From DEPARTMENT OF CELL AND MOLECULAR BIOLOGY
Karolinska Institutet, Stockholm, Sweden

LET'S TALK ABOUT SOX:

DIVERSE KEY PLAYERS IN THE TRANSCRIPTIONAL REGULATION OF NEURAL STEM CELLS

Susanne Klum

Karolinska Institutet

Stockholm 2018

On the cover: “Chromatin” by Jeff Prymowicz, The Jackson Laboratory

Published by Karolinska Institutet.
Printed by E-Print AB 2018

ISBN 978-91-7831-110-1

Let's talk about SOX: diverse key players in the transcriptional regulation of neural stem cells

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Susanne Klum

June 15th 2018, 9:30am
at "Hörsal Vesalius", Berzelius Väg 3

Principal Supervisor:
Prof. Jonas Muhr
Karolinska Institutet
Department of Cell and Molecular Biology

Co-supervisor(s):
Assoc. Prof. Johan Holmberg
Karolinska Institutet
Department of Cell and Molecular Biology

Opponent:
Dr. rer. nat. Ralf Jauch
University of Hong Kong
Li Ka Shing Faculty of Medicine
School of Biomedical Science

Examination Board:
Prof. Mikael Sigvardsson
Lunds Universitet
Division of Molecular Hematology

Docent Margareta Wilhelm
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Docent Elisabet Andersson
Karolinska Institutet
Department of Cell and Molecular Biology

"Biology gives you a brain. Life turns it into a mind."

- Jeffrey Eugenides

ABSTRACT

The development of the central nervous system relies on neural stem cells. How these stem cells either maintain their identity as self-renewing progenitors or differentiate into cells of the neuronal or glial lineage are fundamental questions. The work presented in this thesis investigates how different SOX transcription factors orchestrate the mechanisms and gene expression programs that govern these processes.

In **paper I**, we investigate how SOX2, SOX3 and SOX11 regulate specific gene sets in embryonic stem cells (ESCs), neural progenitor cells (NPCs) and neurons, respectively. We propose a model of sequentially acting SOX transcription factors that control neural lineage-specific gene expression by predisposing gene sets to become activated in NPCs and during neuronal differentiation.

In **paper II**, we expand our sequential binding model to cells of the glial lineage. Binding studies in NPCs and glial precursor cells (GPCs) show that astrocyte and oligodendrocyte specific gene sets are extensively preselected prior to the onset of gliogenesis, through prebinding of SOX3 and SOX9. This prebinding serves to prevent premature activation of a subset of genes, but also promotes the formation of permissive chromatin, which facilitates their activation at a later stage during astrocytic and oligodendrocytic differentiation.

In **paper III**, we use ChIP-seq and RNA-seq analysis to investigate the binding profile of SOX2 in neural stem cells of four different tissues from two germlayers. We demonstrate that although SOX2 binds few regions that are common, the majority of its target regions are cell type specific. The target sites are enriched for distinct binding motifs of putative co-factors that are either commonly expressed or cell type specific. Furthermore, the bound regions function as cis-regulatory modules (CRMs) that instruct tissue specific gene expression.

In **paper IV**, we examine the initial binding events of SOX2 and FOXG1 during early reprogramming. When misexpressed alone in mouse embryonic fibroblasts (MEFs), FOXG1 preferably targets chromatin regions that were already open in native MEFs. In contrast, SOX2 can target regions that were previously closed and inaccessible, hence displaying pioneering activity. When co-expressed, both factors potentiate each other's binding to neural genes, while the regions that are co-targeted by SOX2 and FOXG1 together correspond to genes that are mainly associated with proliferation.

LIST OF SCIENTIFIC PAPERS

I. Maria Bergsland*, Daniel Ramsköld*, Cécile Zaouter, **Susanne Klum**, Rickard Sandberg, Jonas Muhr.
Sequentially acting SOX transcription factors in neural lineage development.
Genes & Development 2011, 25(23):2453-64.

II. **Susanne Klum***, Cécile Zaouter*, Zhanna Alekseenko, Åsa K. Björklund, Daniel W. Hagey, Johan Ericson, Jonas Muhr, Maria Bergsland.
Sequentially acting SOX proteins orchestrate astrocyte and oligodendrocyte specific gene expression.
Manuscript.

III. Daniel W. Hagey, **Susanne Klum***, Idha Kurtsdotter*, Cécile Zaouter*, Danijal Topcic, Olov Andersson, Maria Bergsland, Jonas Muhr.
SOX2 regulates common and specific stem cell features in the CNS and endoderm derived organs.
PLoS Genetic 2018,14(2):e1007224.

IV. **Susanne Klum**, Danijal Topcic, Maria Bergsland, Jonas Muhr:
Reciprocal effects of SOX2 and FOXG1 on their binding specificities to neural genes in mouse fibroblasts.
Manuscript.

*These authors contributed equally to this work.

PUBLICATIONS NOT INCLUDED IN THIS THESIS

Anna Cascante, **Susanne Klum**, Moumita Biswas, Beatriz Antolin-Fontes, Fanie Barnabé-Heider, Ola Hermanson:
Gene-specific methylation control of H3K9 and H3K36 on neurotrophic BDNF versus astroglial GFAP genes by KDM4A/C regulates neural stem cell differentiation.
Journal of Molecular Biology 2014, 426(20):3467-77.

CONTENTS

LIST OF ABBREVIATIONS

bHLH	basic helix-loop-helix
ChIP	chromatin immunoprecipitation
CNS	central nervous system
CRM	cis-regulatory module
DNA	desoxyribonucleic acid
ESC	embryonic stem cell
GPC	glial precursor cell
HMG	high mobility group
ICM	inner cell mass
iPSC	induced pluripotent stem cell
NPC	neural precursor cell
OPC	oligodendrocyte precursor cell
SOX	SRY-related HMG-box
TSS	transcriptional start site

1 INTRODUCTION

"The chief function of the body is to carry the brain around." This statement by Thomas A. Edison illustrates the significance of our most complex organ. Together with the spinal cord it makes up the central nervous system (CNS), which is the central control module of our body and mind. It coordinates our entire activity, from physical functions like motion, reflexes and hormone secretion to the processing of sensation and emotions, like the creation of memories, thoughts and feelings. The building blocks of the CNS are billions of nerve and glial cells. The human brain alone consists of close to 100 billion neurons and 1000 billion glial cells, which are mostly astrocytes and oligodendrocytes. These cells originate from a common ancestor, the neural stem cell. Neural stem cells are self-renewing cells, which means that they divide and give rise to daughter cells while maintaining their undifferentiated state. At the same time they are multipotent, which means they have the capacity to develop into different cell types of the neural lineage, such as neurons, astrocytes or oligodendrocytes. Neural stem cells are mainly found in the embryonic nervous tissue, but some persist also in the adult CNS, where they provide a life-long source of neurons and glia.

The two characteristic properties of neural stem cells, infinite self-renewal capacity and lineage specific differentiation, are tightly controlled by a network of various transcription factors. SOX transcription factors are essential regulators of both features. The differentiation of neural stem cells into neurons and glial cells is a multi-step process that involves lineage specification, the downregulation of proliferation markers, cell cycle exit, migration and the upregulation of neuronal or glial markers while repressing genes of alternative lineages. The involvement and requirement of different SOX proteins in each of these steps underlines their status as key regulators of CNS development and maintenance.

2 THE SOX FAMILY

The founding member of the SOX family was first identified in 1990 as the sex determining gene on the Y-chromosome, *Sry* (Gubbay et al. 1990, Sinclair et al. 1990). As *Sry* is crucial for the male sex determination, mutations in *Sry* gene are associated with sex reversal in mice and humans. Since then, 20 different *Sox* genes have been identified in mice and humans (Schepers et al. 2002), which have been shown to play key roles in a variety of developmental processes. They are expressed during all phases of embryonic development and are well-established regulators of cell fate decisions (Kamachi & Kondoh 2013).

2.1 PROTEIN STRUCTURE

All SOX members are characterized by a 79 amino acid motif that encodes a high-mobility-group (HMG) box, which allows them to bind chromatin in a sequence specific manner and thereby to function as transcription factors. This DNA-binding domain is highly conserved between the different SOX proteins as all SOX members share more than 50% amino acid identity with the SRY HMG box (Schepers et al. 2002). The 20 different mammalian SOX proteins are divided into 8 groups A-H based on their homology (Bowles et al. 2000). Members of the same group share more than 80 % sequence identity within the HMG domain, whereas sequence conservation among SOX proteins from different groups is poor, particularly outside the HMG box. Quite often, members within a group have similar biochemical properties, expression patterns and overlapping functions, while SOX factors from different groups display distinct biological properties. But despite this apparent redundancy, individual members show a selective binding of target genes and exert distinct, cell type specific functions (Bergsland et al. 2011 (paper I), Hagey et al. 2018 (paper III)).

2.2 DOMAINS OUTSIDE THE HMG BOX

Apart from the HMG box, some SOX transcription factors carry other structural elements, such as activation/repression domains and dimerization domains (Table 1). Transactivation domains have been identified in SOXB1, SOXC, SOXE and SOXF proteins (Lefebvre et al. 2007, Bowles et al. 2000). While SOXB1 and SOXC members have only one transactivation domain at the C-terminal end, SOXE and SOXF proteins have two, one at the C-terminal end and one in the central part. In contrast, a repressor domain can be found in SOXB2 group transcription factors (Uchikawa et al. 1999, Sandberg et al. 2005).

Members of the SOXD and SOXE group are the best studied examples of dimerization among SOX proteins. They dimerize via different mechanisms: SOXD proteins have coiled-

coiled domain, which allows them to dimerize in a DNA-independent fashion (Lefebvre 2010). SOXE proteins are monomers in solution, but they have the ability to dimerize in a DNA-dependent fashion by cooperatively binding to adjacent sites. The spacing between the sites is variable, but 3-6 bp are preferred (Peirano et al. 2000, Bridgewater et al. 2003). They form either homodimers or heterodimers with each other, but not with non-SOXE members. Dimerization is achieved via a unique dimerization domain that precedes the HMG box (Schlierf et al. 2002).

Surprisingly, a study by Xia et al. (2015) in neutrophils revealed that the SOXB1 member SOX2 can also form homodimers.

Table 1. Classification of the SOX family in mammals

Group	Members	Regulatory domain	Other domains
SOXA	SRY	Activation	
SOXB1	SOX1, SOX2, SOX3	Activation	
SOXB2	SOX14, SOX21	Repression	
SOXC	SOX4, SOX11, SOX12	Activation	
SOXD	SOX5, SOX6, SOX13	----------	Coiled-coiled
SOXE	SOX8, SOX9, SOX10	Activation	Dimerization
SOXF	SOX7, SOX17, SOX18	Activation	
SOXG	SOX15	----------	
SOXH	SOX30	----------	

2.3 THE MOLECULAR BASIS FOR DNA RECOGNITION BY SOX FACTORS

In contrast to most other transcription factors, SOX proteins bind to the minor groove of the DNA. The binding occurs via the two L-shaped arms of the HMG domain, thereby inducing a bend of the DNA of 60-70°. The minor groove is widened, while the major groove is compressed, which facilitates helical unwinding (Werner et al. 1995). By altering the conformation of the DNA, SOX factors increase its protein accessibility and plasticity; designating them as architectural modifiers of the 3D chromatin structure. Interestingly, the degree of DNA bending was reported to be critical for the transcriptional output. Mutation analysis of the SOX2-HMG domain or the *Fgf4*-enhancer, a classic SOX2 target sequence, demonstrated that alterations of the typical SOX-induced DNA bending angle can change the potency of transcriptional activity and even dictate whether SOX2 functions as a transcriptional activator or repressor (Scaffidi & Bianchi, 2001). However, analysis of DNA elements bound to SOX2, SOX4, SOX9, SOX17 and SOX18 reported highly similar bending angles of the different SOX factors (Klaus et al. 2016). Thus, the selective binding of the different SOX transcription factors cannot be explained by differential DNA bending activity.

All SOX factors bind to a common consensus motif C(A/T)TTG(A/T)(A/T) (Harley et al. 1994). This motif is so short and degenerate that it is found abundantly throughout the genome. Even though this motif is preferred by all 20 Sox proteins, variations in the heptameric consensus site and flanking regions and an overall substantial degeneracy are tolerated (Mertin et al. 1999, Badis et al. 2009). Since there are no apparent differences in the motifs recognized by different SOX factors that would allow us to distinguish or determine target gene selectivity of individual SOX members, other mechanisms must account for their context dependent function.

Indeed, a transcription factor with a non-palindromic six-base-pair recognition sequence will select close to 1,5 million sites in a genome of $3x10^9$ base pairs, sixty times larger than protein-coding genes, while a combination of two factors will select only 360 regulatory target sites out of the entire genome (Kondoh & Kamachi 2010). As recognition sequences of transcription factors are generally degenerated, the actual genomic sites selected by the combination of two factors may correspond to the number of genes that are regulated. Hence, the main contributing factor to the multitude of non-redundant and cell-type specific functions of Sox proteins is their dynamic association with different partner factors. The particular combinations and partner switches dictate genome engagement, gene expression programs and the developmental traits driven by SOX transcription factors (Hou et al. 2017).

2.4 PARTNER FACTORS

SOX proteins bind DNA with weak affinity and low specificity and the binding of a SOX protein alone does normally not lead to activation of gene transcription. The regulatory function that SOX factors exert on their target genes depends heavily on the available partner protein. Those partnerships are not fixed, but rather dynamic and change regularly depending on the cellular context and maturation state of the cell. SOX proteins can interact with a variety of different partner factors, such as lineage specific transcription factors, chromatin modifying enzymes and architectural factors (Kondoh & Kamachi 2010, Wegner 2010). In the following sections I will elaborate on a few selected interactions as exemplified by SOX2.

2.4.1 SOX2 and POU factors

The best characterized example of a SOX partnership is that of SOX2 and the POU factor OCT4 (*Pou5f1*). The functional interaction between these two factors was discovered in 1995, when Yuan et al. demonstrated the requirement for this heterodimeric complex for the embryonic stem cell (ESC) specific activity of the *Ffg4* enhancer. The studies demonstrated that SOX2 and OCT4 bind cooperatively to two adjacent cis-regulatory elements separated by a 3-base spacer and that the spacing of those sites is critical for

complex formation and transactivation of the *Fgf4* enhancer (Ambrosetti et al. 1997). Interestingly, while OCT1 can also form a complex with SOX2 on this *Fgf4* enhancer, it is completely inactive. On the other hand, the combination of SOX11 and the POU factor BRN2 can activate the *Fgf4* enhancer (Kuhlbrodt et al. 1998).

Since then, many cases for a functional SOX2/OCT4 partnership have been found, e.g. at regulatory regions of ESC markers like *Utf1*, *Sox2*, *Nanog* and *Pou5f1*, but in all these cases the SOX2 site and the OCT4 site lack the 3-base spacer in between them (Nishimoto et al. 1999, Tomioka et al. 2002, Rodda et al. 2005, Chew et al. 2005). Those genes are part of the core ESC regulatory circuits; thus, making SOX2 and OCT4 cooperativity fundamental for a stable stem cell state. Their status as core stem cell factors is further emphasized by the discovery that SOX2 and OCT4 can induce pluripotency in somatic cells, either in combination with other transcription factors (Takahashi & Yamanaka 2006) or even alone (Huangfu et al. 2008).

SOX-POU partnerships are not only essential for the induction and maintenance of stem cells, but are also involved in the progression of developmental processes and the transition from one cell state to another. For example, as pluripotent embryonic stem cells transition to multipotent neural progenitor cells, OCT4 expression is silenced and SOX2 partnership switches from OCT4 to BRN2 (*Pou3f2*) (Lodato et al. 2013). This partner switch represents just a glimpse of a much more extensive event. The SOX2 interactome changes dramatically as ESCs differentiate. Of the >60 proteins that are associated with SOX2 in ESCs, only 18 of these interactions were present 24 h after the induction of differentiation (Mallanna et al. 2010).

2.4.2 SOX2 and PAX6

A cooperative binding of SOX2 and a partner factor has also been observed in lens cells of the chicken eye. Here, SOX2 binds together with PAX6 to an enhancer of the ∂-crystallin gene, which results in the activation of this specific gene (Kamachi et al. 1995). Both factors bind to a composite site that consists of a SOX and a PAX binding sequence. A prerequisite for the interaction and subsequent activation of the gene is the precise spacing between the two binding sites (Kamachi et al. 2001). Interestingly, the PAX binding motif is only little similar to the consensus binding sequence of PAX6 and is bound poorly by PAX6 alone. Even though PAX6 binds this degenerate motif with low affinity and only in cooperation with SOX2, the complex shows a strong transactivation potential, whereas the replacement of the deviant PAX6 site with the consensus binding site abolished the transactivation potential despite stronger binding of PAX6. Furthermore, while SOX2 alone was able to bind the isolated enhancer *in vitro*, *in vivo* SOX2 binding was only observed in conjunction with PAX6 (Kamachi et al. 2001).

These findings demonstrate that 1) strong binding does not equal strong cooperativity or high activation potential and 2) the chromatin environment plays an important role in the engagement of transcription factors with DNA.

2.4.3 SOX and chromatin regulators

Multiple studies have shown an association of SOX factors with chromatin modifying enzymes and architectural factors. In 2011, Engelen et al. performed a proteomic analysis for SOX2-interacting proteins in neural progenitor cells. Among the 50 factors that physically interacted with SOX2 were proteins of the chromatin modifying complexes SWI/SNF and SMRT, as well as CHD7, a chromatin remodeling ATPase. In a comprehensive study from 2017, Chronis et al. investigated the molecular mechanisms that underlie the OCT4, SOX2, KLF4 and c-MYC induced reprogramming of mouse embryonic fibroblasts (MEF). They showed that, while the four OSKM factors predominantly bind to active somatic enhancers in the early stages of reprogramming, they immediately initiate their inactivation by recruiting the co-repressor HDAC1, a histone deacetylase. Further evidence for a SOX-co-repressor partnership comes from the Wegner lab, that showed the recruitment of co-repressors HDAC1 and CTBP2 to regulatory regions by SOXD member SOX5 in melanocytes (Stolt et al. 2008).

On the other hand, SOX proteins do not only associate with co-repressors, but also with members of co-activator complexes. In particular, SOX2 and SOX9 have been shown to recruit and physically interact with transcriptional co-activators/histone acetylases, for instance p300/CBP, TIP60 and EYA1 (Tsuda et al 2003, Chen at al. 2008, Chronis et al. 2017, Hattori et al. 2008, Zou et al. 2008).

In summary, SOX proteins alone are normally not sufficient to activate or repress gene expression. Their target gene selectivity and regulatory function rather depends on the combination with a specific partner factor. This molecular versatility is the basis to our understanding of how the same SOX factor can exert different functions and play different roles in different biological contexts.

2.5 REGULATION OF SOX EXPRESSION AND ACTIVITY

The activity of SOX proteins is regulated at multiple levels. The main determinant of their functional activity is their interaction with partner factors, as discussed above. Furthermore, SOX factors act in a dose-dependent fashion (Pevny & Nicolis 2010, Hagey & Muhr 2014, Whittington et al. 2015, Prévostel & Blache 2017), which means their protein levels must be tightly controlled. Both of these traits are regulated via post-translational modifications. These covalent modifications influence whether and how SOX factors homo- and heterodimerize, and affect DNA-binding activity, transactivation, protein stability and

intracellular localization (Hou et al. 2017, Kamachi & Kondoh 2013, Bernard & Harley 2010). Modifications include phosphorylation, acetylation, sumoylation, glycosylation and methylation (Jeong et al. 2010, Van Hoof et al. 2009, Baltus et al. 2009, Tsuruzoe et al. 2006, Jang et al. 2012, Zhao et al. 2011).

In addition, SOX activity is controlled at the transcriptional level. The expression of SOX factors is coordinated by both intrinsic and extrinsic signaling pathways. SOX2 expression, for instance, is stimulated by TGFB-signaling in ESCs and by EGF and SHH signaling, as well as PAX6, in NPCs (Chen et al. 2008, Mullen et al. 2011, Favaro et al. 2009, Lengler et al. 2005). In contrast, thyroid hormone signaling suppresses SOX2 expression as NPCs differentiate (Lopez-Juarez et al. 2012). Furthermore, SOX proteins are known to auto-regulate themselves or other SOX proteins ((Inoue et al. 2007, Kuzmichev et al. 2012, Amador-Arjona et al. 2015). To maintain ESCs in an undifferentiated state, SOX2 is auto-regulated in a positive feedback-loop by itself and in complex with other transcription factors, including OCT4, NANOG, SMAD1 and STAT3. (Uchikawa et al. 2003, Miyagi et al. 2006, Catena et al. 2004, Tomioka et al. 2002, Boyer et al. 2005, Chen et al. 2008).

3 SOX IN EMBRYONIC CNS DEVELOPMENT

3.1 SOX2 IN EMBRYONIC STEM CELLS

The development of a mammalian embryo begins with a totipotent zygote that undergoes a series of rapid cell divisions. The resulting multicellular morula further develops into the blastocyst. The formation of the blastocyst is the first event of cells acquiring spatially derived identities as the cells segregate into the inner cell mass (ICM) and trophectoderm. While the trophectoderm will form placental and extraembryonic tissue, the inner cell mass contains founder cells that will give rise to the embryo. As these founder cells have the ability to self-renew and develop into all three germ layers, they are true pluripotent stem cells. However, their pluripotent potential is a transient state confined to the blastocyst stage; and as embryogenesis proceeds, those cells will become progressively fate restricted. Embryonic stem cells (ESC) in culture are derived from those ICM cells and share many characteristics with them. But, in contrast to their *in vivo* counterparts, they are able to self-renew and amplify themselves during extended culture without compromising their pluripotency (Zhang & Cui 2014).

Pluripotent stem cells are regulated by a comprehensive network of transcription factors. For instance, zygotic deletion of SOX2 leads to a failure to form the pluripotent ICM (Avilion et al., 2003). SOX2 is not only required for the induction, but also the maintenance of pluripotent stem cells, as deletion from already established ESCs results in loss of pluripotency and their aberrant differentiation into trophectoderm-like cells (Masui et al. 2007, Kopp et al. 2008). This phenotype can be rescued by the overexpression of OCT4. One elementary role for SOX2 could therefore be the maintenance of a sufficient OCT4 expression. Surprisingly, overexpression of SOX2 (2-fold) has a similar effect as ablation and also causes ESCs to differentiate. Consequently, SOX2 levels need to be tightly controlled to maintain pluripotency, as either higher or lower expression levels lead to a loss of pluripotency in ESCs (Kopp et al. 2008). Indeed, SOX2 cooperates closely with other dose-dependent transcription factors, such as OCT4 and Nanog. Like SOX2, increasing or decreasing levels of OCT4 will induce differentiation in ESCs (Niwa et al. 2005). Together, SOX2, OCT4 and Nanog form an interdependent, core regulatory network that maintains self-renewal and pluripotency.

This network is furthermore suggested to function as lineage specifiers. SOX2 promotes neuroectoderm differentiation, while OCT4 and Nanog regulate the differentiation of ESCs into mesendoderm, a common precursor of mesoderm and endoderm (Loh & Lim 2011, Thomson et al. 2011, Wang et al. 2012). SOX2 induces neural lineage specification through the repression of alternative fates by suppressing key regulators of the mesoderm, like *Brachyury* (Thomson et al. 2011, Wang et al. 2012). Upon neural specification, ESCs start to differentiate into neural precursor cells (NPCs), which are often referred to as neural stem cells. In contrast to ESCs, these neural stem cells are fate restricted and can only develop

into cells of the neural lineage, i.e. neurons and glia, and are thus multipotent. These neural progenitors maintain SOX2 expression and, as in ESCs, SOX2 functions to maintain self-renewal and proliferation (Avilion et al. 2003).

3.2 SOX IN NEUROGENESIS

In the developing embryo, SOXB1 (SOX1, SOX2, SOX3) expression is largely restricted to the neuroectoderm (Wood & Episkopou 1999), which will give rise to the skin and nervous tissue. The nervous tissue includes neural crest and the neural tube, the latter of which will develop into the brain and spinal cord (= central nervous system, CNS). As most of my work is done in the embryonic mouse/chicken spinal cord, I will focus on SOX expression in this tissue in the following sections.

Neural stem cells (=neural precursor cells, NPCs) are located in the ventricular zone of the developing spinal cord. NPCs express high levels of SOX1-3 (Fig 1), which are not only expressed simultaneously and redundantly, but also have overlapping functions in the maintenance of undifferentiated neural progenitor cells (Bylund et al. 2003, Graham et al. 2003, Wood & Episkopou 1999). The functional redundancy of the SOXB1 factors has been demonstrated in various studies (Nishiguchi et al. 1998, Malas et al. 2003, Ferri et al. 2004). In general, the overexpression of any SOXB1 factor promotes proliferation and hinders neural differentiation, whereas inhibiting SOXB1 function causes cell cycle exit and premature differentiation (Bylund et al. 2003, Graham et al. 2003, Favaro et al. 2009). The mechanisms through which they maintain the undifferentiated state in neural progenitor cells is largely unknown. One mechanism is blocking the activity of proneural bHLH (basic helix-loop-helix) proteins like NGN2 and ASCL1 (Bylund et al. 2003).

Thus, in order for a NPC to differentiate, SOXB1 proteins need to be downregulated. Co-expressed with the SOXB1 proteins is a member of their SOXB2 sister group, SOX21. SOX21 has the opposite function of the SOXB1 members; it promotes cell cycle exit and neuronal differentiation (Sandberg et al. 2005, Uchikawa et al. 1999). Therefore, it is thought that the balance between SOXB1 and SOX21 is one of the determinants of whether a progenitor cell is maintained as a proliferating precursor or commits to neuronal differentiation. When the NPC receives extrinsic and intrinsic differentiation signals, it will exit the cell cycle, leave the ventricular zone, migrate laterally through the adjacent intermediate zone and eventually settle in the marginal zone of the developing spinal cord, where it initiates its physiological function as a neuron (Fig 1). This neuronal maturation is a stepwise process and each step is tightly regulated by a network of transcription factors. Proliferation and NPC genes are gradually downregulated, while neuronal markers are being upregulated. As the SOXB1 factors are downregulated, members of the SOXC (SOX4, SOX11, SOX12) group starting to be expressed. SOXC proteins can also be detected in NPCs, but are mostly expressed in post-mitotic differentiating neurons (Bergsland et al. 2006, Hoser et al. 2008). As SOXB1, SOXC factors are also expressed

redundantly and have overlapping functions. Overexpression of SOX4 and SOX11 in the developing chick spinal cord initiates a premature neuronal differentiation program as shown by the upregulation of pan-neuronal markers TUJ1 and MAP2. Surprisingly, these cells do not down-regulate SOXB1 proteins and thus, do not exit the cell cycle and remain proliferating (Bergsland et al. 2006). Thus, SOXC proteins play an important role in the acquisition of pan-neuronal characteristics.

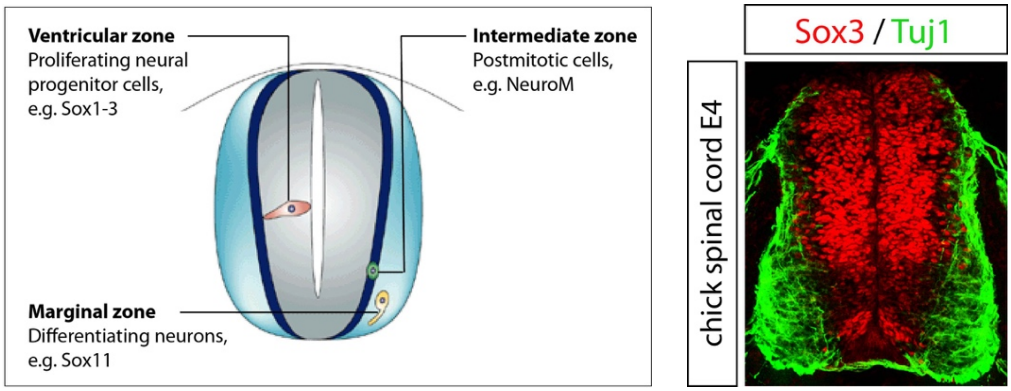

Figure 1. Illustration of the embryonic spinal cord (modified from Diez del Corral & Storey, 2001) and cross-section of a chick E4 spinal cord showing NPCs expressing SOX3 (red) and post-mitotic neurons expressing TUJ1 (green).

3.3 SOX IN GLIOGENESIS

Neural stem cells undergo a competence switch around day E10.5 of mouse embryogenesis (E5 in chick), which is referred to as the neuro- to gliogenic switch. While they give rise to neurons during the early stages of development, they deploy their full multipotent potential after the neuro- to gliogenic switch as they start to generate glial cells, such as astrocytes and oligodendrocytes. The SOXE member SOX9 starts to be expressed just slightly before the onset of gliogenesis (E9.5 in mice) and plays a key role in glial specification. The conditional deletion of SOX9 in the embryonic mouse spinal cord leads to an increase in the number of motor- and interneurons at the expense of glial cells, which demonstrates the importance of SOX9 for the neuro- to gliogenic switch (Stolt et al. 2003). After specification, the glial precursor cells will generate first astrocytes and later oligodendrocytes. In astrocyte precursors, SOX9 induces the expression of Nuclear Factor I-A (NFIA) and together they activate numerous important astrocytic genes, e.g. Apcdd1, Mmd2 and Nfe2l1 (Martini et al, 2013, Kang et al. 2012, Molofsky et al. 2013). SOX9 is the only SOXE member present in astrocytes and its expression is maintained throughout all stages of astrocyte maturation.

The other glial subgroup, oligodendrocytes, originates from the pMN domain of the ventricular zone of the ventral spinal cord. NPCs in this region express the bHLH transcription factor OLIG2 and produce first motor neurons, before they generate

oligodendrocyte precursor cells (OPCs). As in astrocyte precursors, SOX9 induces the expression of NFIA and continues to be expressed in OPCs past the initial specification event (Stolt et al. 2003). However, expression levels are much lower and SOX9 executes the opposite function. In the OPCs of the pMN domain, SOX9 cooperates with OLIG2 to induce the expression of SOX10, which in turn activates many crucial oligodendrocyte genes (Küspert et al. 2011). With increasing levels of SOX10, SOX9 expression is decreasing (Stolt et al. 2003). After the induction of SOX10, NFIA now preferentially interacts with SOX10. Consequentially, SOX10 sequesters NFIA and thereby prevents the activation of astrocytic genes (Glasgow et al. 2014). Hence, even though SOX9 is required for both astrocyte and oligodendrocyte specification, it exerts different functions thereafter depending on the cellular environment and transcriptional co-factors. While it promotes astrocytic differentiation in astrocyte precursors, it promotes oligodendrocytic differentiation and the suppression of astrocytic genes in OPCs.

Unlike in astrocytes, SOX9 is down-regulated in OPCs before the terminal differentiation. In contrast, SOX10 expression continues past the OPC stage and is maintained in fully differentiated oligodendrocytes. Terminal differentiation and myelination start at the time of birth and are characterized by the induction of myelin gene expression, increased lipid biosynthesis and myelin sheath formation. SOX10 is essential for the induction and maintenance of these processes, as it directly targets many key regulatory genes that are involved in myelination, such as *Mbp* and *Myrf* (Stolt et al. 2002, Hornig et al. 2013, Lopez-Anido et al. 2015). Consequentially, the conditional deletion of SOX10 in mice leads to severe dysmyelination (Stolt et al. 2002). It activates its target genes often in cooperation with partner factors, like e.g. OLIG1, MYRF and TCF7L2/TCF4 (Li et al. 2007, Hornig et al. 2013, Zhao et al. 2016).

In summary, gliogenesis relies on SOXE factors. They are involved in and regulate each step from specification until terminal differentiation. Even though they are expressed (and operate) partly redundantly, they exert individual functions in differentiating astrocytes and oligodendrocytes.

4 GENE REGULATION IN THE CONTEXT OF CHROMATIN

4.1 THE CHROMATIN LANDSCAPE

Our genetic information is stored in ca. 3 billion base pairs that form, together with sugar and phosphate molecules, two long strands in the shape of a spiral called a double helix, the DNA. If stretched out, the DNA of a single cell would be about 2 m long. Hence, in order to fit into the microscopic cell nucleus, it does not appear as free linear strands, but is highly condensed and wrapped around proteins, called histones. They compact and organize DNA. This DNA-protein complex is called chromatin and its 3D structure has an enormous impact on transcriptional regulation.

Chromatin exists in different states that are defined according to their compaction, accessibility and transcriptional activity (Fig 2). Heterochromatin, also called type R chromatin, is highly condensed and correlates with silenced transcription. These actively repressed domains are inaccessible and are accompanied by repressive covalent histone modifications, such as H3K9 or H3K27 methylation, although usually type R domains do not carry both marks at the same time (Ho et al. 2014). Its folded structure and the tightly packed histones pose a barrier to transcription factors and polymerases, leaving heterochromatin inaccessible to the transcriptional machinery.

In order to initiate gene expression, the chromatin has to be opened up to allow access to the DNA template. This is achieved either by transiently displacing histones by chromatin remodeling complexes (Smith & Peterson, 2005) or by enzymatically modifying the histones by the addition of acetyl, methyl or phosphate groups (Fischle et al. 2005). The open, active chromatin is also known as chromatin type A. These active domains can usually be accessed by any transcription factor and are accompanied by active histone marks, such as histone H3 methylation on lysine 4 (H3K4me) and H3K9 and H3K27 acetylation (The ENCODE Project Consortium 2011). Large-scale chromatin-mapping studies have revealed that about half of the genome consists of vast stretches that possess neither active nor repressive histone marks (Ho et al. 2014, Roadmap Epigenomics Consortium et al. 2015). These domains are called low signal chromatin or chromatin type L. Although it is inaccessible to most transcription factors, pioneer factors can target this type of chromatin (Soufi et al. 2012, van Oevelen et al. 2015) and attract chromatin modifying enzymes, resulting in either open, active chromatin or further compaction and the formation of repressive heterochromatin (see next chapter). Genome-wide mapping studies showed that active promoters are usually depleted of nucleosomes (Schones et al. 2008, Valouev et al. 2011, Teif et al. 2012). However, promoters tend to open ubiquitously among various cell types, as demonstrated by DNA accessibility studies (Thurman et al. 2012), suggesting that

ubiquitously expressed factors keep promoters accessible (Iwafuchi-Doi & Zaret 2014). In contrast, enhancers open in a tissue specific manner (Thurman et al. 2012).

In summary, chromatin undergoes continuous structural changes and chromatin accessibility is highly dynamic. This variable access to regulatory elements provides the key for context and tissue specific gene expression profiles.

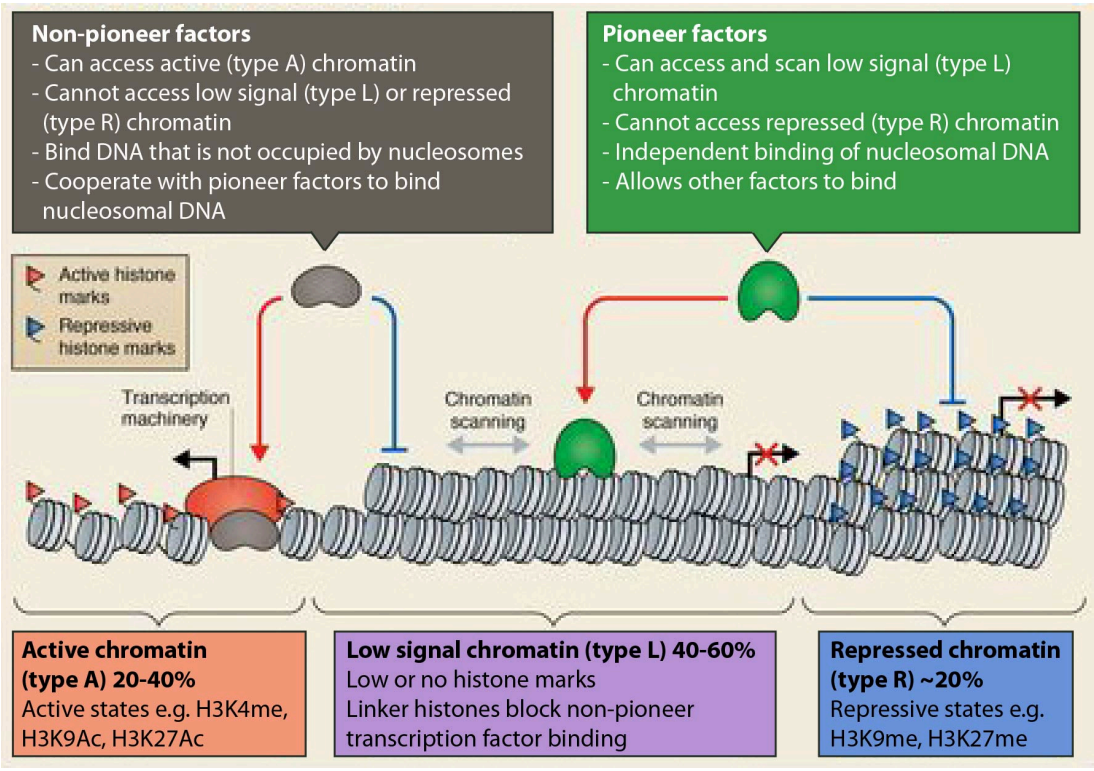

Figure 2. Accessibility of different types of chromatin (Iwafuchi-Doi & Zaret, 2016)

4.2 PIONEER FACTORS IN CELL FATE DECISIONS

Pioneer factors have the unique ability to overcome the constraints of nucleosomal DNA, i.e. DNA wrapped around histones, and can access silent genes that have not been marked for activity. Subsequently, this enables the binding of other transcription factors, cofactors, chromatin modifying and remodeling enzymes, resulting in the assembly of a complex of regulatory proteins that either initiate gene expression or lead to the formation of heterochromatin and gene repression. The hierarchical mechanisms by which transcription factors engage target sites provide insights into ways to modulate the process and hence cell fate control (Iwafuchi-Doi & Zaret 2014). This makes pioneer factors a fundamental tool for understanding and altering lineage specification. Chromatin states and gene expression profiles are cell type specific, and genes specific to the function of one cell type may be antagonistic the function of another cell type (Iwafuchi-Doi & Zaret 2016). Therefore, in order to establish new cell fates during embryonic development, genes specific to the

original cell fate must be suppressed and genes specific to the new cell fate, which are silent and inactive, must be accessed and activated. This concept is also the basis for cellular reprogramming, the conversion of one specific cell type into another, which I will elaborate on more extensively in the next chapter. Thus, cell fate changes, naturally or forcibly induced, usually depend on the pioneering function of certain transcription factors.

The mechanism of how the different pioneer factors recognize their target sites varies. The paradigm pioneer factor, FOXA1, binds to the full motif on one side of the DNA, leaving the other side free to bind core histones (Clark et al. 1993, Cirillo et al., 1998). Its binding displaces linker histones and keeps local nucleosomes accessible (Cirillo et al. 1998, Iwafuchi-Doi et al. 2016). Also SOX2 has been suggested to function as a pioneer factor, along with OCT4, KLF4 (Soufi et al. 2012) and the bHLH protein ASCL1 (Wapinski et a. 2013). In contrast to FOXA1, these factors recognize only a partial DNA sequence of their canonical binding motifs, which is compatible with nucleosome binding (Soufi et al. 2015). The authors of this study analyzed the binding sites of OCT4, SOX2 and KLF4 on nucleosome free versus nucleosome bound loci. While all three factors bound their canonical motifs in nucleosome free DNA, they targeted degenerate or partial motifs in nucleosome wrapped regions. The binding site of SOX2 was degenerated from CTTTGT to CTTTNT.

In summary, pioneer transcription factors are crucial for cell fate control and the regulation of transcriptional profiles in a cell type and context specific manner. Their primary engagement with closed chromatin enables non-pioneer factors to subsequently bind to these regions and initiate the secondary events that lead to appropriate gene expression.

4.3 INDUCED CELL FATE CHANGES (REPROGRAMMING)

Cellular reprogramming is the conversion of a mature and fully differentiated cell type (e.g. fibroblast) into another cell type. In most cases the desired cell types to be generated are proliferative progenitor cells, such as a NPCs or even pluripotent stem cells. However, a mature somatic cell can also be directly converted into a different mature somatic cell type without undergoing an intermediate pluripotent or progenitor state, which is referred to as transdifferentiation. As mentioned before, cellular reprogramming requires massive chromatin restructuring as the genes specific to the original cell fate must be repressed and genes specific to the desired cell fate must be activated.

Cellular reprogramming is typically achieved through forced overexpression of lineage-specific transcription factors. The most famous example of reprogramming is the conversion of fibroblasts into induced pluripotent stem cells (iPSCs) using only four transcription factors: SOX2, OCT4, KLF4 and c-MYC. This discovery by Yamanaka et al. in 2006 had an enormous impact on the field of regenerative medicine and was rewarded with a Nobel

prize in 2012. Since then, several SOX factors have been shown to be required or even sufficient for reprogramming, especially members of the SOXB1 and SOXE group.

SOX2 is widely expressed in the early neuroectoderm and in neural stem and progenitor cells. Adequately, it has been used extensively in various combinations to induce neural stem or progenitor cells, but also to revert cells back to a pluripotent state. Even though it is mostly used in combination with other transcription factors, SOX2 alone was shown to be sufficient to induce NPCs from either fibroblasts (Ring et al. 2012) or astrocytes (Niu et al. 2013). However, the generation of iPSCs with SOX2 alone has not yet been shown, suggesting that cofactor partnerships may be less essential for iNPC than for iPSC generation (Julian et al. 2017). In another interesting study from 2012, Lujan et al. narrowed down a pool of initially 11 neural transcription factors to a triple combination of SOX2, BRN2 and FOXG1 to convert mouse fibroblasts into NPCs. While the NPCs that were induced with this triple combination were tripotent and could be differentiated into functional neurons, astrocytes and oligodendrocytes, NPCs that were induced with only SOX2 and FOXG1 were only bipotent and lacked the potential to differentiate into oligodendrocytes. Moreover, whereas the transcription factors FOXG1 and BRN2 alone (but not SOX2 alone) were also capable of generating neural precursor cells, the addition of SOX2 was required for the generation of functionally mature neurons. In summary, SOX2's role in reprogramming is crucial as it is required for the de-differentiation process of somatic cells, but also functions as a specifier for the neural lineage.

Similarly to SOX2, the members of the SOXE group have also been attributed with the ability to convert fibroblasts. In combination with glial fate specifiers NFIA and NFIB, Caiazzo et al. (2014) showed that SOX9 can convert mouse fibroblasts into astrocytes that are transcriptionally and functionally comparable to native brain astrocytes. Furthermore, SOX9 in conjunction with KLF4 and c-MYC established induced chondrocytes from mouse fibroblasts (Hiramatsu et al. 2011, Tam et al. 2014). Another SOXE member, SOX10, can convert fibroblasts to functional oligodendrocytes when co-expressed with OLIG2 and NKX6.2 or ZFP536 (Najm et al. 2013, Yang et al. 2013). Like SOX2, SOX10 alone has also been demonstrated to be sufficient to autonomously generate multipotent neural progenitor cells (neural crest) from mouse and human fibroblasts (Kim et al. 2014).

In summary, these findings indicate that several SOX factors have the capacity to induce restructuring of chromatin landscapes and initiate a cascade of subsequent events which lead to the establishment of gene expression programs that are aberrant to the original cell fate.

5 AIMS

The main focus of my work lies on the transcriptional regulation of the different stages as a neural stem cell transitions to a neuronal or glial fate. In particular, we address the following questions:

What are the mechanisms that maintain stem cell identity? What triggers cells to commit to a certain fate? What instructs the cell to stop proliferating, downregulate progenitor genes and initiate a cell type specific gene expression program? How is this cell type specific identity maintained? How can certain transcription factors manipulate this mature state and reverse it back to a proliferative state? And overall, how are these processes regulated on a transcriptional level and what genomic features accompany them?

This thesis aims to illuminate the role of different SOX factors in the regulation of these processes and to contribute to our understanding of the mechanisms that drive neuro- and gliogenesis.

6 RESULTS AND DISCUSSION

6.1 PAPER I

Different members of the SOX transcription factor family play key sequential roles in each stage of neurogenesis; the differentiation process of an embryonic stem cell (ESC) to a neural precursor cell (NPC) to a mature neuron. But how these SOX factors regulate the specific gene expression programs of each stage is not well characterized.

We have used ChIP-sequencing to generate and compare genome-wide binding data for SOX2, SOX3 and SOX11 in ESC derived NPCs and immature neurons, respectively. A comparison of the binding pattern of SOX3 in NPCs to the binding pattern of the transcriptional coactivator p300, which occupies active enhancers in the developing embryo (Visel et al. 2009), revealed that 40% of the SOX3 peaks corresponded to regions that were labeled as active enhancers in the brain, but not in ESCs or limb tissue. Similar results were observed for SOX11 peaks. Thus, many of the SOX3 and SOX11 bound regions correspond to functional enhancers that are active in the developing CNS. Indeed, the genes associated with SOX3 peaks were enriched for genes involved in CNS development. Interestingly, SOX3 bound not only genes that are expressed in NPCs, but also genes that are selectively expressed in later stages of neurogenesis, i.e. in SOX3-negative differentiated neurons. Surprisingly, the binding pattern of SOX3 in NPCs overlapped substantially with the binding pattern of SOX11 in immature neurons. This promiscuous binding pattern of the different SOX factors raised the question of how their unique and common target genes are expressed. While genes that were first bound by SOX3 and later by SOX11 were most significantly enriched for genes expressed in NPCs, genes targeted uniquely by SOX3 and SOX11 were most significantly expressed in late populations of neurons and glia. Thus, SOX3 binds genes that are active in NPCs as well as genes that are silent in NPCs and will later be bound and activated by SOX11 in differentiating neurons. This raised the question of how SOX factors regulate their target genes in a cell type restricted fashion despite their broad range of binding. Functional studies with SOX3 and SOX11 on isolated neuronal enhancers showed that, even though these regions were bound by SOX3 in NPCs, only SOX11 could activate theses enhancers and drive the expression of a reporter construct. The presence of SOX3 suppressed the SOX11-mediated activation. Additional *in vitro* experiments indicated that this suppression was due to competitive binding rather than active repression. These results suggest that SOX3 prebinds silent neuronal genes in NPCs to prevent their premature SOX11-mediated induction.

Previous studies have shown similar characteristics for SOX2, as SOX2 in ESCs binds many silent genes that will be activated at later stages of development (Boyer et al. 2005, Lee et al. 2006). Several of these genes carried active as well as repressive histone modifications. When we examined the SOX3 and SOX11 bound targets for these histone marks by ChIP-

qPCR, we discovered that SOX3 bound regions expressed in NPCs carried only the active mark H3K4me3, whereas SOX3 bound neuronal/glial genes carried both the active H3K4me3 and the repressive H3K27me3 mark. As these neuronal genes become bound and activated by SOX11 in immature neurons, the repressive mark is lost and only the active modification remains. In contrast, NPC genes that were bound by SOX11 lost the active mark and were only associated with H3K27me3. This data was complemented with additional histone modification ChIP-data from a myoblast cell line misexpressing SOX3, and together they indicate that these SOX factors can modulate and alter histone modifications and induce local epigenetic changes.

To investigate if a similar prebinding pattern can also be observed at even earlier stages of neurogenesis, we compared our NPC SOX3-ChIP to a SOX2-ChIP in ESCs (Chen et al. 2008). Nearly 50% of the regions that were bound by SOX3 in NPCs were also bound by SOX2 in ESCs. While genes that were bound exclusively by SOX2 were mainly expressed in ESCs, genes sequentially bound by SOX2 and SOX3 were most highly expressed in NPCs and genes targeted exclusively by SOX3 were associated with neuronal and glial expression. Thus, similar to the prebinding of SOX3 to silent neuronal and glial genes, SOX2 prebinds many silent genes in ESCs that will later be targeted and activated by SOX3 in NPCs. Furthermore, genes that were targeted by SOX2 in ESCs and carried bivalent histone marks were strongly expressed in NPCs, but not in progenitor populations of meso- or endodermal origin. Thus, in parallel to its function in maintaining pluripotent stem cells, SOX2 prebinding specifies neural lineage-specific gene programs.

SOX transcription factors have been shown to function as architectural proteins that induce bending of DNA and local unwinding (Ferrari et al. 1992). Therefore, the prebinding of SOX proteins to poised genes may protect these genes from epigenetic repression, such as heterochromatin formation, and promote the formation of permissive chromatin, thus priming them for later activation. Another scenario could be that SOX binding to poised enhancers primes them for either later activation or terminal repression as the bivalent histone marks are resolved into monovalent marks at later developmental stages. Furthermore, SOX2 bound regions in human ESCs have been shown to be depleted of DNA methylation (Lister et al. 2009). Thus, in addition to preventing their premature expression, SOXB1 occupancy could also facilitate the later activation of genes at the proper developmental stage and cellular context. This mechanism has also been demonstrated outside of the neuronal lineage: the liver-specific enhancer of the Alb1 gene is prebound by FOXD3 already in ESCs, which protects this region from methylation (Xu et al. 2007). FOXD3 is later replaced by FOXA1 which activates this enhancer in liver cells (Xu et al. 2009). A prerequisite for such a cell type specific activation of SOX bound enhancers would be the cooperation with specific partner factors, since SOX factors alone bind DNA with low affinity and low specificity and, as demonstrated above, the binding of SOX factors does not lead automatically to an activation of the corresponding gene.

In summary, in addition to maintaining proliferating cells in an undifferentiated state, SOXB1 proteins also function to prepare these cells for differentiation by occupying and epigenetically predisposing genes to be activated at later steps of neurogenesis. The sequential binding of different SOX factors regulates coordinated lineage selection and maintenance from early pluripotency to later post-mitotic differentiation stages.

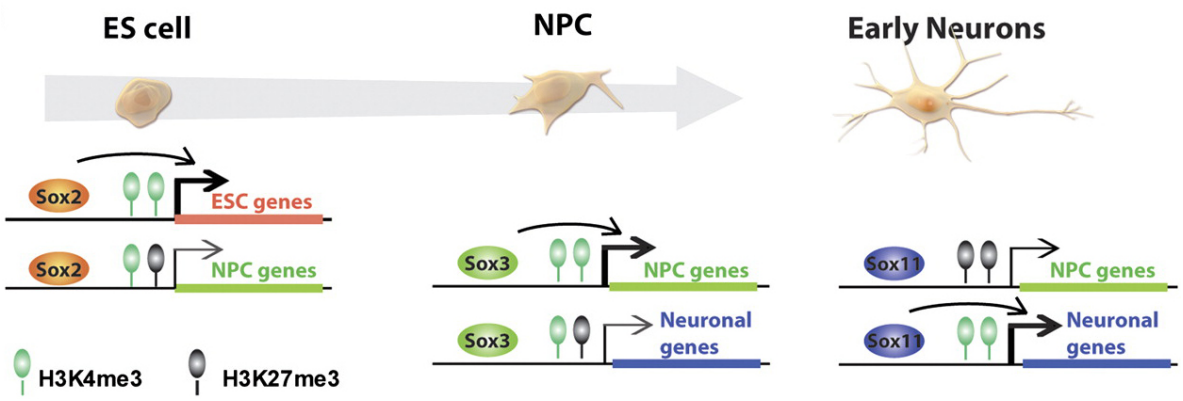

Figure 3. Model depicting the sequential binding of SOX proteins to common target genes during the differentiation of ESCs towards the neuronal lineage, illustrating the association between SOX prebinding and histone modifications (Bergsland et al. 2011).

6.2 PAPER II

In this paper we follow up on our sequential binding model from paper I and expand it to gliogenesis. In paper I we demonstrate that SOX3 binds to silent genes in NPCs that will be expressed at later stages of neuro- and gliogenesis, i.e. in mature neurons and glial cells. We showed that SOX11 will replace SOX3 on those neuronal genes and induce their activation. The question that arose consequentially was, if a similar model could be directing the activation of glial genes. SOX9 is expressed in glial precursor cells (GPCs) and is essential for both astrocyte and oligodendrocyte development (Stolt et al. 2003). While it continues to be expressed in mature astrocytes, it is downregulated during the differentiation of oligodendrocytes as another SOX factor, SOX10, is starting to be expressed. SOX10 is required for the terminal differentiation of oligodendrocytes and the induction of genes involved in myelin production (Stolt et al. 2002, Hornig et al. 2013, Lopez-Anido et al. 2015). Thus, we speculated that SOX9 and SOX10 could replace SOX3 on astrocyte and oligodendrocyte genes, respectively, and activate those. The first step to such an analysis was to define astrocyte and oligodendrocyte specific genes. Therefore, we performed single cell RNA-sequencing (scRNA-seq) on the developing mouse spinal cord and defined the transcriptomes specific to the major cell types, NPCs and GPCs as well as neurons, astrocytes and oligodendrocytes. Then we generated cell cultures from differentiating ESCs, that recapitulate the sequential generation of neurons and glial cells. We performed ChIP-sequencing of SOX3, SOX9, H3K27ac in GPCs (11 days of

differentiation) and compared the target genes to our defined cell type specific transcriptomes. Furthermore we ChIPped H3K27ac in NPCs (4 days of differentiation) and re-analyzed our SOX3 NPC-ChIP (Bergsland et al. 2011).

The genes that were bound by SOX3 in NPCs consisted of genes that are expressed in NPCs/GPCs, neurons, astrocytes and oligodendrocytes. When we mapped these SOX3 NPC-targets onto H3K27ac ChIP reads from NPCs and GPCs, we detected two clusters within these targets: while cluster II carried the active chromatin mark in both NPCs and GPCs, cluster I lost the active mark along the transition from NPCs to GPCs. Cluster I corresponded to genes that are expressed in NPCs, neurons and oligodendrocytes, but not astrocytes. In contrast, genes associated with cluster II were strongly enriched for astrocyte genes. Hence, even though SOX3 prebinds neuronal and glial genes in NPCs, mainly the astrocytic genes seem to remain active in GPCs. Even though the binding pattern of SOX3 changes from NPCs to GPCs, 30% of the GPC targets had been prebound in NPCs. Astrocytic genes were most strongly enriched within these prebound genes, which is in line with the maintained active chromatin mark as NPCs transition to GPCs.

SOX9, which is co-expressed with SOX3 in GPCs, targeted mainly regions that were associated with astrocyte and oligodendrocyte genes. Furthermore, SOX9 co-targeted 20% of the SOX3 bound regions and these shared targets are enriched for astrocytic genes. Thus, while SOX9 prebound both astro- and oligodendrocytic genes in GPCs, a specific feature of astrocytic genes is that they were co-bound by SOX3. To further examine the prebinding of SOX9 to oligodendrocytic genes and a possible subsequent binding of SOX10, we compared our SOX9 GPC-ChIP to a publically available SOX10 ChIP in oligodendrocytes (Gomez-Lopez et al. 2011). SOX9 prebound many regions in GPCs that would later be bound by SOX10 in mature oligodendrocytes. Those genes were mostly enriched for oligodendrocyte and astrocyte specific genes, which could indicate a repressive function of SOX10 since astrocytic genes are not expressed in SOX10 positive oligodendrocytes. Interestingly, no overlap could be detected between SOX10 peaks and SOX3 GPC-peaks. In summary, SOX9 and SOX3 co-prebind astrocyte genes in GPCs, but genes that are only prebound by SOX9 alone correspond to both astro- and oligodendrocyte genes and will subsequently be bound by SOX10 in oligodendrocytes. *De novo* motif analysis revealed the enrichment of NFI motifs among SOX9 and SOX3 co-bound astrocytic regions. SOX9 and SOX10 bound astrocytic regions were enriched for EHF motifs, whereas SOX9 and SOX10 bound oligodendrocytic regions showed an enrichment for ZBTB3 motifs. Among these putative partner factors, only NFIA was significantly expressed in our scRNA-seq cell clusters.

To investigate if SOX3 blocks glial gene expression in a similar fashion as neuronal genes, we cloned enhancers that were either bound by SOX3/SOX9 or SOX9/SOX10 and tested the regulatory function of SOX3, SOX9 and SOX10 on these. The astrocytic SOX3/SOX9 bound elements could be activated by SOX9 and also in synergy by SOX9 and NFIA, but not by SOX3. Instead, SOX3 suppressed their activation by SOX9 and NFIA. While

astrocytic SOX9/SOX10 elements did not show an activation by any SOX factor, oligodendrocytic SOX9/SOX10 elements were activated by SOX9 and SOX10 in an additive manner. Additional experiments *in vitro* and in the chick spinal cord confirmed that SOX3 can prevent the premature onset of astrocytic gene expression.

In this paper, we demonstrate at a genome-wide level how SOX3 and SOX9 preselect gene programs in NPCs and GPCs that are first activated during gliogenesis. The function of this prebinding could be, as discussed for paper I, to facilitate later activation, poise them for either later activation/terminal repression and/or to prevent their premature expression. The fact that SOX9 is an essential specifier of both the astrocytic and oligodendrocytic fate and the observation that it prebinds genes of both glial sublineages, raises the question of how these different target genes are regulated. We could identify a selective enrichment of the NFI motif only in sequences that were associated with astrocyte genes. This is in line with NFIA's cooperation with SOX9 during astrocyte differentiation (Martini et al, 2013, Kang et al. 2012, Molofsky et al. 2013).

6.3 PAPER III

SOX2 is expressed in a variety of stem cells and has key functions in regulating fundamental processes of stem cells, including their maintenance, proliferation and cell fate decision. However, whether the common characteristics of different stem cells are regulated through similar transcriptional mechanisms is not well understood. To examine if and how SOX2 regulates common and cell type specific gene sets in different stem cell populations, we performed SOX2 ChIP-seq on E11.5 mouse stomach and lung/esophagus tissue and compared it to SOX2 ChIP-seq data from E11.5 mouse cortex and spinal cord (Hagey et al. 2016).

We found that SOX2 binding was primarily cell-type specific and only 32 peaks were common between all four tissues. The SOX2 binding profile was more similar between tissues of the same germ layer than between endodermal and neuroectodermal tissues. This cell type specificity could possibly be due to interaction with distinct, tissue specific partner factors. *De novo* motif analyses revealed an enriched OTX1 motif in regions that were bound by SOX2 specifically in the cortex and a PAX2 motif in the regions bound in the spinal cord. DNA-regions targeted specifically by SOX2 in the stomach were enriched for GATA4 and HNF1A motifs, while those targeted in the lung/esophagus included FOXA1 and TEAD4 motifs. The regions that were commonly bound by SOX2 in neural and endodermal cells were enriched for ZEB1 and ZBTB33 binding motifs. These transcription factors belong to families that have been shown to interact with SOX proteins and they are appropriately expressed within the different tissues, which led us to investigate a potential physical and functional interaction with SOX2. Co-immunoprecipitations with the neuro specific OTX1, the endoderm specific FOXA1 and the common ZEB1 revealed a physical interaction between SOX2 and these transcription factors. As DNaseI treatment abolished

the SOX2-OTX1 complex formation, these factors' interaction appears to be DNA-dependent. To examine a functional interaction between SOX2 and the putative partner factors, we cloned enhancers that were commonly or specifically bound by SOX2 in neural and endodermal tissues and tested the activation potential of SOX2 in combination with OTX1, FOXA1 and ZEB1. CNS specific regions could be activated by SOX2 and OTX1 in an additive manner, while stomach specific regions were activated by SOX2 and FOXA1. Regions that were commonly bound by SOX2 in neural and endodermal tissues were activated by SOX2, but co-expression of ZEB1 repressed this activation.

To investigate how the distinct binding patterns of SOX2 are correlated with the gene expression profiles of neural and endodermal tissues, we performed an RNA-seq of SOX2-GFP+ cells isolated from the cortex, spinal cord, stomach and lung/esophagus of *Sox2-Gfp* mice. The comparison of the SOX2 bound targets with genes differentially expressed in cortex, spinal cord, stomach and lung/esophagus showed that cell type specific SOX2 binding was significantly enriched around genes specifically expressed in the corresponding tissue. The gene ontology for those genes was enriched for cell type specific terms such as "pallium development", "cell differentiation in spinal cord", "embryo digestive tract development" and "lung alveolus development". Interestingly, genes commonly bound by SOX2 in neural and endodermal tissue were enriched for more general GO terms such as "regulation of stem cell proliferation" and "regulation of stem cell differentiation". The observation that the binding pattern of SOX2 reflects the gene expression profiles of the corresponding tissue indicated that SOX2 binding could be instructive and SOX2-bound regions could drive cell type specific gene expression. To examine this possibility we tested reporter constructs containing tissue specific or commonly bound DNA regions in zebrafish embryos. Even though these regions were randomly integrated into the genome, neural specific enhancers were predominantly activated in the CNS and endoderm specific enhancers predominantly in the endoderm. Common regions were activated in both tissues. Thus, SOX2 bound regions can function as CMRs (cis-modulatory region) and drive cell type specific gene expression.

As the commonly bound genes were enriched for stem cell proliferation terms and SOX2 has previously been shown to regulate proliferation in the developing mouse cortex in a dose-dependent fashion (Hagey et al. 2014), we next sought to investigate if SOX2 regulates stem cells outside the cortex in a similar fashion. The correlation of SOX2 expression levels and cell proliferation in the E11.5 mouse spinal cord showed that cells with low levels of SOX2 incorporated more BrdU than cells with high SOX2 levels, indicating a higher proliferation and turnover rate. As a common feature of stem cells is a slowly proliferative state, high levels of SOX2 promote "stemness" and maintenance of stem cells. In the posterior stomach, decreasing SOX2 levels between E11.5 and E15.5 correlated with an increase of BrdU+ cells. Additional *in vivo* experiments in the chick spinal cord and mouse stomach confirmed that SOX2 dictates the rate of proliferation via cell cycle regulation, a mechanism similar to that in the cortex.

In summary, SOX2 displays functions that are common to all stem cells as well as cell type specific gene regulation. By maintaining stem cell identity and "stemness" via the repression of CyclinD1, SOX2 controls global functions in stem cells. On the other hand, SOX2 targets chromatin in a cell type specific fashion and its bound CRMs are instructive for cell type specific gene expression, likely due to cooperation with partner factors.

6.4 PAPER IV

The differentiation of a stem cell into a mature cell type is a stepwise process that is regulated by lineage-specific transcription factors which orchestrate changes in chromatin and gene expression profiles, specific for the generated cell type. However, by reactivating molecular programs typical of stem and progenitor cells, various studies have demonstrated that the differentiated state is more dynamic than previously thought. Reprogramming studies have shown the conversion of mature cells, e.g. fibroblast, to proliferating neural progenitor cells through forced overexpression of various transcription factors, but how these transcription factors function to impose neural fates is unclear.

In this paper we misexpressed SOX2 and FOXG1 either alone or in combination in mouse embryonic fibroblasts (MEFs) to examine their initial engagement with chromatin in the early stages of reprogramming. Principal component analysis (PCA) and peak overlap revealed that the binding pattern of both factors was significantly altered in the presence of the other factor. Furthermore, we observed that FOXG1's binding pattern changed more drastically than SOX2's and became more similar to the binding profile of SOX2. This indicates that SOX2 attracts FOXG1 and instructs it to bind in close proximity. Consistently, we found a degenerated SOX motif enriched among sequences that were newly bound by FOXG1 in the double factor condition. A co-immunoprecipitation confirmed a physical interaction between these two factors. As SOX2 has been shown to function as a pioneer factor, we speculated that SOX2 can bind DNA regions that are closed in MEF cells and allow FOXG1 to subsequently target those regions. We analyzed our ChIP data in respect to publically available ATAC-seq, H3K27ac and H3K9me3 ChIP data from WT MEF cells (Chronis et al. 2017) and found that SOX2 displayed the ability to bind regions that are inaccessible in WT MEFs and devoid of the active chromatin mark H3K27ac. In contrast, FOXG1 bound regions that were open and carried the active mark. Interestingly, in the presence of SOX2, FOXG1's binding became more similar to that of SOX2 and it bound to regions that were previously closed and depleted of the active mark. Notably, neither SOX2 nor FOXG1 were able to bind to DNA regions that were enriched for the heterochromatin mark H3K9me3.

Next, we were interested in the function of the genes that were targeted by these two factors. Surprisingly, gene ontology (GO) analysis showed that SOX2's target genes in MEF cells were highly enriched for neuronal terms, both in the absence and presence of FOXG1, even though this enrichment was stronger in the presence of FOXG1. Similarly, terms

associated with the developing nervous system were strongly enriched among genes that were targeted by FOXG1 in the presence of SOX2, but underrepresented in the absence of SOX2. Strikingly, genes co-targeted by SOX2 and FOXG1 in the double-factor condition were strongly enriched for terms related to cellular proliferation, which is consistent with the finding that FOXG1 is necessary for the proper proliferation of cortical NPCs (Manuel et al. 2011). Hence, SOX2 and FOXG1 bind to genes in MEF cells that are involved in the regulation of cellular proliferation and nervous system development.

Since the expression of FOXG1 is mainly found in NPCs of the developing telencephalon (Tao & Lai, 1992), we next examined if this transcription factor could be involved in the regulation of forebrain-specific genes. To address this question, we generated transgenic mice that misexpress either FOXG1-IresGFP or just GFP under the control of the PAX6 enhancer. We collected spinal cord tissue at E10.5 and the FACS sorted cells were subjected to RNA-seq analysis. Notably, of the 20 most deregulated genes 18 were upregulated including proteins involved in chromatin modification, such as the helicase DDX3Y, the histone demethylases DM5D and UTY, while the two most downregulated genes consisted of ATOH1 and WNT1. Genes that were upregulated in the FOXG1 spinal cord compared to the GFP spinal cord were enriched for GO terms associated with CNS development such as the top three terms "axon guidance", "CNS development" and "synapse assembly". Gene set enrichment analysis (GSEA) revealed several significantly enriched gene sets associated with brain development, for instance "forebrain generation of neurons", "cerebellar cortex formation" and "telencephalon glial cell migration". Furthermore, when we compared the genes that were upregulated by FOXG1 to genes specifically expressed in the developing mouse cortex (compared to spinal cord, Hagey et al. 2016), we found that 34 of the cortex specific genes were upregulated in the FOXG1-positive spinal cord. Consistent with these findings, FOXG1 enhanced the binding of SOX2 to genes involved in the development of anterior neural structures in MEF cells. Thus, apart from synergizing with SOX2 in the targeting of neural genes during early reprogramming events in MEFs, FOXG1 appears to also confer anterior characteristics to the targeted gene sets.

7 CONCLUSIONS AND FUTURE PERSPECTIVES

This thesis aims to contribute to our understanding of the regulatory programs that orchestrate the differentiation of a stem cell into a nerve or glial cell. We have investigated the role of different SOX proteins in neural stem cells and subsequent cellular stages during neuro- and gliogenesis. We present a model of sequentially binding SOX factors that target genes that are active as well as genes that will become activated or repressed at later stages along the differentiation cascade, upon subsequent binding of another SOX factor. Thus, SOX proteins function to both activate specific gene sets while keeping others in a primed state. These primed genes are associated with bivalent chromatin marks which will be resolved into monovalent marks upon activation or repression. This indicates that SOX proteins can modulate and alter histone modifications and induce local epigenetic changes. Furthermore, we provide further evidence for SOX2's role as a pioneer factor. We show that SOX2 can bind to regions in fibroblast cells that were silent and inaccessible and alter the DNA occupancy pattern of a second transcription factor, which are key indicators of a pioneering function. These findings only cover the initial events in the engagement of SOX proteins with DNA. Further studies are needed to investigate the secondary events that lead to subsequent changes in local chromatin structure and the assembly of large protein complexes at gene regulatory sequences. Another question that needs to be explored further is how SOX proteins confer cell type specificity to their binding and their regulatory function. We show that SOX factors bind relatively promiscuous and to a broad range of gene sets, yet they exert distinct functions depending on the cellular context and developmental stage. This can most likely be attributed to the association with specific partner factors. We have identified a few putative candidates in our studies, but more extensive work is needed to expand this circle and help us to gain more insight into how SOX factors function in the specification of neural sublineages and the mechanisms through which they regulate neural stem cells and their progeny.

8 ACKNOWLEDGEMENTS

Even though I decided to go back to Germany, I am thankful for the opportunity to have lived in the beautiful city of Stockholm for 10 years. I'm truly grateful for all the experiences I gathered here and all of you people have contributed to these moments that I now call memories. Thank you for your support and for making my time here colorful and interesting. I'm glad to have met so many fantastic people, many of whom I call friends now. My time at the **Ludwig Institute for Cancer Research** has shaped my life forever and will always be a big part of me.

First and foremost, I would like to thank my supervisor, **Jonas.** Thank you for giving me the opportunity to do my PhD in your lab, for your time and for the chance and freedom to navigate every step of a project, from the basic footwork to acquiring the bioinformatical know-how. These years have been truly educational.

To my co-supervisor, **Johan**, for all the laughs and nonsense and making the day to day PhD routine less dull.

Thanks to all the present and past members of the Muhr lab, including **Idha, Florie, Danny, Maria, Daniel, Magnus, Bhumica, Gaëlle** and **Alex**. Thank you for a friendly and helpful environment, for all your input over the years and for creating memorable moments in the lab, in the office, on lab retreats or over some drinks outside of work.

Thanks to all the present and past members of all the other groups at LICR: Perlmann, Holmberg, Sandberg, Schlisio, Deng and Stenman. Special thanks to the people with whom I shared the office, in particular **Linda, Hilda, Lina** and **Fredrik**, for good laughs, nice chats about life and beyond and for creating a joyful atmosphere. A big thank you also to: **Omid**, **Sissy**, **Chris**, **Sven, M'takai**, **Nick** and **Katarina** for sharing and caring, kind words and valuable discussions.

And great big thank you to **Charlotta, Mats, Eliza, Soheilla, Erika** and **Jorge**, the indispensable staff at the Ludwig Institute for Cancer Research. You were the backbone of Ludwig!

Cécile, **Eva**, **Vilma** and **Konstantinos**, you were not only colleagues, but also dear friends. **Cécile**, you were part of this journey from the beginning and I'm truly grateful for all the memories we created over the years, from working together to traveling together to sharing the ups and downs, both professional and personal. **Eva**, my girl, I'm so glad our paths crossed. You made Stockholm so much more fun! **Vilma**, you are my second-longest friend here in Stockholm. Thanks for all the silly and sincere moments. **Konstantinos**, Mr. Bootylicious, I honestly can't remember how it was before you joined us. It just feels like you were always part of our innermost circle. You brought lots of fun and you will always

have a place in my heart. Thank you guys for making working at Ludwig fun! I will miss you like crazy.

A big shout out to the bagpipers gang for good company and so many great nights filled with loud (very loud!) laughter, heartfelt conversations, utter nonsense and lots of beer. Cheers to all you beautiful people! **Michael** and **Eliane**, keep on being such nice, sweet, caring and, most of all, fun people. I will be forever grateful! **Gert-Jan**, you are a great asset to our crew, I hope Stockholm will treat you well. **Sophie**, you are an inspiration when it comes to work-life balance, hard working and still enjoying life to the fullest, rock on! **Antje and Martin**, I'm happy to have met you crazy people! **Johan**, our trophy Swede, I'm glad you joined us that evening and the many more evenings thereafter. **Aga**, thanks for nice chats and good company.

Danke **Alex** und **Birgit** für die vielen lustigen Mädelsabende und die aufmunternden Worte.

Ralf, du hast den Grundstein for meine wissenschaftliche Laufbahn gelegt. An einem meiner größten Tiefpunkte hast du einmal zu mir gesagt: "Denk dran, auch kleine Schritte bringen einen voran." Daran hab ich noch oft zurückgedacht. Vielen Dank!

Zu guter Letzt danke ich meinen **Eltern**, die mir die Grundsteine für alles gelegt haben. Danke für die "Wenn du denkst du denkst, dann denkst du nur du denkst"-Mentalität!

9 REFERENCES

Amador-Arjona A, Cimadamore F, Huang CT, Wright R, Lewis S, Gage FH, Terskikh AV: SOX2 primes the epigenetic landscape in neural precursors enabling proper gene activation during hippocampal neurogenesis. Proc Natl Acad Sci USA (2015), 112(15):E1936-45.

Ambrosetti DC, Basilico C, Dailey L: 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 (1997), 17:6321–9.

Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R: Multipotent cell lineages in early mouse development depend on SOX2 function. Genes Dev (2003), 17:126-140.

Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA, Chan ET, Metzler G, Vedenko A, Chen X, et al: Diversity and complexity in DNA recognition by transcription factors. Science 324, 5935, (2009) 1720–1723.

Baltus GA, Kowalski MP, Zhai H, Tutter AV, Quinn D, Wall D, Kadam S: Acetylation of Sox2 induces its nuclear export in embryonic stem cells. Stem Cells (2009) 27, 2175-2184.

Bergsland M, Ramsköld D, Zaouter C, Klum S, Sandberg R, Muhr J: Sequentially acting Sox transcription factors in neural lineage development. Genes Dev (2011), 25(23):2453-64.

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

Bernard P, Harley VR: Acquisition of SOX transcription factor specificity through protein–protein interaction, modulation of Wnt signalling and post-translational modification, Int J Biochem Cell Biol (2010), 42 (3) 400–410.

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

Boyer LA, Lee TI, Cole MF, Johnstone SE, Levine SS, Zucker JP, Guenther MG, Kumar RM, Murray HL, Jenner RG, et al: Core transcriptional regulatory circuitry in human embryonic stem cells. Cell (2005), 122: 947-956.

Bridgewater LC, Walker MD, Miller GC, Ellison TA, Holsinger LD, Potter JL, et al: Adjacent DNA sequences modulate Sox9 transcriptional activation at paired Sox sites in three chondrocyte-specific enhancer elements. Nucleic Acids Res (2003), 31:1541–53.

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

Caiazzo M, Giannelli S, Valente P, Lignani G, Carissimo A, Sessa A, Colasante G, Bartolomeo R, Massimino L, Ferroni S et al: Direct conversion of fibroblasts into functional astrocytes by defined transcription factors. Stem Cell Rep (2015), 4:25-36.

Chen X, Xu H, Yuan P, Fang F, Huss M, Vega VB, Wong E, Orlov YL, Zhang W, Jiang J, Loh YH, Yeo HC et al: Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. Cell (2008), 133(6):1106-17.

Chew JL, Loh YH, Zhang W, Chen X, Tam WL, Yeap LS, Li P, Ang YS, Lim B, Robson P, Ng HH: Reciprocal transcriptional regulation of Pou5f1 andSox2 via the Oct4/Sox2 complex in embryonic stem cells. Mol. Cell. Biol. 25 (14) (2005) 6031–6046.

Chronis C, Fiziev P, Papp B, Butz S, Bonora G, Sabri S, Ernst J, Plath K: Cooperative Binding of Transcription Factors Orchestrates Reprogramming. Cell (2017), 168(3):442-459.e20.

Cirillo LA, McPherson CE, Bossard P, Stevens K, Cherian S, Shim EY, Clark KL, Burley SK, Zaret KS: Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. EMBO J (1998) 17, 244-254.

Clark KL, Halay ED, Lai E, Burley SK: Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. Nature (1993) 364, 412-420.

Diez del Corral R, Storey KG: Markers in vertebrate neurogenesis. Nat Rev Neurosci (2001), 2(11):835-9.

Engelen E, Akinci U, Bryne JC, Hou J, Gontan C, Moen M, Szumska D, Kockx C, van Ijcken W, Dekkers DH et al: Sox2 cooperates with Chd7 to regulate genes that are mutated in human syndromes. Nat Genet. 2011 Jun;43(6):607-11.

Favaro R, Valotta M, Ferri AL, Latorre E, Mariani J, Giachino C, Lancini C, Tosetti V, Ottolenghi S, Taylor V, Nicolis SK: Hippocampal development and neural stem cell maintenance require Sox2-dependent regulation of Shh. Nat Neurosci (2009), 12, 1248–1256.

Ferrari S, Harley VR, Pontiggia A, Goodfellow PN, Lovell-Badge R, Bianchi ME: SRY, like HMG1, recognizes sharp angles in DNA. EMBO J (1992), 11: 4497–4506.

Ferri AL, Cavallaro M, Braida D, Di Cristofano A, Canta A, Vezzani A, Ottolenghi S, Pandolfi PP, et al: Sox2 deficiency causes neurodegeneration and impaired neurogenesis in the adult mouse brain. Development (2004), 131(15):3805-19.

Fischle W, Wang Y, Allis CD: Histone and chromatin cross-talk. Current Opinion in Cellular Biology (2003) 15, 172–183.

Gubbay J, Collignon J, Koopman P, Capel B, Economou A, Münsterberg A, Vivian N, Goodfellow P, Lovell-Badge R: A genemapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. Nature (1990) 346, 245–250.

Glasgow SM, Zhu W, Stolt CC, Huang TW, Chen F, LoTurco JJ, Neul JL, Wegner M, Mohila C, Deneen B: Mutual antagonism between Sox10 and NFIA regulates diversification of glial lineages and glioma subtypes. Nat Neurosci 17 (2014) 1322–1329.

Graham V, Khudyakou J, Ellis P, Pevny L: SOX2 functions to maintain neural progenitor identity. Neuron (2003) 39:749–765.

Hagey DW, Klum S, Kurtsdotter I, Zaouter C, Topcic D, Andersson O, Bergsland M, Muhr J: SOX2 regulates common and specific stem cell features in the CNS and endoderm derived organs. PLoS Genet (2018), 14(2):e1007224.

Hagey DW, Muhr J: Sox2 acts in a dose-dependent fashion to regulate proliferation of cortical progenitors. Cell Rep (2014), 9(5):1908-1920.

Hagey DW, Zaouter C, Combeau G, Lendahl MA, Andersson O, Huss M, et al. Distinct transcription factor complexes act on a permissive chromatin landscape to establish regionalized gene expression in CNS stem cells. Genome Res Cold Spring Harbor Lab (2016), 26: 908±917.

Harley VR, Lovell-Badge R, Goodfellow PN: Definition of a consensus DNA binding site for SRY. Nucleic Acids Res (1994), 22:1500–1.

Hattori T, Coustry F, Stephens S, Eberspaecher H, Takigawa M, Yasuda H, et al: Transcriptional regulation of chondrogenesis by coactivator Tip60 via chromatin association with Sox9 and Sox5. Nucleic Acids Res (2008), 36:3011–24.

Hiramatsu K, Sasagawa S, Outani H, Nakagawa K, Yoshikawa H, Tsumaki N: Generation of hyaline cartilaginous tissue from mouse adult dermal fibroblast culture by defined factors. J Clin Invest (2011), 121:640-657.

Ho JWK, Jung YL, Liu T, Alver BH, Lee S, Ikegami K, Sohn KA, Minoda A, Tolstorukov MY, Appert A: Comparative analysis of metazoan chromatin organization. Nature (2014), 512, 449-452.

Hornig J, Fröb F, Vogl MR, Hermans-Borgmeyer I, Tamm ER, Wegner M: The transcription factors sox10 and myrf define an essential regulatory network module in differentiating oligodendrocytes. PLoS Genet 9 (2013) e1003644.

Hoser M, Potzner MR, Koch JM, Bosl MR, Wegner M, Sock E: Sox12 deletion in the mouse reveals nonreciprocal redundancy with the related Sox4 and Sox11 transcription factors. Mol Cell Biol (2008), 28: 4675–4687.

Hou L, Srivastava Y, Jauch R: Molecular basis for the genome engagement by Sox proteins. Semin Cell Dev Biol (2017), 63:2-12.

Huangfu D, Osafune K, Maehr R, Guo W, Eijkelenboom A, Chen S, Muhlestein W, Melton DA: Induction of pluripotent stem cells from primary human fibroblasts with only Oct4 and Sox2. Nat Biotechnol 26 (11) (2008) 1269–1275.

Inoue M, Kamachi Y, Matsunami H, Imada K, Uchikawa M, Kondoh H: PAX6 and SOX2-dependent regulation of the Sox2 enhancer N-3 involved in embryonic visual system development. Genes Cells (2007) 12 (9) 1049–1061.

Iwafuchi-Doi M, Donahue G, Kakumanu A, Watts JA, Mahony S, Pugh BF, Lee D, Kaestner KH, Zaret KS: The pioneer transcription factor FoxA maintains an accessible nucleosome configuration at enhancers for tissue-specific gene activation. Mol Cell (2016) 62, 79-91.

Iwafuchi-Doi M, Zaret KS: Cell fate control by pioneer transcription factors. Development (2016), 143(11):1833-7.

Iwafuchi-Doi M, Zaret KS: Pioneer transcription factors in cell reprogramming. Genes Dev (2014), 28(24): 2679–2692.

Jang H, Kim TW, Yoon S, Choi SY, Kang TW, Kim SY, Kwon YW, Cho EJ, Youn HD: O-GlcNAc regulates pluripotency and reprogramming by directly acting on core components of the pluripotency network. Cell Stem Cell (2012) 11, 62-74.

Jeong CH, Cho YY, Kim MO, Kim SH, Cho EJ, Lee SY, Jeon YJ, Lee KY, Yao K, Keum YS, et al: Phosphorylation of SOX2 cooperates in reprogramming to pluripotent stem cells. Stem Cells (2010) 28, 2141-2150.

Julian LM, McDonald ACH, Stanford WL: Direct reprogramming with SOX factors: masters of cell fate. Current Opinion in Genetics & Development (2017), 46:24–36.

Kamachi Y, Kondoh H: Sox proteins: regulators of cell fate specification and differentiation. Development (2013), 140(20):4129-44.

Kamachi Y, Sockanathan S, Liu Q, Breitman M, Lovell-Badge R, Kondoh H: Involvement of SOX proteins in lens-specific activation of crystallin genes. EMBO J (1995), 14:3510–9.

Kamachi Y, Uchikawa M, Tanouchi A, Sekido R, Kondoh H: Pax6 and SOX2 form a co-DNA-binding partner complex that regulates initiation of lens development. Genes Dev (2001), 15:1272–86.

Kang P, Lee HK, Glasgow SM, Finley M, Donti T, Gaber ZB, Graham BH, Foster AE, Novitch BG, Gronostajski RM, Deneen B: Sox9 and NFIA coordinate a transcriptional regulatory cascade during the initiation of gliogenesis. Neuron 74 (2012) 79–94.

Kim YJ, Lim H, Li Z, Oh Y, Kovlyagina I, Choi IY, Dong X, Lee G: Generation of multipotent induced neural crest by direct reprogramming of human postnatal fibroblasts with a single transcription factor. Cell Stem Cell (2014), 15:497-506.

Klaus M, Prokoph N, Girbig M, Wang X, Huang YH, Srivastava Y, Hou L, Narasimhan K, Kolatkar PR, Francois M, Jauch R: Structure and decoy-mediated inhibition of the SOX18/Prox1-DNA interaction. NucleicAcids Res 44 (8) (2016) 3922–3935.

Kopp JL, Ormsbee BD, Desler M, Rizzino A: Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. Stem Cells (2008), 26: 903-911.

Kuhlbrodt K, Herbarth B, Sock E, Enderich J, Hermans-Borgmeyer I, Wegner M: Cooperative function of POU proteins and SOX proteins in glial cells. J Biol Chem (1998), 273:16050–7.

Küspert M, Hammer A, Bösl MR, Wegner M: Olig2 regulates SOX10 expression in oligodendrocyte precursors through an evolutionary conserved distal enhancer. Nucleic Acids Res 39 (2011) 1280–1293.

Kuzmichev AN, Kim SK, D'Alessio AC, Chenoweth JG, Wittko IM, Campanati L, McKay RD: Sox2 acts through Sox21 to regulate transcription in pluripotent and differentiated cells. Curr Biol (2012), 22(18):1705-10.

Lee TI, Jenner RG, Boyer LA, Guenther MG, Levine SS, Kumar RM, Chevalier B, Johnstone SE, Cole MF, Isono K, et al: Control of developmental regulators by Polycomb in human embryonic stem cells. Cell (2006), 125: 301–313.

Lefebvre V: The SoxD transcription factors-Sox5, Sox6, and Sox13-are key cell fate modulators. Int. J. Biochem. Cell Biol. 42 (3) (2010) 429–432.

Lefebvre V, Dumitriu B, Penzo-Mendez A, Han Y, Pallavi B: Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. Int J Biochem Cell Biol (2007), 39:2195–214.

Lengler J, Bittner T, Münster D, Gawad Ael-D, Graw J. Agonistic and antagonistic action of AP2, Msx2, Pax6, Prox1 AND Six3 in the regulation of Sox2 expression. Ophthalmic Res (2005), 37: 301-309.

Li H, Lu Y, Smith HK, Richardson WD: Olig1 and Sox10 interact synergistically to drive myelin basic protein transcription in oligodendrocytes. J Neurosci 27 (2007) 14375–14382.

Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, et al: Human DNA methylomes at base resolution show widespread epigenomic differences. Nature (2009), 462: 315–322.

Lodato MA, Ng CW, Wamstad JA, Cheng AW, Thai KK, Fraenkel E, Jaenisch R, Boyer LA: SOX2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specify cell state. PLoS Genet. 9 (2) (2013) e1003288.

Loh KM, Lim B: A precarious balance: pluripotency factors as lineage specifiers. Cell Stem Cell (2011), 8: 363-369.

Lopez-Anido C, Sun G, Koenning M, Srinivasan R, Hung HA, Emery B, Keles S, Svaren J: Differential Sox10 genomic occupancy in myelinating glia. Glia 63 (2015) 1897–1914.

López-Juárez A, Remaud S, Hassani Z, Jolivet P, Pierre Simons J, Sontag T, Yoshikawa K, Price J, Morvan-Dubois G, Demeneix BA: Thyroid hormone signaling acts as a neurogenic switch by repressing Sox2 in the adult neural stem cell niche. Cell Stem Cell (2012) 10, 531–543.

Lujan E, Chanda S, Ahlenius H, Südhof TC, Wernig M: Direct conversion of mouse fibroblasts to self-renewing, tripotent neural precursor cells. Proc Natl Acad Sci USA (2012) 109:2527–2532.

Malas S, Postlethwate M, Ekonomou A, Whalley B, Nishiguchi S, Wood H, Meldrum B, Constanti A, Episkopou V: SOX1-deficient mice suffer from epilepsy associated with abnormal ventral forebrain development and olfactory cortex hyperexitability. Neuroscience (2003) 119:421–432.

Mallanna SK, Ormsbee BD, Iacovino M, et al: Proteomic analysis of Sox2-associated proteins during early stages of mouse embryonic stem cell differentiation identifies Sox21 as a novel regulator of stem cell fate. Stem Cells (2010), 28:1715–1727.

Martini S, Bernoth K, Main H, Ortega GD, Lendahl U, Just U, Schwanbeck R: A critical role for Sox9 in notch-induced astrogliogenesis and stem cell maintenance. Stem Cells 31 (2013) 741–751.

Manuel MN, Martynoga B, Molinek MD, Quinn JC, Kroemmer C, Mason JO, Price DJ: The transcription factor Foxg1 regulates telencephalic progenitor proliferation cell autonomously, in part by controlling Pax6 expression levels. Neural Dev (2011) 18, 6:9.

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

Maucksch C, Jones KS, Connor B: Concise review: the involvement of SOX2 in direct reprogramming of induced neural stem/precursor cells. Stem Cells Transl Med (2013) 2(8): 579–583.

Mertin S, McDowall SG, Harley VR: The DNA-binding specificity of SOX9 and other SOX proteins. Nucleic Acids Res (1999), 27:1359–64.

Molofsky AV, Glasgow SM, Chaboub LS, Tsai HH, Murnen AT, Kelley KW, Fancy SP, Yuen TJ, Madireddy L, Baranzini S, Deneen B, Rowitch DH, Oldham MC: Expression profiling of Aldh1l1-precursors in the developing spinal cord reveals glial lineage-specific genes and direct Sox9-Nfe2l1 interactions. Glia 61 (2013) 1518–1532.

Mullen AC, Orlando DA, Newman JJ, Lovén J, Kumar RM, Bilodeau S, Reddy J, Guenther MG, DeKoter RP, Young RA: Master transcription factors determine cell-type-specific responses to TGF-b signaling. Cell (2011) 147, 565–576.

Najm FJ, Lager AM, Zaremba A, Wyatt K, Caprariello AV, Factor DC, Karl RT, Maeda T, Miller RH, Tesar PJ: Transcription factor-mediated reprogramming of fibroblasts to expandable, myelinogenic oligodendrocyte progenitor cells. Nat Biotechnol (2013), 31:426-433.

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

Nishimoto M, Fukushima A, Okuda A, Muramatsu M: 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) (1999) 5453–5465.

Niu W, Zang T, Zou Y, Fang S, Smith DK, Bachoo R, Zhang CL: In vivo reprogramming of astrocytes to neuroblasts in the adult brain. Nature cell biology (2013), 15(10):10.1038/ncb2843.

Niwa H, Toyooka Y, Shimosato D et al: Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. Cell (2005), 123:917–929.

Peirano RI, Goerich DE, Riethmacher D, Wegner M: Protein zero expression is regulated by the glial transcription factor Sox10. Mol Cell Biol (2000), 20:3198–209.

Peirano RI, Wegner M: The glial transcription factor Sox10 binds to DNA both as monomer and dimer with different functional consequences. Nucleic Acids Res (2000), 28:3047–55.

Pevny LH, Nicolis SK: Sox2 roles in neural stem cells. Int J Biochem Cell Biol (2010) 42, 421-424.

Prévostel C, Blache P: The dose-dependent effect of SOX9 and its incidence in colorectal cancer. Eur J Cancer (2017), 86:150-157.

Ring KL, Tong LM, Balestra ME, Javier R, Andrews-Zwilling Y, Li G, Walker D, Zhang WR, Kreitzer AC, Huang Y: Direct reprogramming of mouse and human fibroblasts into multipotent neural stem cells with a single factor. Cell Stem Cell (2012), 11:100-109.

Roadmap Epigenomics Consortium, et al.: Integrative analysis of 111 reference human epigenomes. Nature (2015) 518, 317-330.

Rodda DJ, Chew JL, Lim LH, Loh YH, Wang B, Ng HH, Robson P: Transcriptional regulation of nanog by OCT4 and SOX2. J. Biol. Chem. 280 (26) (2005) 24731–24737.

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

Scaffidi P, Bianchi ME: Spatially precise DNA bending is an essential activity of the sox2 transcription factor. J Biol Chem (2001), 276:47296-47302.

Schepers GE, Teasdale RD, Koopman P: Twenty pairs of sox: extent, homology, and nomenclature of the mouse and human sox transcription factor gene families. Dev Cell (2002) 3, 167–170.

Schlierf B, Ludwig A, Klenovsek K, Wegner M: Cooperative binding of Sox10 to DNA: requirements and consequences. Nucleic Acids Res (2002) 30:5509–16.

Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, Wei G, Zhao K: Dynamic regulation of nucleosome positioning in the human genome. Cell (2008) 132: 887–898.

Scott CE, Wynn SL, Sesay A, Cruz C, Cheung M, Gomez Gaviro MV, Booth S, Gao B, Cheah KS, Lovell-Badge R, Briscoe J: SOX9 induces and maintains neural stem cells. Nat Neurosci 13 (2010) 1181–1189.

Sinclair AH1, Berta P, Palmer MS, Hawkins JR, Griffiths BL, Smith MJ, Foster JW, Frischauf AM, Lovell-Badge R, Goodfellow PN: A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature (1990), 346(6281):240-4.

Smith CL, Peterson CL: ATP-dependent chromatin remodeling. Current Topics in Developmental Biology (2005) 65, 115–148.

Soufi A, Donahue G, Zaret KS: Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. Cell (2012) 151, 994-1004.

Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS: Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. Cell (2015) 161, 555-568.

Stolt CC, Lommes P, Friedrich RP, Wegner M: Transcription factors Sox8 and Sox10 perform non-equivalent roles during oligodendrocyte development despite functional redundancy. Development 131 (2004) 2349–2358.

Stolt CC, Lommes P, Hillgärtner S, Wegner M: The transcription factor Sox5 modulates Sox10 function during melanocyte development. Nucleic Acids Res (2008), 36(17): 5427–5440.

Stolt CC, Lommes P, Sock E, Chaboissier MC, Schedl A, Wegner M: TheSox9 transcription factor determines glial fate choice in the developing spinal cord. Genes Dev 17 (2003) 1677–1689.

Stolt CC, Rehberg S, Ader M, Lommes P, Riethmacher D, Schachner M, Bartsch U, Wegner M: Terminal differentiation of myelin-forming oligodendrocytes depends on the transcription factor Sox10. Genes Dev (2002), 16(2):165-70.

Stolt CC, Schmitt S, Lommes P, Sock E, Wegner M: Impact of transcription factor Sox8 on oligodendrocyte specification in the mouse embryonic spinal cord. Dev Biol 281 (2005) 323–331.

Takahashi K, Yamanaka S: Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell (2006), 126(4):663-76.

Tam WLODF, Hiramatsu K, Tsumaki N, Luyten FP, Roberts SJ: Sox9 reprogrammed dermal fibroblasts undergo hypertrophic differentiation in vitro and trigger endochondral ossification in vivo. Cell Reprogram (2014), 16:29-39.

Tao W & Lai E: Telencephalon-restricted expression of BF-1, a new member of the HNF-3/fork head gene family, in the developing rat brain. Neuron (1992) 8: 957–966.

Teif VB, Vainshtein Y, Caudron-Herger M, Mallm JP, Marth C, Hofer T, Rippe K: Genome-wide nucleosome positioning during embryonic stem cell development. Nat Struct Mol Biol (2012) 19: 1185–1192.

The ENCODE Project Consortium: A user's guide to the encyclopedia of DNA elements (ENCODE). PLoS Biol. 9 (2011), e1001046.

Tomioka M, Nishimoto M, Miyagi S, Katayanagi T, Fukui N, Niwa H, Muramatsu M, Okuda A: Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. Nucleic Acids Res. 30 (14) (2002) 3202–3213.

Thomson M, Liu SJ, Zou LN, Smith Z, Meissner A, Ramanathan S: Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. Cell (2011), 145: 875-889.

Thurman RE, Rynes E, Humbert R, Vierstra J, Maurano MT, Haugen E, Sheffield NC, Stergachis AB, Wang H, Vernot B, et al: The accessible chromatin landscape of the human genome. Nature (2012) 489: 75–82.

Tsuda M, Takahashi S, Takahashi Y, Asahara H: Transcriptional co-activators CREB binding protein and p300 regulate chondrocyte-specific gene expression via association with Sox9. J Biol Chem (2003), 278:27224–9.

Tsuruzoe S, Ishihara K, Uchimura Y, Watanabe S, Sekita Y, Aoto T, Saitoh H, Yuasa Y, Niwa H, Kawasuji M, et al: Inhibition of DNA binding of Sox2 by the SUMO conjugation. Biochem Biophys Res Commun (2006) 351, 920-926.

Uchikawa M, Kamachi Y, Kondoh H: Two distinct subgroups of Group B Sox genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken. Mech Dev 84 (1999), 103-120.

Valouev A, Johnson SM, Boyd SD, Smith CL, Fire AZ, Sidow A: Determinants of nucleosome organization in primary human cells. Nature (2011) 474: 516–520.

Van Hoof D, Munoz J, Braam SR, Pinkse MW, Linding R, Heck AJ, Mummery CL, Krijgsveld J: Phosphorylation dynamics during early differentiation of human embryonic stem cells. Cell Stem Cell (2009) 5, 214-226.

van Oevelen C, Collombet S, Vicent G, Hoogenkamp M, Lepoivre C, Badeaux A, Bussmann L, Sardina JL, Thieffry D, Beato M, et al.: C/EBPalpha activates pre-existing and de novo macrophage enhancers during induced pre-B cell transdifferentiation and myelopoiesis. Stem Cell Rep (2015) 5, 232-247.

Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, Plajzer-Frick I, Shoukry M, Wright C, Chen F, et al: ChIP-seq accurately predicts tissue-specific activity of enhancers. Nature (2009), 457: 854–858.

Wang Z, Oron E, Nelson B, Razis S, Ivanova N: Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. Cell Stem Cell (2012), 10: 440-454.

Wapinski OL, Vierbuchen T, Qu K, Lee QY, Chanda S, Fuentes DR, Giresi PG, Ng YH, Marro S, Neff NF, et al.: Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons. Cell (2013), 155: 621–635.

Wegner M: All purpose Sox: The many roles of Sox proteins in gene expression. Int J Biochem Cell Biol (2010), 42(3):381-90.

Werner MH, Huth JR, Gronenborn AM, Clore GM: Molecular basis of human 46X,Y sex reversal revealed from the three-dimensional solution structure of the human SRY-DNA complex. Cell 81 (5) (1995) 705–714.

Whittington N, Cunningham D, Le TK, De Maria D, Silva EM: Sox21 regulates the progression of neuronal differentiation in a dose-dependent manner. Dev Biol (2015), 397(2):237-47.

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

Xia P, Wang S, Ye B, Du Y, Huang G, Zhu P, Fan Z: Sox2 functions as a sequence-specific DNA sensor in neutrophils to initiate innate immunity against microbial infection. Nat Immunol (2015), 16(4):366-75.

Xu J, Pope SD, Jazirehi AR, Attema JL, Papathanasiou P, Watts JA, Zaret KS, Weissman IL, Smale ST: Pioneer factor interactions and unmethylated CpG dinucleotides mark silent tissue-specific enhancers in embryonic stem cells. Proc Natl Acad Sci (2007), 104: 12377–12382.

Xu J, Watts JA, Pope SD, Gadue P, Kamps M, Plath K, Zaret KS, Smale ST: Transcriptional competence and the active marking of tissue-specific enhancers by defined transcription factors in embryonic and induced pluripotent stem cells. Genes Dev (2009), 23: 2824–2838.

Yang N, Zuchero JB, Ahlenius H, Marro S, Ng YH, Vierbuchen T, Hawkins JS, Geissler R, Barres BA, Wernig M: Generation of oligodendroglial cells by direct lineage conversion. Nat Biotechnol (2013), 31:434-439.

Yuan HB, Corbi N, Basilico C, Dailey L: Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. Genes Dev (1995), 9:2635–45.

Zhang S, Cui W: Sox2, a key factor in the regulation of pluripotency and neural differentiation. World J Stem Cells (2014), 6(3): 305–311.

Zhao C, Deng Y, Liu L, Yu K, Zhang L, Wang H, He X, Wang J, Lu C, Wu LN, Weng Q, Mao M, Li J, vanEs JH, Xin M et al: Dual regulatory switch through interactions of Tcf7l2/Tcf4 with stage-specific partners propels oligodendroglial maturation. Nat Commun 7 (2016) 10883.

Zhao HY, Zhang YJ, Dai H, Zhang Y, Shen YF: CARM1 mediates modulation of Sox2. PLoS ONE (2011) 6, e27026.

Zhu Q, Zhao X, Zheng K, Li H, Huang H, Zhang Z, Mastracci T, Wegner M, Chen Y, Sussel L, Qiu M: Genetic evidence that Nkx2.2 and Pdgfra are major determinants of the timing of oligodendrocyte differentiation in the developing CNS. Development 141 (2014) 548–555.

Zou D, Erickson C, Kim EH, Jin D, Fritzsch B, Xu PX: Eya1 gene dosage critically affects the development of sensory epithelia in the mammalian inner ear. Hum Mol Genet (2008), 17(21):3340-56.