 2000). Two major features suggest that Sox proteins have unique and important roles during development of the vertebrate embryo. First, Sox proteins are expressed in most organs, at least at some step during development (Wegner, 1999).



Figur 3. Sox proteins bind to the minor groove of DNA and induce a 70-85° bend. This allows contact of Sox proteins with distantly bound partner factors.

Second, these HMG-box containing proteins are unique in their way of binding to and bending DNA. Unlike most transcription factors that bind DNA in major groove, Sox proteins bind to minor groove that induces a 70-85° bend of DNA (Ferrari *et al.*, 1992; Connor *et al.*, 1994). This allows Sox proteins to bind DNA in a close proximity to other transcription factors that bind the major groove, which has lead to the hypothesis that Sox proteins, at least in part, function by organizing local chromatin structure to make it accessible for other transcription complexes (Werner and Burley, 1997; Wolffe, 1994) (Figure 3). Support for this hypothesis is that Sox proteins are often found to bind DNA together with other transcription factors, such as members of the Pit-Oct-Unc (POU) family (Kamachi *et al.*, 2000; Wegner, 1999).

Group	*M. Musculus*	*D. Melanogaster*	Transactivating domain
SoxA	Sry		Activating
SoxB1	Sox1 Sox2 Sox3	SoxN	Activating
SoxB2	Sox14 Sox21	SoxB2.1 SoxB2.2 SoxB2.3	Repressing
SoxC	Sox4 Sox11 Sox12	SoxC	Activating
SoxD	Sox5 Sox6 Sox13	SoxD	None
SoxE	Sox8 Sox9 Sox10	Sox100B/SoxE	Activating
SoxF	Sox7 Sox17 Sox18	SoxF	Activating
SoxG	Sox15		Activating
SoxH	Sox30		Activating

Table 1. Classification of the different members of the Sox family of transcription factors in mouse and Drosophila and the type of transcriptional function for the different groups.

Generally, Sox proteins are expressed in pairs or even in triplicates and they are often suggested to have redundant functions (Wegner, 1999; Wegner and Stolt, 2005). In early mouse and chick CNS, Sox1, 2 and 3 (SoxB1) are highly expressed and their expression patterns are overlapping. These proteins have a crucial function during embryonic neurogenesis to maintain progenitor cells in a self-renewing state. Over-expression, *in ovo*, of either SoxB1 protein, effectively reduces differentiation of neuronal progenitors by blocking the proneural activity of bHLH proteins (Bylund *et al.*, 2003; Graham *et al.*, 2003). This is separate from Notch signaling, which functions to repress proneural gene expression (Holmberg *et al.*, 2008). On the other hand, Sox21, which is co-expressed with SoxB1 in neural progenitor cells, have the opposite function and promote differentiation. When Sox21 is over-expressed in chick spinal cord, cells rapidly exit the cell cycle, down-regulate progenitor proteins and start to express neuronal markers. The balance between SoxB1 proteins and Sox21 in progenitor cells appears to be a determinant whether cells commit to differentiation or not (Sandberg *et al.*, 2005; Uchikawa *et al.*, 1999).

However, SoxB1 members are not completely overlapping in their expression arguing against a complete functional redundancy. Sox2 is already expressed before the blastula stage in mouse (Zappone *et al.*, 2000; Miyagi *et al.*, 2004) and continuously in the ICM of the blastula (Wood and Episkopou, 1999; Avilion *et al.*, 2003). Sox2 deficient mice lack epiblast cells, indicating the first critical requirement for SoxB1 proteins (Avilion

et al., 2003). Mouse epiblast cells also express a low amount of Sox3, however, loss of Sox3 appears to be compensated by Sox2 (Collignon *et al.*, 1996; Wood and Episkopou, 1999). In chick, Sox3 is the dominantly expressed SoxB1 protein prior to gastrulation, suggesting evolutionary diversity between species (Rex *et al.*, 1997).

Shortly after gastrulation, SoxB1 expression is restricted to the neuroectoderm in both vertebrates and *Drosophila* (Cremazy *et al.*, 2000; Avilion *et al.*, 2003). In addition to the SoxB1 already expressed, additional SoxB1 are turned on in the neuroectoderm. In mouse, for instance, Sox2 and Sox3 are already expressed and Sox1 expression is induced (Collignon *et al.*, 1996; Wood and Episkopou, 1999). As neurogenesis proceeds, SoxB1 protein expression becomes restricted to the progenitor cells in the proliferative zones of the CNS. Mice deficient for one SoxB1 member show only minor effects of the developing brain and spinal cord, suggesting a redundant function for SoxB1 during this particular time and place in development (Nishiguchi *et al.*, 1998; Ferri *et al.*, 2004; Rizzoti *et al.*, 2004).

Sox1, Sox2 and Sox3 can also be found in neurogeneic regions of the adult mouse CNS such as the subventricular zone of the lateral ventricle, the rostral migratory stream and the hippocampus (Pevny and Placzek, 2005). Moreover, adult neurogenesis is reduced in Sox2 deficient mice (Ferri *et al.*, 2004), indicating a role for SoxB1 in this process.

Sox2 also has a function in the developing retina. In Sox2 deficient mice, retinal progenitor cells fail to maintain as proliferative progenitors and are differentiating (Kamachi *et al.*, 1998; Kamachi *et al.*, 2001). Interestingly, in the retina Sox2 has been suggested to block differentiation by increasing Notch expression. Chromatin immunoprecipitation (ChIP) of Sox2 in mouse retina revealed binding of Sox2 to a Notch1 enhancer and decreased levels of Sox2 were followed by reduced Notch1 levels (Taranova *et al.*, 2006).

However, SoxB1 members are not only having a role in preventing premature differentiation, but are also involved in neuronal maturation. Sox2 is required for maturation of GABAergic neurons in cortex and olfactory bulb (Cavallaro *et al.*, 2008). Also, expression of other SoxB1 proteins is found in distinct subset of neurons, for example Sox1 in selective areas of striatum and Sox3 in hypothalamus (Malas *et al.*,

2003; Rizzoti *et al.*, 2004). Thus, although SoxB1 proteins play overlapping roles in progenitor cells, they might have additional functions in differentiating cells.

Sox in Gliogenesis

SoxE proteins (Sox8, 9 and 10) are important for the formation of myelinating glial cells; Schwann cells in the PNS and oligodendrocytes in the CNS (Chew and Gallo, 2009). Sox10 is required for specification, lineage progression and terminal differentiation of Schwann cells (Schreiner *et al.*, 2007). During the development of these cells, Sox10 seems to exert its function through interaction with other transcription factors, for example the POU protein Oct6, which are stage specifically expressed during Schwann cell differentiation (Ghislain and Charnay, 2006).

Regarding oligodendrocyte development, Sox9 is initially important for their specification and generation (Stolt *et al.*, 2003). Sox10 expression is initiated in the oligodendrocyte precursors and is together with Sox9 involved in maintaining this state (Finzsch *et al.*, 2008). Sox10 alone is required for terminal differentiation of oligodendrocytes and interacts with at least one member of the bHLH Olig protein family to induce the expression of *Mbp* (Myelin binding protein) (Li *et al.*, 2007).

SoxC proteins (Sox4, Sox11 and Sox12) have, compared to Sox9 and Sox10, an opposite role in oligodendrocyte differentiation. Prolonged expression of Sox4 *in vivo* under the *Mbp* promoter resulted in reduced myelin gene expression in the spinal cord (Potzner *et al.*, 2007). Sox11 was in another study found to be a repressor of the oligodendroglial lineage (He *et al.*, 2007). SoxC proteins do not interfere with the Sox10 mediated up-regulation of myelin gene expression and it is believed that the SoxC inhibitory function of myelination is an indirect effect on the *Mbp* gene (Potzner *et al.*, 2007).

SOX AND THEIR PARTNERS

As previously mentioned, Sox proteins are expressed in most organs of the developing embryo and have been identified as one of the most important groups of developmental regulators in both vertebrates and invertebrates (Bowles *et al.*, 2000; Wegner, 1999). They are suggested to select their target gene repertoire as well as in their partner factors in a cell specific manner (Miyagi *et al.*, 2009). Although different Sox proteins show no or little distinction in recognizing DNA sequences *in vitro,* Sox proteins show a distinction in their sequence binding specificity *in vivo* (Miyagi *et al.*, 2009). This could imply that the specificity of Sox binding to DNA lies in the selection of partner factors, rather than in the Sox binding sequence itself. This section will bring up some situations describing the interactions between Sox and their partners.

In embryonic stem (ES) cells, Sox2 forms a complex with Oct4 to activate a number of target genes. This is also the case in embryonic carcinoma (EC) cells, where binding of both factors to their recognition sites is essential for gene activation since mutation in either one or the other completely abolish target gene transcription (Yuan *et al.*, 1995; Ambrosetti *et al.*, 1997; Ambrosetti *et al.*, 2000). For example, the embryonic cell transcription factor 1 (UTF1) gene has been demonstrated to be regulated by the Sox2/Oct4 complex in both ES and EC cells and the UTF1 regulatory region can selectively recruit the Sox2/Oct4 complex and prevent the binding of similar complexes such as Sox2/Oct1 and Sox2/Oct6 (Nishimoto *et al.*, 1999).

In lens cells, Sox2 binds together with Pax6 to activate δ-crystallin through the DC5 enhancer and *in vivo* co-expressed Sox2 and Pax6 was found on the DC5 enhancer using ChIP. However, there was no binding of either one of them when expressed alone, suggesting that Sox2 nor Pax6 are not able to stably bind DNA on their own. This cooperation seems to be restricted to SoxB1, since Sox9 fails to activate DC5 enhancer *in vitro* even though it binds as strongly as SoxB1 members to the enhancer (Kamachi *et al.*, 2001; Kondoh *et al.*, 2004). Thus, interactions between Sox and their partner factors seem also important for the stability of the DNA bound complex and not only for the target sequence selection.

EMBRYONIC STEM CELLS

The first human embryonic stem (ES) cell line was established in 1998 (Thomson *et al.*, 1998). Although public criticism and ethical debates concerning the use of human material in science, these cells can provide an unlimited source of human cells for therapeutic use. Moreover, *in vitro* cultured ES cells have become an important tool for understanding mammalian development and specifically neuronal differentiation. Derived from the inner cell mass of the blastocyst, they are the precursors of all cells that form the embryo and during the last decade, efforts have been put into understanding the pluripotent capacity of these cells.

Sox together with Oct and Nanog

Functional studies have been important in identifying a group of transcription factors that affect the pluripotent identity of ES cells (Niwa *et al.*, 2000; Chambers *et al.*, 2003; Ema *et al.*, 2008). Within this group Oct4 (Pou5f1), Nanog and Sox2 are found to be essential (Chambers *et al.*, 2007; Masui *et al.*, 2007). These three transcription factors frequently bind closely located sites in the ES cell chromatin (Boyer *et al.*, 2005). Oct4 belongs to the POU family of transcription factors that binds DNA through two domains, a high affinity and a low affinity domain that binds on either side of the helix to encircle the DNA (Phillips and Luisi, 2000). Nanog binds DNA through a single homeodomain that is similar to the homeodomain of the Nkx transcription factors, but not to the extent that would make them part of the same family (Kappen *et al.*, 1993).

Deletion of either Sox2 or Oct4 in ES cells results in failure to maintain proliferative and pluripotent capacities and cells differentiate into trophoectoderm (Niwa *et al.*, 2000; Masui *et al.*, 2007). Surprisingly, many of the Oct/Sox target genes were not greatly affected by Sox2 deletion and the key contribution of Sox2, according to this study, is to maintain Oct4 expression. This is supported by the finding that Oct4 can rescue ES cells from differentiation upon Sox2 loss (Masui *et al.*, 2007). It has also been suggested that Sox2 function could be compensated by other Sox proteins expressed in ES cells, such as Sox4 and -11 and Sox15 (Chambers and Tomlinson, 2009). However, Sox2 deficient ES cells have indeed lost their pluripotency (Masui *et al.*, 2007) and the expression levels of at least Sox4 and Sox11 are relatively low in ES cells (Shen *et al.*, 2008). Reduction of Nanog levels also increases the ability of cells to

differentiate (Hatano *et al.*, 2005; Ivanova *et al.*, 2006; Chambers *et al.*, 2007), however, ES cells completely lacking Nanog continue to proliferate (Chambers *et al.*, 2007). Oct4 over expression in ES cells results in differentiation into a mixed population of cells, expressing both endodermal and mesodermal markers (Niwa *et al.*, 2000). This differs from Sox2 over expression where the cells differentiate mainly towards the ectodermal and specifically the neuronal lineage (Zhao *et al.*, 2004; Kopp *et al.*, 2008).

Even though phenotypic effects are clearly seen after manipulating the levels of these transcription factors, the mechanisms for their impact in pluripotency and self-renewal of ES cells remains unclear. A number of genome wide binding screens and gene expression arrays have been preformed the last couple of years as an attempt to solve this question. One of the pioneering studies includes promoter regions (8kb (kilobases) upstream to 2kb downstream of transcription start site (TSS)) of ~17,000 human genes (Boyer *et al.*, 2005). It revealed that Sox, Oct and Nanog co-occupy a number of human promoter regions that correspond to both actively transcribed and silent genes in ES cells. Since their function is to maintain the ES cell state, this suggests that loss of these factors upon differentiation results in decreased levels of ES cell maintenance factors and increased levels of proteins important for development. However, Sox and Oct and in particular Nanog also bind to a large set of promoter regions on their own, which could correspond to the phenotype differences observed in gain-and-loss of function experiments.

Other transcription factors are also associated with Oct, Sox and Nanog in ES cells. Several studies have shown that the close localization on DNA of Nanog to Sox/Oct is not Nanog specific but is also seen with other transcription factors including Klf4 (Krupple-like factor 4), Esrrb (estrogen-related receptor beta) and Tcf4 (T cell factor 4) (Chen *et al.*, 2008; Cole *et al.*, 2008; Kim *et al.*, 2008). A biotin-based affinity assay revealed several potential binding partners for Oct4 and Nanog, including proteins of nuclear receptor and estrogen receptor families, Hdac2 (histone deacetylase 2) and the transcriptional co-factor Sp1 (Wang *et al.*, 2006). However, the most studied co-factor of Oct4, namely Sox2, was not identified in this screen, possibly because the Oct/Sox complex is most stable when bound to DNA (Ambrosetti *et al.*, 1997; Chambers and Tomlinson, 2009). Several of these proteins were further confirmed to bind DNA close to Oct4 and/or Nanog in a following study by Liang et al. (Liang *et al.*, 2008). To

summarize, more than 1000 promoter regions are bound by four or more of these factors, Oct, Sox and Nanog included (Chen *et al.*, 2008; Kim *et al.*, 2008).

Epigenetic mechanisms

Recently, several studies have focused on epigenetics and chromatin modifications in human and mouse ES cells. It has been found that a proportion of silent genes are kept in a 'bivalent' or 'poised' state, in order to allow a fast induction upon differentiation. This section will present the most important findings and concepts regarding this issue.

Proteins from the Polycomb-group (PcG) complex stabilize transcriptionally repressed chromatin. Polycomb repressive complex 2 (PCR2) functions as a histone methyltransferase on lysine 27 (K27) of histone 3 (H3), which results in trimethylation of H3K27 (H3K27me3) and locally repressed chromatin (Cao and Zhang, 2004). H3K27me3 recruits the PRC1 complex to the chromatin, which in turn mediates gene silencing through mechanisms not completely understood, but involving histone H2A ubiquitination (Schuettengruber *et al.*, 2007). The H3K27me3 mark is usually separated from trimethylation of lysine 4 of hisone 3 (H3K4me3), which is connected to actively transcribed regions (Strahl and Allis, 2000). Interestingly, these two chromatin marks were found to be co-localized at certain genomic regions in mouse ES cells (Bernstein *et al.*, 2006). Furthermore, these bivalently marked regions were associated with silent genes or genes that were expressed at low levels and half of them contained binding sites for Oct, Sox and Nanog (Bernstein *et al.*, 2006). In a simultaneous study, Lee et al. (2006) showed that members of the PRC2 complex, Suz12 and Eed, co-localized with Oct4, Sox2 and Nanog in human ES cells, and that PRC2 target genes were found to be preferentially activated upon ES cell differentiation (Lee *et al.*, 2006). The authors emphasize that a specific set of genes must be actively repressed in ES cells to maintain the pluripotent capacity and that these genes are poised for activation upon future differentiation.

Above mentioned findings indicate that epigenetics are required for the maintenance of pluripotency. However, PRC complexes do not seem to be the only regulatory pathway in this context. This is clearly evident in ES cells deficient for Eed, a member of the PRC2 complex, where self-renewal is still occurring despite a dramatic decrease of H3K27me3 and increased expression level of PRC target genes (Boyer *et al.*, 2005;

Azuara *et al.*, 2006). Furthermore, ES cells lacking *Ring1β,* a component of PRC1 complex, still proliferate even though they show a decreased amount of histone H2A ubiquitination (de Napoles *et al.*, 2004).

The DNA methylation pattern in ES cells can also be correlated to the activity of genes, according to a study in mouse ES cells (Fouse *et al.*, 2008). Here, the majority of methylated genes have been classified as late differentiation-associated genes that are repressed in ES cells. In contrast, unmethylated DNA is highly correlated to genes that are active in ES cells and are largely associated with transcriptional activities and cell survival processes that are likely to contribute to the self-renewal of ES cells. Interestingly, a group of the unmethylated genes were differentiation-associated genes that were not expressed in ES cells. However, these genes were marked by the alternative silencing mark H3K27me3. It was further found that the majority of Polycomb-group repressed genes were unmethylated in these cells and that DNA methylation levels are dependent on histone methylation (Fouse *et al.*, 2008; Hattori and Shiota, 2008). Unlike PRC2-mediated repression, DNA methylation does not appear to be dependent on the binding on the Oct/Sox and Nanog, thus it seems possible that DNA methylation is a silencing mechanism that operates on a different set of genes from Polycomb repressors.

AIMS

My main focus during these years has been to gain insight into the molecular mechanisms that regulates neurogenesis. How are cells committed to become neurons? How are neural progenitor cells instructed to stop proliferate and to start a differentiation program? How is a neural lineage specific gene expression program regulated?

More specifically:

- What are the regulatory roles of SoxC proteins during the transition of differentiating progenitor cell into mature neurons?

- What roles have SoxB1 and SoxC transcription factors in the specification and maintenance of neural lineage specific gene expression?

RESULTS AND DISCUSSION

Papers will be referred to according to their roman numbers.

SOXB1 TARGET GENES IN NEURAL PROGENITOR CELLS

As previously described, SoxB1 proteins function to maintain the progenitor cell state (Bylund *et al.*, 2003; Graham *et al.*, 2003), but how this is achieved at the molecular level is not understood. In paper II, we focused on the target genes of SoxB1 transcription factors. ChIP-seq (Chromatin immunoprecipitation combined with high throughput sequencing) revealed 9,719 binding sites (peaks) for Sox3 in mouse ES cell derived neural progenitor cells (NPCs). By correlating binding sites to the closest transcriptional start site (TSS), 4,006 target genes were identified, as many genes were surrounded by several binding sites. A number of findings support that our genome wide screen was successful. First, the fairly high number of target genes is comparable to what has been found for other transcription factors using similar techniques (Kim *et al.*, 2008). Second, a binding motif that corresponded to the Sox consensus site was found in ~70% of the peaks (Wiebe *et al.*, 2003). The motif was found less frequent among peaks that were closely located to a TSS, indicating that Sox3 mainly binds to distal enhancers. In addition, this finding support the suggestion that the binding prediction of HMG domain proteins can not only be based on the DNA sequence itself (Lefebvre *et al.*, 2007).

Moreover, a comparison to binding sites of the transcriptional co-factor p300 in various tissues (Visel *et al.*, 2009) revealed that the target genes for Sox3 were highly enriched for p300 targets in forebrain compared to limbs. Given that p300 is used as a mark for actively transcribed genes, these data suggest that Sox3 binds to enhancer regions that are specifically active in developing CNS.

Sox3 targets both active and silent gene in NPCs

Interestingly, Sox3 targets both active and silent genes within progenitor cells, a finding that is comparable to the situation for Sox2 in ES cells (Boyer *et al.*, 2005; Chen *et al.*, 2008). We found that Sox3 target genes in NPCs overlap with >50% to Sox2 target genes in ES cells (Chen *et al.*, 2008), an interesting finding due to their similar role in the two different cell stages. This finding also defines three sets of genes; genes bound

by Sox2 in ES cells only, genes occupied by Sox2 in ES cells and later bound by Sox3 in the NPCs and, finally, genes that are bound by Sox3 in NPCs that are never occupied by Sox2 in ES cells. Importantly, since a specific antibody against Sox3 was used in the NPC screen and Sox3 is not expressed in ES cells, we can exclude contamination of SoxB1 target genes from ES cells. By comparing all Sox2-bound and Sox3-bound target genes in ES cells and NPCs respectively to previously published gene expression arrays for some major cell types and stages (Shen *et al.*, 2008; Doyle *et al.*, 2008; Hartl *et al.*, 2008), we found that the three groups of genes had diverse expression patterns. Genes bound by Sox2 only were highly enriched for genes specifically expressed in ES cells, whereas genes shared by Sox2 and Sox3 were mainly expressed in NPCs. Genes bound by Sox3 in NPCs and that had not been previously occupied by Sox2 were primarily expressed in mature neurons and/or glial cells. Thus, Sox2 and Sox3 bind to genes that are active in ES cells and NPCs, respectively, but they also bind to a significant number of silent genes that are activated at subsequent stages of neurogenesis. Genes poised for later activation have been associated with a bivalent chromatin signature, consisting of both activating H3K4me3 (histone 3 lysine 4 trimethylation) and repressing H3K27me3 (histone 3 lysine 27 trimethylation) modifications. We found that Sox2 binding to bivalent genes in ES cells preselects for genes activated along the neural lineage. Taken together, this argues that Sox2 regulates preferably ectodermal genes in ES cells rather than mesodermal or endodermal, where silent genes are later bound by Sox3 and become active during the neural progenitor state. This could indicate the initial molecular steps regulating neural lineage selection.

It is important to note that despite the similarities in the binding pattern of Sox2 in ES cells and Sox3 in NPCs, they were targeted to different sites in the majority of their common target genes. The binding motifs found for Sox2 (in ES cells) and Sox3 (in NPCs) bound regions were highly similar to each other and to previously identified Sox consensus site (van Beest *et al.*, 2000) and can therefore not explain the distinct target site selection. This argues again that the binding prediction can not only be based on the DNA sequence itself, but also on context-dependent differences such as the access to partner factors (Kondoh *et al.*, 2004; Lefebvre *et al.*, 2007).

Sox3 activates NPC genes

To better understand the functional role of SoxB1 on its bound target genes, we used a gain-of-function strategy, where Sox3 was over-expressed in ES cell derived NPCs under the control of the Nestin enhancer (Nes-Sox3) (Lothian and Lendahl, 1997). In line with previously published data for SoxB1 (Bylund *et al.*, 2003; Graham *et al.*, 2003), Nes-Sox3 cells fail to differentiate and even after 12 days of differentiation cells were maintained in a progenitor state, as shown by the expression of Sox1. These cells also failed to differentiate towards the glial lineage. Gene expression microarray analysis revealed that aproximately 350 genes had an increased expression level after Sox3 over-expression, whereas over 800 genes had a decreased expression level. However, correlation between the ChIP-seq experiments and the microarray data showed that genes increased in Nes-Sox3 cells were strongly over-represented of genes bound directly by Sox3. Moreover, although Sox3 binds both to NPC genes and genes that are expressed in neurons, only NPC genes were up-regulated in Nes-Sox3 cells.

SOXC ESTABSLISHES NEURONAL GENE EXPRESSION

Proneural bHLH transcription factors have key roles in inducing cell cycle exit and committing neural progenitor cells to neuronal differentiation (Farah *et al.*, 2000; Bertrand *et al.*, 2002; Lo *et al.*, 2002). Even though proneural proteins are able to promote a full differentiation program within the progenitor cell, the molecular mechanisms for their activity are not fully understood. Interestingly, the expression of proneural genes is generally down-regulated at an early step of neurogenesis and before differentiating cells start to up-regulate neuronal characters. This implies the existence of a downstream regulator, connecting proneural protein activity with the establishment of a neuronal phenotype.

In paper I, we examined whether SoxC proteins act downstream of proneural protein function. We found that SoxC gene expression is initiated as cells have become post-mitotic and are in the process of down-regulating progenitor specific genes, including proneural proteins. In the developing chick spinal cord, the expression of SoxC is continued in mature neurons. The inactivation of Sox4, Sox11 or Sox12 in mice does not reveal any significant role of these genes during neurogenesis, however functional

redundancy has been suggested (Cheung *et al.*, 2000; Sock *et al.*, 2004; Hoser *et al.*, 2008; Dy *et al.*, 2008). By using a gain-and-loss of function approach in the chick spinal cord, we could show that SoxC proteins indeed have a neurogenic effect. Mis-expression of Sox4 or -11 forced progenitor cells to rapidly up-regulate the expression of a subset of neuronal genes. This up-regulation was independent of cell cycle exit and down-regulation of progenitor specific genes, which mechanistically separates cell cycle withdrawal from the later parts of neurogenesis. By using an enhancer that has been shown to be sufficient to drive the expression of the neuronal gene Tubb3 (Dennis *et al.*, 2002), we found SoxC activity to directly regulate neuronal enhancers. This will be further discussed in section '*Sox11 target genes in early neurons*'.

In line with our gain-of-function experiments, siRNA knock-down experiments of Sox4 and Sox11 resulted in reduced levels of neuronal markers, suggesting that SoxC proteins are required for the induction of neuronal properties in differentiating post-mitotic neuroblasts. The third SoxC member, Sox12, was not yet identified at the time of the experiments for paper I, but has later shown to be functionally redundant to Sox4 and -11 (Dy *et al.*, 2008; Hoser *et al.*, 2008). In our loss-of-function experiments functional redundancy is still likely to exist but is probably blocked due to the high sequence similarities between Sox12 and the two other SoxC members in the regions targeted by the siRNAs that were used.

At what stage during the differentiation process is SoxC crucial? We found that the reduction of SoxC expression did not unable NPCs to exit the cell cycle. However, we noted that when SoxC levels were reduced, there were decreasing amounts of cells expressing the intermediate marker NeuroM. We took advantage of the fact that NeuroM is expressed in the transition period when cells go from expressing progenitor specific genes to express neuronal markers. The amount of cells expressing NeuroM that also still expressed progenitor genes was unchanged after reduction of SoxC proteins. On the other hand, the amount of further differentiated cells expressing NeuroM together with neuronal genes was decreased during these conditions. Thus, we conclude that upon loss of SoxC proteins, cells fail to differentiate any further from the stage where they normally start to express neuronal genes and instead they enter an apoptotic program as indicated by TUNEL staining.

PRONEURAL FUNCTION REQUIRES SOXC

Proneural bHLH proteins function to initiate and promote neurogenesis whereas SoxC proteins function to induce the expression of neuronal proteins in differentiating cells. What is then the connection between SoxC and proneural proteins? In paper I, several experiments show that SoxC appear to be downstream of proneural protein activity. Over-expression in chick spinal cord of either Neurogenin2 (Ngn2) or Mash1 induces ectopic expression of Sox4 and Sox11. The connection between proneural proteins and SoxC gene expression is further exemplified by mis-expression of Id that acts to block the activity of proneural proteins, which resulted in Sox4 and Sox11 decreased expression. Furthermore, we showed that SoxC proteins are also under the control of the neuronal gene repressor NRSF/REST, possibly to prevent premature expression of SoxC in cells expressing low levels of proneural proteins. Interestingly, NRSF/REST is known as a repressor of neuronal genes in non-neuronal tissue as well as in neuronal progenitor cells (Chong *et al.*, 1995; Schoenherr and Anderson, 1995; Bruce *et al.*, 2004).

Since proneural proteins are sufficient to induce a complete neuronal differentiation program, over-expression of Ngn2 in chick spinal cord will promote cells to exit the cell cycle, down-regulate progenitor specific genes and start to express neuronal proteins. However, in the absence of SoxC proteins, Ngn2 was able to promote cell cycle exit but failed to induce the expression of neuronal proteins. Thus, Ngn2 was able to promote its functions upstream of normal SoxC expression, but failed specifically to induce the expression of neuronal genes in the absence of SoxC proteins. These results, presented in paper I, indicate that SoxC proteins are required for the complete neurogenic activity of Ngn2

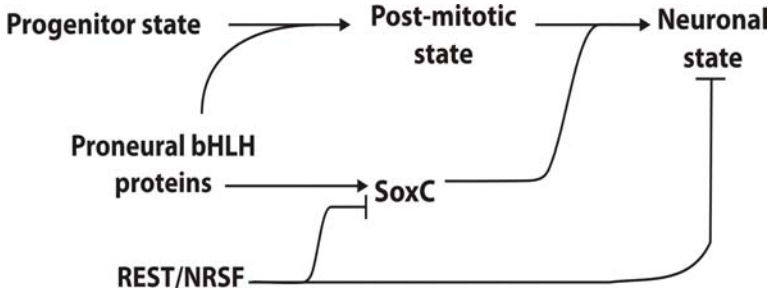


Figure 4. SoxC is expressed under the control of proneural function and REST activity and is required for proneural proteins to promote the expression of neuronal genes.

Although SoxC proteins promote neuronal differentiation, these proteins have an opposite repressing role during glial cell development (Potzner *et al.*, 2007; He *et al.*, 2007). This further demonstrates the diverse context-dependent roles of Sox proteins during development.

SOXC TARGET GENES IN EARLY NEURONS

In our ChIP-seq screen we found that Sox3 binds to a proportion of silent genes in NPCs that are, later during the differentiation process, expressed in neurons or glial cells. Since SoxC proteins have a role in activating neuronal genes upon differentiation, we examined if the silent neuronal genes that are occupied by Sox3 in NPCs are targeted by Sox11 in differentiating neurons.

ChiP-seq analysis revealed 3,644 target genes for Sox11 in ES cell-derived post-mitotic early neurons. We found that Sox11 binding sites (peaks) highly overlapped with binding sites for Sox3 in NPCs (92%), which corresponded to 69% of all found Sox3 binding sites. The shared target genes were primarily found to be expressed in two populations of cells. We found that the target genes were expressed in early neurons, which is in line with our previous data that SoxC can activate neuronal genes. Surprisingly, we found that Sox11 also binds genes expressed in the progenitor state that normally are down-regulated at the time when Sox11 expression is initiated. This raises the question of two possible roles for SoxC proteins, one in activating the genes that are silent and bound by Sox3 in NPCs and another role on the enhancers of progenitor genes that are silent in Sox11-expressing differentiating neurons.

The activating role of Sox11 was examined by cloning Sox3-bound and Sox11-bound enhancers corresponding to the neuronal genes Tubb3 and Lhx2 into a luciferase reporter construct. From the reporter assays we could conclude that Sox11 was able to activate both Tubb3 and Lhx2 enhancers. However, Sox3 had an inhibitory effect of the Sox11 mediated trans-activation. This effect was confirmed in chick spinal cord where Sox3 blocked the Sox11 induced premature expression of the Tubb3 gene. Thus, while Sox11 activates neuronal gene expression, Sox3 may act to prevent premature induction of neuronal genes.

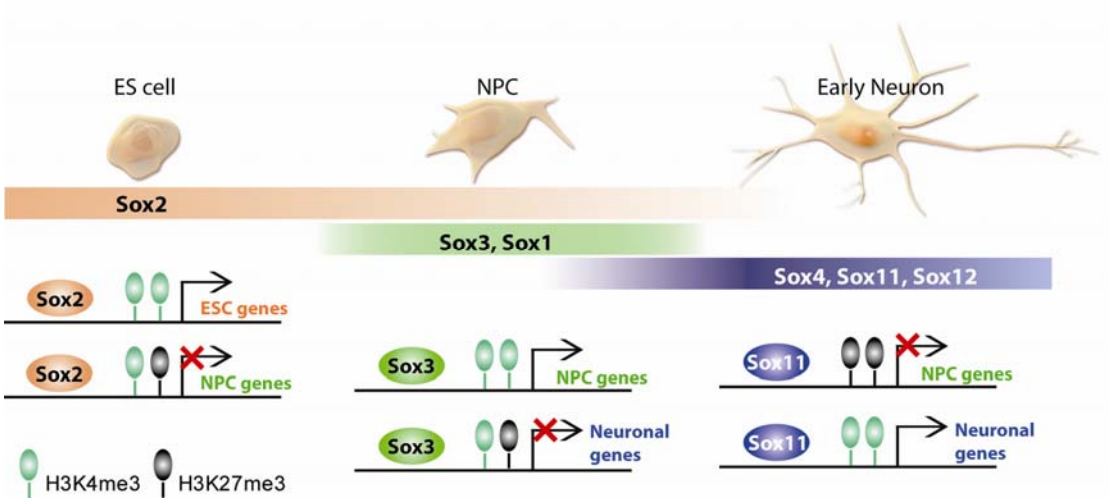
Approximately 30% of the Sox3-bound genes in NPCs are not bound by Sox11 in neurons. These genes were found to be expressed preferably by neurons and glial cells of the adult CNS. It is possible that Sox transcription factors, other than SoxC proteins, occupy these enhancers in neurons and glial cells of the adult CNS. For instance Sox10 may bind these sites in glial cells. According to our data, the oligodendrocyte gene *Mbp* is not a target gene for Sox11 but it has been shown that Sox10 is crucial for activation of *Mbp* at terminal steps of gliogenesis (Li *et al.*, 2007), which could suggest that Sox10 is replacing SoxB1 protein binding at this enhancer. It is also possible that, at a later cell stage than what was used for the Sox11 ChIP-seq analysis, Sox11 would have a different repertoire of target sites, targeting more genes that are expressed at adult CNS stages. Furthermore, a common Sox-Tead binding motif was found among these enhancers, suggesting a combinatorial role of Sox and transcription factors of the Tead family (Kaneko and DePamphilis, 1998) in the regulation of these genes.

EPIGENETIC CHANGES FOLLOW SOX BINDING DURING NEUROGENESIS

Our findings indicate a similar situation for Sox transcription factors during different developmental stages of neurogenesis, as Sox2 in ES cells, Sox3 in NPCs and Sox11 in post-mitotic early neurons occupy both active and silent genes. Interestingly, many Sox2-bound genes that are silent in ES cells are associated with a bivalent chromatin signature (containing both H3K4me3 and H3K27me3 modifications), which is resolved upon differentiation as the genes become activated in NPCs (Bernstein *et al.*, 2006). These findings compelled us to characterize the chromatin modifications of Sox3-bound genes that are either active or silent in NPCs. Sequential preformed ChIP experiments revealed that Sox3-bound active genes in NPCs were associated with H3K4me3 only, which is a mark for actively transcribed genes. However, the Sox3 occupied genes that are silent in NPCs were bivalent marked with both H3K4me3 and H3K27me3, where H3K27me is often associated with silent genes. Interestingly, sequential ChIP experiments for Sox11 revealed that upon Sox11 binding in differentiating neurons, these bivalent histone characters had been resolved and these actively transcribed genes were now associated with H3K4me3 only. In contrast, the

binding of Sox11 to promoters of NPC expressed genes that are silent in neurons was associated with a replacement of H3K4me3 with H3K27me3.

To summarize, Sox3 works in an activating fashion on the previously Sox2-bound progenitor genes whereas Sox3-bound neuronal and glial genes are transcriptional silent and kept in a bivalent state. Later in differentiation, Sox11 is bound to the progenitor genes which are now silent. The differentiation genes which were previously bound by Sox3, silent and bivalent are now bound by Sox11 and are in an active state. These data indicate that Sox transcription factors can act to coordinate neural gene expression as development proceeds, from the early neural lineage specification in pluripotent stem cells to the gene regulatory control operating in maturing post-mitotic neurons (Figure 5).



Figur 5. Sox transcription factors bind both active and silent genes in their different cellular stages. The switch of Sox proteins at the different enhancers is associated with a change in transcriptional activity as well as in histone modifications.

Since sequential ChIP experiments were preformed for the chromatin modification study presented in paper II, we could exclude signals from contaminating cell types or cell stages not expressing the transcription factor of interest. Important to note is that it is not possible to define the origin of the signal after ChIP from a mixed population of cells. For example, a signal for H3K4me3 in a population that consists of 50% neurons and 50% neural progenitor cells could be derived from either cell stages or from both.

CONCLUDING REMARKS

We have with paper I and II contributed to the knowledge concerning vertebrate neurogenesis and more specifically to the role of SoxB1 proteins in neural progenitor cells and to that of SoxC proteins in differentiating neuronal cells. We suggest a sequential regulatory role for Sox transcription factors during the course of neurogenesis, where specific members of the Sox protein family both activate genes as well as prepare other genes for the following differentiation steps. The alternation in Sox protein binding could also be correlated with a change in chromatin modifications, where bivalent chromatin marks of silent genes were resolved as the genes were bound by activating Sox proteins.

Sox transcription factors have not previously been shown to have a role in recruiting chromatin modifying proteins. However, it is possible that Sox proteins work as 'pioneering factors' that act as the primary transcription factor that binds to DNA and thereby induces a series of events which facilitate the binding of later acting transcriptional regulators. For instance, Sox proteins induce a bend of DNA when binding the minor groove, which in turn gives space for other transcription factors to bind and catalyzes histone modifications. This could support the possibility for Sox proteins to function as 'pioneering factors'.

In paper II we present that bivalent, silent and Sox2-bound genes in ES cells are preferably genes that are later expressed in neural tissue compared to other organs. This could imply that Sox2 specifies the ectodermal lineage in ES cells and that other transcription factors are required to specify and prepare for differentiation toward the other two lineages. Similar to the situation of Sox proteins during neuronal lineage differentiation, it has been described that members of the forkhead family of proteins are important during endodermal development. The forkhead family member FoxD3 is essential for unmethylated DNA mark observed at the *Alb1* (liver-specific) enhancer (Gualdi *et al.*, 1996) in ES cells, where *Alb1* is silent. During endodermal differentiation, another member of this family, FoxA1, replaces FoxD3 at the *Alb1* enhancer in order to keep the gene poised for expression in a more mature cell stage (Zaret *et al.*, 2008; Xu *et al.*, 2009). Furthermore, FoxD3 is the only forkhead family member expressed in ES cells and has been shown to be necessary for ES cell maintenance (Hanna *et al.*, 2002; Pan *et al.*, 2006).

In a recent study, 38 different types of histone modifications were investigated to be correlated with binding of the NRSF/REST repressor in the human genome (Zheng *et al.*, 2009). It was found that several of the histone modifications were correlated to NRSF/REST binding and the publication provides information regarding the most frequent histone marks within NRSF/REST binding sites. This information regarding combinatorial codes of histone modifications could be correlated with the DNA binding of other transcription factors, such as factors from the Sox protein family. The binding and different modifications could also be correlated to the activity of the certain transcription factor. Hence, to investigate local chromatin modifications beyond H3K4me3 and H3K27me3 could provide more information regarding the function of Sox proteins.

ACKNOWLEDGEMENTS

De senaste sju åren har varit en speciell tid i mitt liv. Jag har mött människor som delar mitt intresse och som jag kommer att sakna oerhört mycket. Speciellt vill jag tacka följande personer.

Dr. Jonas Muhr. Du introducerade forskningsvärlden för mig med entusiasm och ett imponerande sätt att tänka. Tack för att du har kämpat med mig och för mig. Och för att du fick mig att känna att det alltid fanns tid för frågor.

Mina fina vänner som gått igenom doktorandtiden med mig. **Banafsheh** för din lojalitet. **Stina** för din energi. **Nick** för din humor. Men framförallt, tack för sällskapet under den här resan. Vardagar, semestrar, konferenser, bra dagar och dåliga dagar. Och alla fester.

Professor Thomas Perlmann. Tack för ditt intresse. Och tack för det initiala antagandet till LICR!

Mats, **Charlotta**, **Birgitta** och **Erika**. Världens bästa markservice!

Inger Tollman för din excellence.

Magnus och **Magda**. De första välkomnande Muhrarna.

Min vän **Idha**.

Johan H. Tack för din vänskap och ditt stöd. Det har under vissa stunder gjort avgörande skillnad.

Muhr/Perlmann lab, past and present members. Thanks for creating an amazing atmosphere at LICR

Ulf Eriksson lab.

Cécile. Thank you for this last year and enjoy your PhD.

Alexandra, **Eva** och **Stina**. För att ni finns i mitt liv. För alla skratt som påminner mig om att det även finns annat i livet än forskning.

Martina. Jag är så glad att vi möttes den där söndagen i Uppsala.

Kusin **Linda**. Tack för ditt annorlunda synsätt. Tack för att du alltid har varit intresserad. Du är fantastisk.

De vikigaste av alla:
Underbara **familjen Bergsland**. Mitt skyddsnät. Mamma och Pappa. Det finns inte ord. Ni gav mig en känsla av att ingenting är omöjligt. På riktigt.

Mina allierade. Mina systrar. Tack för allt stöd. Där den ena inte räcker till tar den andre vid. **Rebecca** för ditt eviga och orubbade support, när som helst på dygnet, var som helst i världen. **Lina** för att du ger mig trygghet och stabilitet när jag behöver det.

Jonny. Tack för backup och din eviga förståelse. Det finns inte många som du.

REFERENCES

Ambrosetti, D. C., Basilico, C. &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(11): 6321-6329.

Ambrosetti, D. C., Scholer, H. R., Dailey, L. &Basilico, C. (2000). Modulation of the activity of multiple transcriptional activation domains by the DNA binding domains mediates the synergistic action of Sox2 and Oct-3 on the fibroblast growth factor-4 enhancer. *J Biol Chem* 275(30): 23387-23397.

Artavanis-Tsakonas, S., Rand, M. D. &Lake, R. J. (1999). Notch signaling: cell fate control and signal integration in development. *Science* 284(5415): 770-776.

Avilion, A. A., Nicolis, S. K., Pevny, L. H., Perez, L., Vivian, N. &Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* 17(1): 126-140.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H. F., John, R. M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M. &Fisher, A. G. (2006). Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* 8(5): 532-538.

Bernstein, B. E., Mikkelsen, T. S., Xie, X., Kamal, M., Huebert, D. J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., Jaenisch, R., Wagschal, A., Feil, R., Schreiber, S. L. &Lander, E. S. (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2): 315-326.

Bertrand, N., Castro, D. S. &Guillemot, F. (2002). Proneural genes and the specification of neural cell types. *Nat Rev Neurosci* 3(7): 517-530.

Botto, L. D., Moore, C. A., Khoury, M. J. &Erickson, J. D. (1999). Neural-tube defects. *N Engl J Med* 341(20): 1509-1519.

Bowles, J., Schepers, G. &Koopman, P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev Biol* 227(2): 239-255.

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., Gifford, D. K., Melton, D. A., Jaenisch, R. &Young, R. A. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122(6): 947-956.

Bray, S. J. (2006). Notch signalling: a simple pathway becomes complex. *Nat Rev Mol Cell Biol* 7(9): 678-689.

Briscoe, J., Pierani, A., Jessell, T. M. &Ericson, J. (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. *Cell* 101(4): 435-445.

Bruce, A. W., Donaldson, I. J., Wood, I. C., Yerbury, S. A., Sadowski, M. I., Chapman, M., Gottgens, B. &Buckley, N. J. (2004). Genome-wide analysis of repressor element 1 silencing transcription factor/neuron-restrictive silencing factor (REST/NRSF) target genes. *Proc Natl Acad Sci U S A* 101(28): 10458-10463.

Bylund, M., Andersson, E., Novitch, B. G. &Muhr, J. (2003). Vertebrate neurogenesis is counteracted by Sox1-3 activity. *Nat Neurosci* 6(11): 1162-1168.

Campuzano, S. &Modolell, J. (1992). Patterning of the Drosophila nervous system: the achaete-scute gene complex. *Trends Genet* 8(6): 202-208.

Cao, R. &Zhang, Y. (2004). The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* 14(2): 155-164.

Cavallaro, M., Mariani, J., Lancini, C., Latorre, E., Caccia, R., Gullo, F., Valotta, M., DeBiasi, S., Spinardi, L., Ronchi, A., Wanke, E., Brunelli, S., Favaro, R., Ottolenghi, S. &Nicolis, S. K. (2008). Impaired generation of mature neurons by neural stem cells from hypomorphic Sox2 mutants. *Development* 135(3): 541-557.

Chambers, I., Colby, D., Robertson, M., Nichols, J., Lee, S., Tweedie, S. &Smith, A. (2003). Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell* 113(5): 643-655.

Chambers, I., Silva, J., Colby, D., Nichols, J., Nijmeijer, B., Robertson, M., Vrana, J., Jones, K., Grotewold, L. &Smith, A. (2007). Nanog safeguards pluripotency and mediates germline development. *Nature* 450(7173): 1230-1234.

Chambers, I. &Tomlinson, S. R. (2009). The transcriptional foundation of pluripotency. *Development* 136(14): 2311-2322.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V. B., Wong, E., Orlov, Y. L., Zhang, W., Jiang, J., Loh, Y. H., Yeo, H. C., Yeo, Z. X., Narang, V., Govindarajan, K. R., Leong, B., Shahab, A., Ruan, Y., Bourque, G., Sung, W. K., Clarke, N. D., Wei, C. L. &Ng, H. H. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133(6): 1106-1117.

Cheung, M., Abu-Elmagd, M., Clevers, H. &Scotting, P. J. (2000). Roles of Sox4 in central nervous system development. *Brain Res Mol Brain Res* 79(1-2): 180-191.

Chew, L. J. &Gallo, V. (2009). The Yin and Yang of Sox proteins: Activation and repression in development and disease. *J Neurosci Res* 87(15): 3277-3287.

Chong, J. A., Tapia-Ramirez, J., Kim, S., Toledo-Aral, J. J., Zheng, Y., Boutros, M. C., Altshuller, Y. M., Frohman, M. A., Kraner, S. D. &Mandel, G. (1995). REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. *Cell* 80(6): 949-957.

Cole, M. F., Johnstone, S. E., Newman, J. J., Kagey, M. H. &Young, R. A. (2008). Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. *Genes Dev* 22(6): 746-755.

Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N. &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(2): 509-520.

Connor, F., Cary, P. D., Read, C. M., Preston, N. S., Driscoll, P. C., Denny, P., Crane-Robinson, C. &Ashworth, A. (1994). DNA binding and bending properties of the post-meiotically expressed Sry-related protein Sox-5. *Nucleic Acids Res* 22(16): 3339-3346.

Corbin, J. G., Gaiano, N., Juliano, S. L., Poluch, S., Stancik, E. &Haydar, T. F. (2008). Regulation of neural progenitor cell development in the nervous system. *J Neurochem* 106(6): 2272-2287.

Cremazy, F., Berta, P. &Girard, F. (2000). Sox neuro, a new Drosophila Sox gene expressed in the developing central nervous system. *Mech Dev* 93(1-2): 215-219.

de Napoles, M., Mermoud, J. E., Wakao, R., Tang, Y. A., Endoh, M., Appanah, R., Nesterova, T. B., Silva, J., Otte, A. P., Vidal, M., Koseki, H. &Brockdorff, N. (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Dev Cell* 7(5): 663-676.

Dennis, K., Uittenbogaard, M., Chiaramello, A. &Moody, S. A. (2002). Cloning and characterization of the 5'-flanking region of the rat neuron-specific Class III beta-tubulin gene. *Gene* 294(1-2): 269-277.

Doyle, J. P., Dougherty, J. D., Heiman, M., Schmidt, E. F., Stevens, T. R., Ma, G., Bupp, S., Shrestha, P., Shah, R. D., Doughty, M. L., Gong, S., Greengard, P. &Heintz, N. (2008). Application of a translational profiling approach for the comparative analysis of CNS cell types. *Cell* 135(4): 749-762.

Dy, P., Penzo-Mendez, A., Wang, H., Pedraza, C. E., Macklin, W. B. &Lefebvre, V. (2008). The three SoxC proteins--Sox4, Sox11 and Sox12--exhibit overlapping expression patterns and molecular properties. *Nucleic Acids Res* 36(9): 3101-3117.

Ema, M., Mori, D., Niwa, H., Hasegawa, Y., Yamanaka, Y., Hitoshi, S., Mimura, J., Kawabe, Y., Hosoya, T., Morita, M., Shimosato, D., Uchida, K., Suzuki, N., Yanagisawa, J., Sogawa, K., Rossant, J., Yamamoto, M., Takahashi, S. &Fujii-Kuriyama, Y. (2008). Kruppel-like factor 5 is essential for blastocyst development and the normal self-renewal of mouse ESCs. *Cell Stem Cell* 3(5): 555-567.

Farah, M. H., Olson, J. M., Sucic, H. B., Hume, R. I., Tapscott, S. J. &Turner, D. L. (2000). Generation of neurons by transient expression of neural bHLH proteins in mammalian cells. *Development* 127(4): 693-702.

Ferrari, S., Harley, V. R., Pontiggia, A., Goodfellow, P. N., Lovell-Badge, R. &Bianchi, M. E. (1992). SRY, like HMG1, recognizes sharp angles in DNA. *Embo J* 11(12): 4497-4506.

Ferri, A. L., Cavallaro, M., Braida, D., Di Cristofano, A., Canta, A., Vezzani, A., Ottolenghi, S., Pandolfi, P. P., Sala, M., DeBiasi, S. &Nicolis, S. K. (2004). Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. *Development* 131(15): 3805-3819.

Finzsch, M., Stolt, C. C., Lommes, P. &Wegner, M. (2008). Sox9 and Sox10 influence survival and migration of oligodendrocyte precursors in the spinal cord by regulating PDGF receptor alpha expression. *Development* 135(4): 637-646.

Fouse, S. D., Shen, Y., Pellegrini, M., Cole, S., Meissner, A., Van Neste, L., Jaenisch, R. &Fan, G. (2008). Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell* 2(2): 160-169.

Ghislain, J. &Charnay, P. (2006). Control of myelination in Schwann cells: a Krox20 cis-regulatory element integrates Oct6, Brn2 and Sox10 activities. *EMBO Rep* 7(1): 52-58.

Gilbert, S. F. (2003). Developmental Biology: sinauer associates inc.

Golden, J. A. &Chernoff, G. F. (1993). Intermittent pattern of neural tube closure in two strains of mice. *Teratology* 47(1): 73-80.

Graham, V., Khudyakov, J., Ellis, P. &Pevny, L. (2003). SOX2 functions to maintain neural progenitor identity. *Neuron* 39(5): 749-765.

Gualdi, R., Bossard, P., Zheng, M., Hamada, Y., Coleman, J. R. &Zaret, K. S. (1996). Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. *Genes Dev* 10(13): 1670-1682.

Hanna, L. A., Foreman, R. K., Tarasenko, I. A., Kessler, D. S. &Labosky, P. A. (2002). Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. *Genes Dev* 16(20): 2650-2661.

Hartl, D., Irmler, M., Romer, I., Mader, M. T., Mao, L., Zabel, C., de Angelis, M. H., Beckers, J. &Klose, J. (2008). Transcriptome and proteome analysis of early embryonic mouse brain development. *Proteomics* 8(6): 1257-1265.

Hatano, S. Y., Tada, M., Kimura, H., Yamaguchi, S., Kono, T., Nakano, T., Suemori, H., Nakatsuji, N. &Tada, T. (2005). Pluripotential competence of cells associated with Nanog activity. *Mech Dev* 122(1): 67-79.

Hattori, N. &Shiota, K. (2008). Epigenetics: the study of embryonic stem cells by restriction landmark genomic scanning. *FEBS J* 275(8): 1624-1630.

He, Y., Dupree, J., Wang, J., Sandoval, J., Li, J., Liu, H., Shi, Y., Nave, K. A. &Casaccia-Bonnefil, P. (2007). The transcription factor Yin Yang 1 is essential for oligodendrocyte progenitor differentiation. *Neuron* 55(2): 217-230.

Hitoshi, S., Alexson, T., Tropepe, V., Donoviel, D., Elia, A. J., Nye, J. S., Conlon, R. A., Mak, T. W., Bernstein, A. &van der Kooy, D. (2002). Notch pathway molecules are essential for the maintenance, but not the generation, of mammalian neural stem cells. *Genes Dev* 16(7): 846-858.

Hollyday, M. (2001). Neurogenesis in the vertebrate neural tube. *Int J Dev Neurosci* 19(2): 161-173.

Holmberg, J., Hansson, E., Malewicz, M., Sandberg, M., Perlmann, T., Lendahl, U. &Muhr, J. (2008). SoxB1 transcription factors and Notch signaling use distinct mechanisms to regulate proneural gene function and neural progenitor differentiation. *Development* 135(10): 1843-1851.

Hoser, M., Potzner, M. R., Koch, J. M., Bosl, M. R., Wegner, M. &Sock, E. (2008). Sox12 deletion in the mouse reveals nonreciprocal redundancy with the related Sox4 and Sox11 transcription factors. *Mol Cell Biol* 28(15): 4675-4687.

Hunt, P. &Krumlauf, R. (1991). Deciphering the Hox code: clues to patterning branchial regions of the head. *Cell* 66(6): 1075-1078.

Ivanova, N., Dobrin, R., Lu, R., Kotenko, I., Levorse, J., DeCoste, C., Schafer, X., Lun, Y. &Lemischka, I. R. (2006). Dissecting self-renewal in stem cells with RNA interference. *Nature* 442(7102): 533-538.

Jessell, T. M. (2000). Neuronal specification in the spinal cord: inductive signals and transcriptional codes. *Nat Rev Genet* 1(1): 20-29.

Kamachi, Y., Uchikawa, M., Collignon, J., Lovell-Badge, R. &Kondoh, H. (1998). Involvement of Sox1, 2 and 3 in the early and subsequent molecular events of lens induction. *Development* 125(13): 2521-2532.

Kamachi, Y., Uchikawa, M. &Kondoh, H. (2000). Pairing SOX off: with partners in the regulation of embryonic development. *Trends Genet* 16(4): 182-187.

Kamachi, Y., Uchikawa, M., Tanouchi, A., Sekido, R. &Kondoh, H. (2001). Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. *Genes Dev* 15(10): 1272-1286.

Kaneko, K. J. &DePamphilis, M. L. (1998). Regulation of gene expression at the beginning of mammalian development and the TEAD family of transcription factors. *Dev Genet* 22(1): 43-55.

Kappen, C., Schughart, K. &Ruddle, F. H. (1993). Early evolutionary origin of major homeodomain sequence classes. *Genomics* 18(1): 54-70.

Kim, J., Chu, J., Shen, X., Wang, J. &Orkin, S. H. (2008). An extended transcriptional network for pluripotency of embryonic stem cells. *Cell* 132(6): 1049-1061.

Kondoh, H., Uchikawa, M. &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(8-9): 819-827.

Kopp, J. L., Ormsbee, B. D., Desler, M. &Rizzino, A. (2008). Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells* 26(4): 903-911.

Lee, T. I., Jenner, R. G., Boyer, L. A., Guenther, M. G., Levine, S. S., Kumar, R. M., Chevalier, B., Johnstone, S. E., Cole, M. F., Isono, K., Koseki, H., Fuchikami, T., Abe, K., Murray, H. L., Zucker, J. P., Yuan, B., Bell, G. W., Herbolsheimer, E., Hannett, N. M., Sun, K., Odom, D. T., Otte, A. P., Volkert, T. L., Bartel, D. P., Melton, D. A., Gifford, D. K., Jaenisch, R. &Young, R. A. (2006). Control of developmental regulators by Polycomb in human embryonic stem cells. *Cell* 125(2): 301-313.

Lefebvre, V., Dumitriu, B., Penzo-Mendez, A., Han, Y. &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(12): 2195-2214.

Li, H., Lu, Y., Smith, H. K. &Richardson, W. D. (2007). Olig1 and Sox10 interact synergistically to drive myelin basic protein transcription in oligodendrocytes. *J Neurosci* 27(52): 14375-14382.

Liang, J., Wan, M., Zhang, Y., Gu, P., Xin, H., Jung, S. Y., Qin, J., Wong, J., Cooney, A. J., Liu, D. &Songyang, Z. (2008). Nanog and Oct4 associate with unique transcriptional repression complexes in embryonic stem cells. *Nat Cell Biol* 10(6): 731-739.

Lo, L., Dormand, E., Greenwood, A. &Anderson, D. J. (2002). Comparison of the generic neuronal differentiation and neuron subtype specification functions of mammalian achaete-scute and atonal homologs in cultured neural progenitor cells. *Development* 129(7): 1553-1567.

Lothian, C. &Lendahl, U. (1997). An evolutionarily conserved region in the second intron of the human nestin gene directs gene expression to CNS progenitor cells and to early neural crest cells. *Eur J Neurosci* 9(3): 452-462.

Louvi, A. &Artavanis-Tsakonas, S. (2006). Notch signalling in vertebrate neural development. *Nat Rev Neurosci* 7(2): 93-102.

Lupo, G., Harris, W. A. &Lewis, K. E. (2006). Mechanisms of ventral patterning in the vertebrate nervous system. *Nat Rev Neurosci* 7(2): 103-114.

Malas, S., Postlethwaite, M., Ekonomou, A., Whalley, B., Nishiguchi, S., Wood, H., Meldrum, B., Constanti, A. &Episkopou, V. (2003). Sox1-deficient mice suffer from epilepsy associated with abnormal ventral forebrain development and olfactory cortex hyperexcitability. *Neuroscience* 119(2): 421-432.

Massari, M. E. &Murre, C. (2000). Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol Cell Biol* 20(2): 429-440.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A. A., Ko, M. S. &Niwa, H. (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* 9(6): 625-635.

Miyagi, S., Kato, H. &Okuda, A. (2009). Role of SoxB1 transcription factors in development. *Cell Mol Life Sci* 66(23): 3675-3684.

Miyagi, S., Saito, T., Mizutani, K., Masuyama, N., Gotoh, Y., Iwama, A., Nakauchi, H., Masui, S., Niwa, H., Nishimoto, M., Muramatsu, M. &Okuda, A. (2004). The Sox-2 regulatory regions display their activities in two distinct types of multipotent stem cells. *Mol Cell Biol* 24(10): 4207-4220.

Nakada, Y., Hunsaker, T. L., Henke, R. M. &Johnson, J. E. (2004). Distinct domains within Mash1 and Math1 are required for function in neuronal differentiation versus neuronal cell-type specification. *Development* 131(6): 1319-1330.

Nakatsu, T., Uwabe, C. &Shiota, K. (2000). Neural tube closure in humans initiates at multiple sites: evidence from human embryos and implications for the pathogenesis of neural tube defects. *Anat Embryol (Berl)* 201(6): 455-466.

Nishiguchi, S., Wood, H., Kondoh, H., Lovell-Badge, R. &Episkopou, V. (1998). Sox1 directly regulates the gamma-crystallin genes and is essential for lens development in mice. *Genes Dev* 12(6): 776-781.

Nishimoto, M., Fukushima, A., Okuda, A. &Muramatsu, M. (1999). The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol Cell Biol* 19(8): 5453-5465.

Niwa, H., Miyazaki, J. &Smith, A. G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24(4): 372-376.

Novitch, B. G., Chen, A. I. &Jessell, T. M. (2001). Coordinate regulation of motor neuron subtype identity and pan-neuronal properties by the bHLH repressor Olig2. *Neuron* 31(5): 773-789.

Pachikara, A., Dolson, D. K., Martinu, L., Riccomagno, M. M., Jeong, Y. &Epstein, D. J. (2007). Activation of Class I transcription factors by low level Sonic hedgehog signaling is mediated by Gli2-dependent and independent mechanisms. *Dev Biol* 305(1): 52-62.

Pan, G., Li, J., Zhou, Y., Zheng, H. &Pei, D. (2006). A negative feedback loop of transcription factors that controls stem cell pluripotency and self-renewal. *FASEB J* 20(10): 1730-1732.

Pevny, L. &Placzek, M. (2005). SOX genes and neural progenitor identity. *Curr Opin Neurobiol* 15(1): 7-13.

Phillips, K. &Luisi, B. (2000). The virtuoso of versatility: POU proteins that flex to fit. *J Mol Biol* 302(5): 1023-1039.

Potzner, M. R., Griffel, C., Lutjen-Drecoll, E., Bosl, M. R., Wegner, M. &Sock, E. (2007). Prolonged Sox4 expression in oligodendrocytes interferes with normal myelination in the central nervous system. *Mol Cell Biol* 27(15): 5316-5326.

Rex, M., Orme, A., Uwanogho, D., Tointon, K., Wigmore, P. M., Sharpe, P. T. &Scotting, P. J. (1997). Dynamic expression of chicken Sox2 and Sox3 genes in ectoderm induced to form neural tissue. *Dev Dyn* 209(3): 323-332.

Rizzoti, K., Brunelli, S., Carmignac, D., Thomas, P. Q., Robinson, I. C. &Lovell-Badge, R. (2004). SOX3 is required during the formation of the hypothalamo-pituitary axis. *Nat Genet* 36(3): 247-255.

Sandberg, M., Kallstrom, M. &Muhr, J. (2005). Sox21 promotes the progression of vertebrate neurogenesis. *Nat Neurosci* 8(8): 995-1001.

Schoenherr, C. J. &Anderson, D. J. (1995). The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. *Science* 267(5202): 1360-1363.

Schreiner, S., Cossais, F., Fischer, K., Scholz, S., Bosl, M. R., Holtmann, B., Sendtner, M. &Wegner, M. (2007). Hypomorphic Sox10 alleles reveal novel protein functions and unravel developmental differences in glial lineages. *Development* 134(18): 3271-3281.

Schuettengruber, B., Chourrout, D., Vervoort, M., Leblanc, B. &Cavalli, G. (2007). Genome regulation by polycomb and trithorax proteins. *Cell* 128(4): 735-745.

Shen, X., Liu, Y., Hsu, Y. J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G. C. &Orkin, S. H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. *Mol Cell* 32(4): 491-502.

Smith, J. L. &Schoenwolf, G. C. (1997). Neurulation: coming to closure. *Trends Neurosci* 20(11): 510-517.

Sock, E., Rettig, S. D., Enderich, J., Bosl, M. R., Tamm, E. R. &Wegner, M. (2004). Gene targeting reveals a widespread role for the high-mobility-group transcription factor Sox11 in tissue remodeling. *Mol Cell Biol* 24(15): 6635-6644.

Stern, C. D. (2005). Neural induction: old problem, new findings, yet more questions. *Development* 132(9): 2007-2021.

Stolt, C. C., Lommes, P., Sock, E., Chaboissier, M. C., Schedl, A. &Wegner, M. (2003). The Sox9 transcription factor determines glial fate choice in the developing spinal cord. *Genes Dev* 17(13): 1677-1689.

Strahl, B. D. &Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403(6765): 41-45.

Sun, Y., Nadal-Vicens, M., Misono, S., Lin, M. Z., Zubiaga, A., Hua, X., Fan, G. &Greenberg, M. E. (2001). Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms. *Cell* 104(3): 365-376.

Taranova, O. V., Magness, S. T., Fagan, B. M., Wu, Y., Surzenko, N., Hutton, S. R. &Pevny, L. H. (2006). SOX2 is a dose-dependent regulator of retinal neural progenitor competence. *Genes Dev* 20(9): 1187-1202.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S. &Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. *Science* 282(5391): 1145-1147.

Uchikawa, M., Kamachi, Y. &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(1-2): 103-120.

Waddington, C. H. &Waterman, A. J. (1933). The Development in vitro of Young Rabbit Embryos. *J Anat* 67(Pt 3): 355-370.

van Beest, M., Dooijes, D., van De Wetering, M., Kjaerulff, S., Bonvin, A., Nielsen, O. &Clevers, H. (2000). Sequence-specific high mobility group box factors recognize 10-12-base pair minor groove motifs. *J Biol Chem* 275(35): 27266-27273.

Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D. N., Theunissen, T. W. &Orkin, S. H. (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444(7117): 364-368.

Wegner, M. (1999). From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res* 27(6): 1409-1420.

Wegner, M. &Stolt, C. C. (2005). From stem cells to neurons and glia: a Soxist's view of neural development. *Trends Neurosci* 28(11): 583-588.

Werner, M. H. &Burley, S. K. (1997). Architectural transcription factors: proteins that remodel DNA. *Cell* 88(6): 733-736.

Wiebe, M. S., Nowling, T. K. &Rizzino, A. (2003). Identification of novel domains within Sox-2 and Sox-11 involved in autoinhibition of DNA binding and partnership specificity. *J Biol Chem* 278(20): 17901-17911.

Wilson, P. A., Lagna, G., Suzuki, A. &Hemmati-Brivanlou, A. (1997). Concentration-dependent patterning of the Xenopus ectoderm by BMP4 and its signal transducer Smad1. *Development* 124(16): 3177-3184.

Visel, A., Blow, M. J., Li, Z., Zhang, T., Akiyama, J. A., Holt, A., Plajzer-Frick, I., Shoukry, M., Wright, C., Chen, F., Afzal, V., Ren, B., Rubin, E. M. &Pennacchio, L. A. (2009). ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457(7231): 854-858.

Wolffe, A. P. (1994). Architectural transcription factors. *Science* 264(5162): 1100-1101.

Wood, H. B. &Episkopou, V. (1999). Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages. *Mech Dev* 86(1-2): 197-201.

Xu, J., Watts, J. A., Pope, S. D., Gadue, P., Kamps, M., Plath, K., Zaret, K. S. &Smale, S. T. (2009). Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. *Genes Dev* 23(24): 2824-2838.

Yoon, K. &Gaiano, N. (2005). Notch signaling in the mammalian central nervous system: insights from mouse mutants. *Nat Neurosci* 8(6): 709-715.

Yuan, H., Corbi, N., Basilico, C. &Dailey, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev* 9(21): 2635-2645.

Zappone, M. V., Galli, R., Catena, R., Meani, N., De Biasi, S., Mattei, E., Tiveron, C., Vescovi, A. L., Lovell-Badge, R., Ottolenghi, S. &Nicolis, S. K. (2000). Sox2 regulatory sequences direct expression of a (beta)-geo transgene to telencephalic neural stem cells and precursors of the mouse embryo, revealing regionalization of gene expression in CNS stem cells. *Development* 127(11): 2367-2382.

Zaret, K. S., Watts, J., Xu, J., Wandzioch, E., Smale, S. T. &Sekiya, T. (2008). Pioneer factors, genetic competence, and inductive signaling: programming liver and pancreas progenitors from the endoderm. *Cold Spring Harb Symp Quant Biol* 73: 119-126.

Zhao, S., Nichols, J., Smith, A. G. &Li, M. (2004). SoxB transcription factors specify neuroectodermal lineage choice in ES cells. *Mol Cell Neurosci* 27(3): 332-342.

Zheng, D., Zhao, K. &Mehler, M. F. (2009). Profiling RE1/REST-mediated histone modifications in the human genome. *Genome Biol* 10(1): R9.