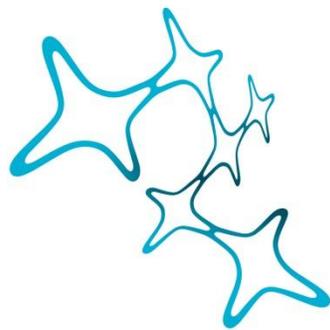

IDENTIFICATION AND MANIPULATION OF CHROMATIN BARRIERS IN TRANSCRIPTIONAL REPROGRAMMING

Valentin Claude Vincent Baumann



Graduate School of
Systemic Neurosciences
LMU Munich



Dissertation der
Graduate School of Systemic Neurosciences
der Ludwig-Maximilians-Universität München

März 2019

Supervisor:

Dr. Stefan H. Stricker

MCN Junior Research Group, Munich Center for Neurosciences, Ludwig-Maximilian-Universität,
BioMedical Center
Grosshaderner Strasse 9
82152 Planegg-Martinsried
Germany

Epigenetic Engineering, Institute of Stem Cell Research, Helmholtz Zentrum,
German Research Center for Environmental Health

First Reviewer: Dr. Stefan H. Stricker

Second Reviewer: Prof. Dr. Wolfgang Wurst

Third Reviewer: Dr. Jörg Betschinger

Date of Submission: 27.03.2019

Date of Defense: 03.07.2019

Abstract

Master transcription factors are cell fate determining factors that can force new identities onto cells by evoking distinct gene transcription patterns. This potential can be detrimental for an organism when activated erroneously, which is why genes of master transcription factors have to be regulated tightly. Chromatin features like DNA methylation have been implied in transcriptional regulation, the direct causalities and the underlying mechanisms are however still largely unclear.

Here I investigated the role of the neurogenic transcription factor Sex-determining-region-y-box 1 (Sox1) in directing neural stem cell identity. I employed a targeted trans-activating domain (dCas9-VP64) to induce Sox1 expression *in vitro* in neural progenitor cells (NPCs) and characterized the invoked phenotypic changes. Inducing Sox1 expression in NPCs restored their neuronal differentiation potential, and transcriptome analysis revealed a shift in cell identity towards neural stem cells (NSCs), underlining the role of Sox1 as cell fate determining factor.

Analysis on single cell basis however revealed that only a small subset of NPCs responded to the targeted gene induction with Sox1 upregulation. Using a GFP knock in as reporter, I separated responsive from unresponsive cells and investigated differences in chromatin features at the Sox1 promoter. I identified DNA methylation as a strong barrier against trans-activation by combining transcriptional engineering and epigenome editing via dCas9-Tet1. Furthermore, I found similar barriers at the promoters of other master transcription factor genes, including Oct4 and Nkx2-2.

Lastly, I employed a screening approach to identify potential regulatory regions distal of the Sox1 gene. By transducing NPCs with a gRNA library of high complexity, I was able to identify targeting sites for dCas9-VP64 in the locus of Sox1 that have the potency to induce gene transcription even outside of the promoter.

In conclusion, I have confirmed Sox1 as a neurogenic master transcription factor and identified mechanisms that control expression of this gene. These findings could serve to optimize future trans-activation approaches and underline the importance of chromatin features in the regulation of cell fate determining factors.

TABLE OF CONTENT

	Page
Abstract	4
1 Introduction	10
1.1 Cell types and cellular identity	10
1.1.1 Master transcription factors.....	11
1.1.2 Reprogramming factors	12
1.2 Cell fate determinants in neural development.....	13
1.3 The master transcription factor Sox1	14
1.4 Epigenetic gene regulation	17
1.4.1 DNA modifications	17
1.4.2 Histone modifications.....	18
1.4.3 Other chromatin features	19
1.4.3.1 DNA Topology.....	19
1.4.3.2 RNA modification and non-coding RNAs	20
1.5 Epigenetic engineering	21
1.5.1 dCas9	21
1.5.2 Transcription activating factors	23
1.5.3 Targeted epigenetic modifiers.....	24
1.5.4 gRNA systems	26
1.5.4.1 gRNA design	26
1.5.4.2 gRNA multiplexing.....	26
1.5.4.3 dCas9 screens	27
1.5.5 Barriers to dCas9	28

1.6 Aim of study.....	29
2 MATERIAL AND METHODS.....	30
2.1 Materials.....	30
2.1.1 Chemicals.....	30
2.1.2 Cell Culture Media and Supplements.....	31
2.1.3 Kits.....	32
2.1.4 Cell lines and bacterial strains.....	33
2.1.5 Plasmids.....	33
2.1.6 Primers.....	33
2.1.7 Antibodies.....	37
2.1.8 gRNA Sequences.....	38
2.1.9 Buffers.....	39
2.1.9.1 Western Blot.....	39
2.1.9.2 Bacterial culture.....	41
2.1.9.3 Other.....	41
2.1.10 Software and statistics.....	41
2.2 Methods.....	42
2.2.1 Plasmid generation.....	42
2.2.1.1 dCas9-Effectors.....	42
2.2.1.2 gRNA design and plasmid generation.....	43
2.2.1.3 Gel electrophoresis and DNA isolation.....	44
2.2.1.4 Transformation, plasmid isolation, and Sanger Sequencing.....	45
2.2.2 Cell culture.....	45
2.2.2.1 Generation and cultivation of murine Sox1-GFP cells.....	45

2.2.2.2	Transfection	46
2.2.2.3	Flow cytometry and fluorescence activated cell sorting (FACS)	46
2.2.3	Lentiviral work	46
2.2.3.1	Lentivirus generation	46
2.2.3.2	Titer measurement	47
2.2.3.3	Establishment of stable cell lines	48
2.2.3.4	Transduction of gRNA lentivirus	48
2.2.4	Molecular methods	49
2.2.4.1	Reverse transcription and qPCR	49
2.2.4.2	Western Blotting	49
2.2.4.3	Differentiation assay	50
2.2.4.4	Immunofluorescence staining and microscopic analysis ...	51
2.2.4.5	Chromatin Immunoprecipitation (ChIP)-qPCR	51
2.2.5	Sequencing	52
2.2.5.1	Bisulfite and oxidative bisulfite sequencing	52
2.2.5.2	RNA sequencing	54
2.2.5.3	Sequencing of gRNA amplicons	54
3	Results	56
3.1	Generation and characterization of dCas9-expressing cell lines	56
3.2	Targeted activation of Sox1	57
3.2.1	Sox1 upregulation following VP64 targeting	57
3.2.2	Characterization of Sox1 positive cells	63
3.3	Investigation of technical barriers to targeted gene activation	68

3.4 Investigation of chromatin barriers to targeted gene activation.....	72
3.4.1 Characterization of chromatin features at the Sox1 locus.....	72
3.4.2 Combining transcriptional editing with epigenome engineering enhances efficiency of gene induction	76
3.4.3 DNA methylation as barrier to transcriptional engineering is not exclusive to Sox1	79
3.5 Identification of regulatory domains at the Sox1 locus.....	82
3.5.1 VP64 Screen: Identification of distal regulatory elements	82
3.5.2 Future Experiments.....	88
4 Discussion.....	89
4.1 Sox1 is instructive for neuroepithelial cell identity.....	89
4.2 DNA methylation serves as transcriptional barrier at several master transcription factors	92
4.3 Additional epigenetic barriers against activation of transcription	96
4.4 The chromatin model of epigenetic gene regulation	98
4.5 Implications for the dCas9 tool	100
4.6 Conclusion.....	102
Bibliography	104
Appendix: Chromatin modifier screen gRNA enrichment	117
List of tables	118
List of Figures.....	119
List of abbreviations	120
List of Publications	123
Copyright information	124

Acknowledgements	125
Eidesstattliche Versicherung/Affidavit	126
Declaration of author contributions.....	127

1 Introduction

1.1 Cell types and cellular identity

Since their discovery as the smallest structural unit of a living organism in 1824, it has become clear that cells not only carry out a function as frameworks for organs, but also conduct physiologic processes (Harris 1999). In highly complex multicellular organisms like the human body, which consists of approximately 10^{13} cells (Bianconi et al. 2013), they vary strongly in their characteristics, enabling their classification in different cell types that are primarily defined by morphology, functionality, or potential to generate cells of other types. During development, these different types are all derived from a single, totipotent stem cell (Mitalipov and Wolf 2009), a circumstance that was first postulated by Conrad Waddington in the Waddington landscape (Waddington 2012) (Figure 1): This schematic depicts the process of cellular differentiation as a marble rolling down a hill towards terminally differentiated cells. The different paths that are available become more and more restricted along the way, representing a gradual commitment of the stem cell to first a lineage, and ultimately a cell type. While first it was postulated that this loss of potency might originate from a loss of genetic information in differentiating cells, John Gurdon showed in 1958 that indeed even terminally differentiated cells contain the same genetic material as their ancestry (Gurdon, Elsdale, and Fischberg 1958). By now we know that different cell types make use of different parts of the same rather than possessing different genetic material, i.e. they exhibit differential transcription patterns (Arendt et al. 2016).

During development, the activation of certain genes, so called cell fate determining factors, is sufficient to trigger different transcriptional programs that direct a stem cell towards a certain lineage and subsequently to a specified terminally differentiated cell (Smith, Sindhu, and Meissner 2016). The expression of distinct genetic networks in a cell in turn stabilizes and protects its identity. Collectively, this concept of changing transcriptional programs that guide the increasing specialization of a cell during development reflects the formation of the valleys and hills in the Waddington landscape, and how they define the available paths for the marble. The

identification of cell types based on uniquely expressed marker genes (Arendt et al. 2016, Regev et al. 2017) can thus improve the definition of the ever increasing number of cell types and subtypes being discovered in the human body (Vickaryous and Hall 2006).

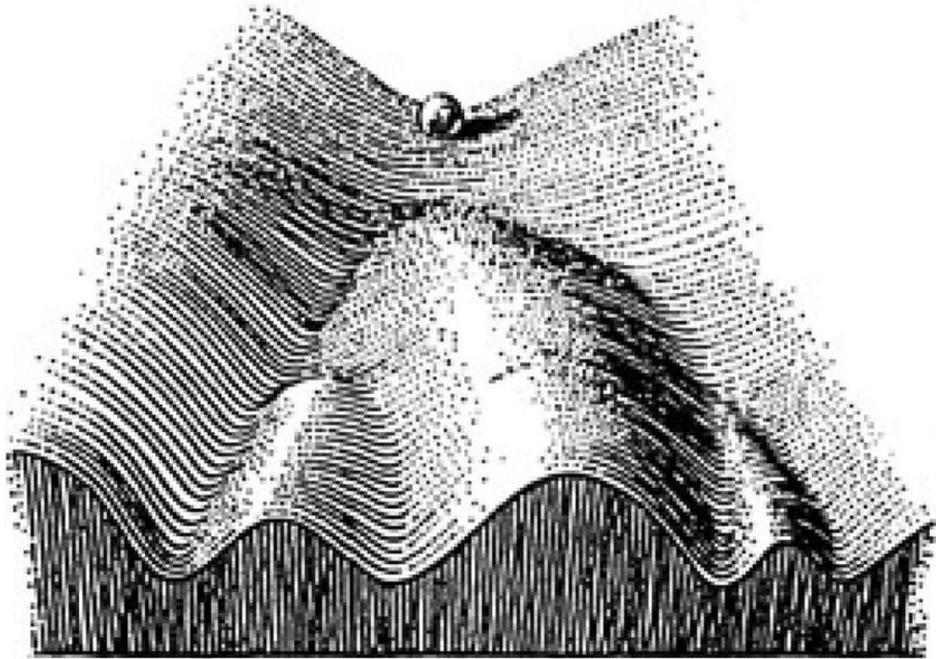


Figure 1: Waddington's epigenetic landscape. The schematic depicts a pluripotent stem cell at the beginning of development. Preferential transcriptional states are depicted as valley. These guide cells during differentiation to specific cell fates, while transcriptional barriers, depicted as hills, interfere with spontaneous change to a cell identity from a different lineage. From Bard, 2008 (permission received from Springer Nature).

1.1.1 Master transcription factors

While genes of various gene families have been described to be of considerable importance for cell fate choices during reprogramming, transcription factors are clearly the most abundant in this context. These are proteins that can induce gene expression by interaction with the DNA (Spitz and Furlong 2012). The locus of their binding is defined by short sequences, so called binding motifs that are often highly specific for certain transcription factors (Whyte et al. 2013).

Individual transcription factors that trigger cell type specific programs and therefore direct and protect cell identities, are called master transcription factors, and because of their potency to direct cellular fate are highly specific in their expression pattern (Vaquerizas et al. 2009). Derivation from this pattern can be highly detrimental to multicellular organisms. For example, erroneous activation of a master transcription factor of a different lineage in a terminally differentiated cell can overwrite existing transcriptional patterns with a new ones, thereby changing the cell fate determining features of this cell and defining a different cell type (Iwafuchi-Doi and Zaret 2014).

1.1.2 Reprogramming factors

The ability of some master transcription factors to induce a new cell type outside of their natural context even in terminally differentiated cells can be exploited in a process called cellular reprogramming. The first direct reprogramming of one somatic cell type into another was performed 1987 in fibroblasts (Davis, Weintraub, and Lassar 1987). When MyoD, a basic helix loop helix transcription factor was ectopically expressed in these cells, they changed their identity to myocytes. By now, this concept has been widely adapted to generate cells of various lineages from fibroblasts (Guo and Morris 2017). Of note, Yamanaka and colleagues managed to generate induced pluripotent stem cells by combining four reprogramming factors, Oct4, Klf4, Sox2, and c-Myc (OKSM, (Takahashi and Yamanaka 2006)).

Interestingly, once acquired, the new identity is stable even upon removal of the reprogramming factor(s). This distinguishes them from terminal selector genes that can define a cell type while transcribed; upon loss of their expression, a cell however loses the existing identity and switches to a new one in an undirected manner. This indicates that reprogramming factors fundamentally change transcriptional patterns and instruct changes even to the cell fate determining factors (Hobert 2008).

It should be noted that under physiological conditions, such switches in cell identity occur only rarely. It is therefore likely that master transcription factors with such potency are tightly regulated in their expression in order to safeguard this specificity.

1.2 Cell fate determinants in neural development

The development of the mammalian brain is a complex process that starts as early as gastrulation at embryonic day 6.5 (E6.5) in mice (third gestational week in humans) (Stiles and Jernigan 2010). During gastrulation, the embryonic structure changes essentially – while hypoblasts, cells from the lower layer of the embryo form extra-embryonic tissues, epiblasts, cells from the upper layer, differentiate into the three germ layers of the embryo (Stiles and Jernigan 2010). During this process, a part of the epiblasts migrate through the primitive streak and give later rise to the endoderm and mesoderm, while cells that stay in the former epidermal layer give rise to the two parts of the ectoderm – the epidermal ectoderm and the neurectoderm (Stiles and Jernigan 2010). Cells of the neurectoderm form the neural plate at E7.5 and are the first appearance of neural stem cells (NSCs). During the next major step of brain development, the neural plate folds and forms the neural tube, a hollow structure filled with cerebrospinal fluid (CSF) and lined by a monolayer of NSCs. While initially cylindrical, the hollow center of the neural tube will eventually give rise to the ventricular system of the brain, which is why the bordering region containing the NSCs is called ventricular zone. Because NSCs form the epithelium of the neural tube, they are also called neuroepithelial cells (NECs) at this developmental stage. Just before neural tube closure at E8.0, its anterior end expands and forms the three primary brain vesicles, precursors of the three major brain regions forebrain, midbrain, and hindbrain (Stiles and Jernigan 2010). To expand the stem cell pool, NECs undergo a number of symmetric divisions, each producing two new NECs. At the onset of neurogenesis (E9.5 in spinal cord, shortly after in the CNS), NECs turn into apical radial glia cells (aRGCs) that produce the first output of neurons in asymmetric divisions (producing one aRGC and one neuron, or neuronal progenitor; (Jiang and Nardelli 2016). At birth, neurogenesis, as well as formation of the general architecture of brain structures is largely complete; there are however NSCs that continue to give rise to newborn neurons throughout life.

1.3 The master transcription factor Sox1

It is obvious that a vastly complex process like neural development has to be tightly regulated and that disruption of its control mechanisms can lead to tremendously detrimental consequences, and ultimately dysfunction of the CNS (Jiang and Nardelli 2016). Indeed, neural development is controlled by a cascade of master transcription factors. The earliest marker of NSCs, and an essential regulating factor in this sequence is the sex-determining-region-y-box transcription factor 1 (Sox1) (Wood and Episkopou 1999). This transcription factor is already expressed during the neural plate stage, as early as E7.5. During development it is specifically expressed in neural tissue, with the lens as only exception. Here, Sox1 protein can be detected starting mid-gestation, and throughout embryonic development. In neural tissue, Sox1 expression is lost when NSCs exit the neuronal lineage to become glial progenitors (precursors of astrocytes and oligodendrocytes), and when neuronal progenitors exit mitosis during differentiation towards mature neurons, while it is continuously expressed in the proliferative ventricular zone (Pevny et al. 1998); see Figure 2). Even in the adult brain, Sox1 expression can be found, albeit in a very limited number of NSCs, and only in regions of ongoing neurogenesis (i.e. subventricular zone and dentate gyrus), where it marks progenitor cells with long-term neurogenic potential (Venere et al. 2012).

Sox1 is a member of the SoxB1 family of transcription factors (along with Sox2 and Sox3), and shares ca. 50% amino acid sequence identity with these factors. Furthermore, Sox1 can replace Sox2 in inducing pluripotency (Nakagawa et al. 2008) and it is a strong read-out for efficiency in the reprogramming of astrocytes to induced pluripotent stem cells (iPSCs) (Nakajima-Koyama et al. 2015). It is however not expressed in PSCs or extraembryonic tissue like Sox2, and the phenotype of Sox1 knock-out mice underlines the hypothesis that these factors are indeed not redundant: while Sox2 expression can compensate the lack of Sox1 during embryonic development, it later on leads to an epileptic phenotype and is eventually lethal. Of note, the placeholder model indeed suggests that these two members of the SoxB1 subfamily in some cases bind subsequently to the same site: during pluripotency Sox2 is bound and keeps the chromatin accessible, however without inducing gene transcription of

targets; when cells exit pluripotency and enter the neural lineage, Sox2 is then replaced by Sox1 binding, which in turn is necessary for the propagation of NSCs (Buecker and Wysocka 2012). *In vitro* the knock-down of Sox1 in NSCs has been shown to lead to deficiency to produce neurons, and overexpression *in vivo* and *in vitro* leads to increased neuronal output (Kan et al. 2004, Kan et al. 2007). This has been linked not only to the promotion of neuronal differentiation in NSCs, but also to stimulation of proliferation (Kan et al. 2007, Venere et al. 2012). Furthermore, it has been shown that only those astrocytes that upregulate Sox1 transiently during reprogramming with only two factors, namely Oct4 and Klf4, are able to give rise to pluripotent colonies (Nakajima-Koyama et al. 2015). These studies strongly suggest a function for Sox1 as lineage specifying master transcription factor; however its exact function is still poorly understood (especially compared to Sox2 and Sox3; (Julian, McDonald, and Stanford 2017)).

The expression pattern of Sox1 *in vivo* during development can be paralleled *in vitro*; in 2003, Aubert et al. isolated embryonic stem cells (ESCs) from transgenic mice carrying a heterozygous GFP knock-in in the open reading frame (ORF) of Sox1. While ESCs do not express Sox1, cells start to be strongly Sox1^{GFP} positive when neural differentiation is induced (Aubert et al. 2003). This is accompanied by the formation of neural rosettes (NRs), the *in vitro* correlate of NSCs of the neural tube. At this stage, Sox1^{GFP} positive cells can give rise to cells of the neuronal and the glial lineage, when further differentiated. However, if they are kept under proliferative conditions they change their morphology significantly and lose Sox1^{GFP} expression. In line with knock-down experiments, these Sox1^{GFP} negative cells also lose their neurogenic differentiation potential. Due to the limited potential of these cells compared to NSCs with strong Sox1 expression, these cells are termed neural progenitor cells (NPCs) (Pollard et al. 2006).

These findings underline the importance of Sox1 during neural development and suggest its role as neurogenic master transcription factor. In turn, it is very probable that its gene is subject to strict regulatory mechanisms to prevent erroneous transcription, a hypothesis further corroborated by its highly specific expression patterns during neural development.

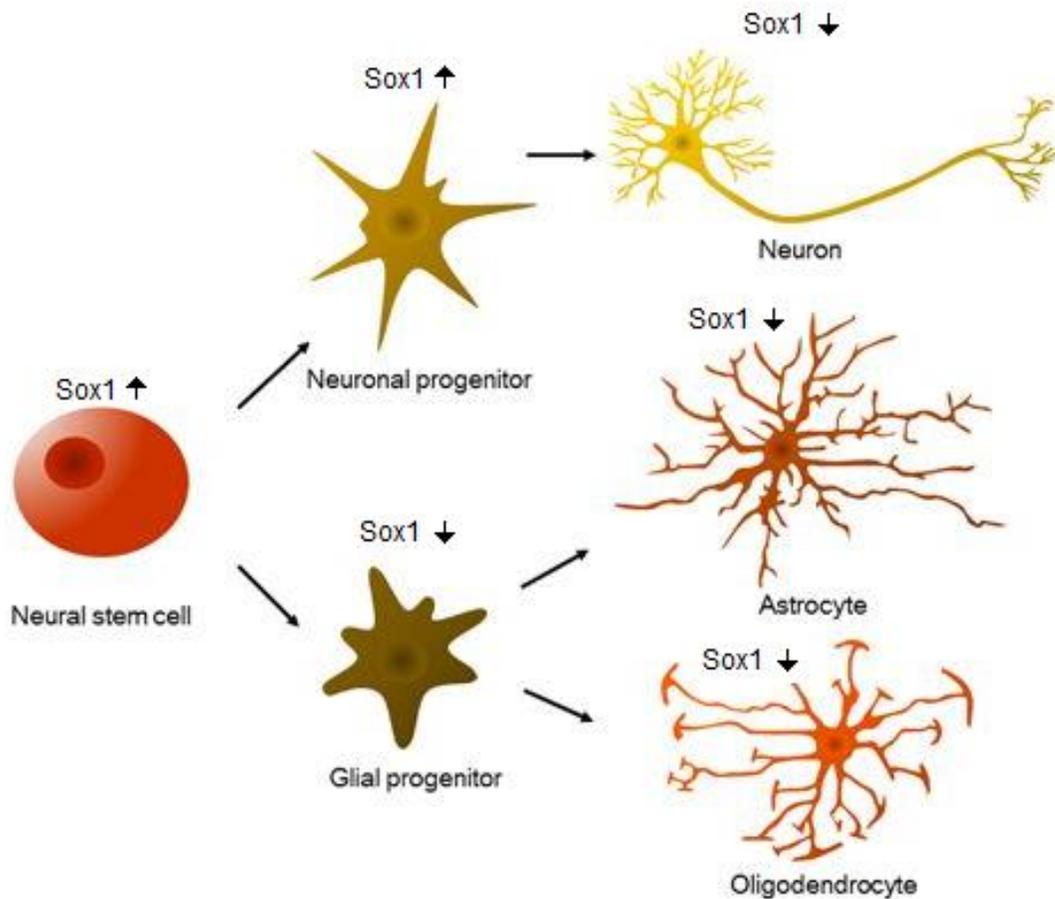


Figure 2: Expression of Sox1 during neural development. Sox1 is the earliest marker of NSCs, already expressed in cells of the neural plate. Sox1 is downregulated when cells exit the neural lineage to become progenitors of glial cells, or when neural progenitors terminally differentiate into post-mitotic neurons. Modified from Tang et al., 2017 (open access).

1.4 Epigenetic gene regulation

A possible mechanism that has often been implied in gene regulation, especially in the case of tightly regulated genes, is the chromatin complex and its modifications (Stricker, Köferle, and Beck 2017, Bultmann and Stricker 2018). Chromatin is a multimolecular complex in the nuclei of eukaryotic cells that consists of DNA and proteins. Its components can be modified to form chromatin features, some of which have been shown to correlate to the transcriptional status of genes. This postulates a suppositive link between the genetic code and the cell type defining transcriptional patterns that form the Waddington epigenetic landscape (Waddington 2012) (Figure 1). Despite many studies that show a correlation between certain chromatin features and gene regulation, it has proven difficult to establish a definitive causal role for such features in the induction of phenotypic outcomes (Bultmann and Stricker 2018).

1.4.1 DNA modifications

The most prevalent modification of DNA is methylation of the Carbon atom at position 5 of the cytosine base (5-methylcytosine, 5mC). It was also the first chromatin feature to be discovered (Hotchkiss 1948). Methylation of cytosine is conducted by DNA methyltransferases (Dnmts), which either set the mark on unmodified DNA strands (Dnmt1) or methylate the newly generated strand during DNA replication (Dnmt3) (Smith et al. 1992, Kho et al. 1998). Methylation of cytosine is correlated to repressed transcription, when present in the promoter of a gene. However, more complex correlations have been suggested for other genomic contexts (Colot and Rossignol 1999, Grosjean 2009).

Demethylation can occur indirectly, through oxidation of the methyl group. This is facilitated over several steps by enzymes of the ten-eleven translocation family (Tet1, 2, and 3, (Ito et al. 2011). While the first product of oxidation, hydroxymethylation (resulting in 5 hydroxymethylcytosine, 5hmC) is far more dynamic and quickly further processed, it has been reported to carry out its own regulatory functions, as it has been linked to impaired self-renewal in ESCs upon removal (Freudenberg et al. 2012, Teif et al. 2014). Further oxidation of the hydroxyl-residue leads to 5-formylcytosine (5fC) and 5-

carboxycytosine (5caC). These modifications are however only transient and unmodified cytosine is then restored by base excision repair (BER, (He et al. 2011)). Demethylation of the unoxidized methyl group can be passive (by DNA replication in the absence of Dnmt1) or facilitated through deamination, followed by BER (He et al. 2011).

Furthermore, methylation of other DNA bases, specifically adenine has been reported (m6A, (Luo et al. 2015))

1.4.2 Histone modifications

More complex than modifications of the DNA are those of histones. Histones are DNA binding proteins that form the structural core of chromatin. While H1 is a linker histone and binds DNA by itself, the other histones form octamers of H2A, H2B, H3, and H4 (each octamer consists of 2 of each histone). These octamers bind 147bp of DNA to form together the nucleosomes, the main way of a cell to store DNA (Lowary and Widom 1997). Of each histone, there are genetic variants with distinct functional roles, even though they are structurally very similar (Khare et al. 2012). Furthermore, the proteins can be modified at in total 130 amino acid residues, mostly located in the N-terminal tail (Tessarz and Kouzarides 2014). Currently, 12 different modifications are known, that carry different regulatory function, in part dependent on the modified residue (Tan et al. 2011). Even though we still lack clear understanding of all modifications and their function, some features have been shown to clearly correlate to certain states of gene transcription. Above all, acetylation at different residues, and methylation of the lysine 4 residue of Histone 3 (H3K4) are connected to active genes, while methylation of lysine 9 and lysine 27 of histone 3 (H3K9 and H3K27 respectively) are primarily located at repressed genes (Zhou, Goren, and Bernstein 2011). The vast amount of potential combinations of modifications and modified residues enables highly differential and specific gene regulation. It is therefore clear, that these chromatin features themselves underlie tight regulation by histone modifying enzymes (Zhou, Goren, and Bernstein 2011). A variety of such enzymes has been discovered to date, and the most common ones are listed in Table 1.

Function	Location	Enzyme	Source
Methylation	H3R2	Carm1	(Chen et al. 1999)
	H3K4	Set1A	(Schneider, Bannister, and Kouzarides 2002)
		Set1B	(Schneider, Bannister, and Kouzarides 2002)
		Set7	(Wang et al. 2001)
	H3K9	Suv39	(Jenuwein et al. 1998)
		G9a	(Tachibana et al. 2001)
	H3K27	Ezh2	(Laible et al. 1997)
H3K36	Set2	(Strahl et al. 2002)	
Demethylation	H3K4	Lsd1	(Shi et al. 2004)
	H3K9	Jmjd2a	(Huang et al. 2006)
		Jmjd2b	(Fodor et al. 2006)
		Jmjd2c	(Loh et al. 2007)
	H3K36	Jhdm1	(Tsukada et al. 2006)
		Jhdm1b	(Tsukada et al. 2006)
Gasc1		(Cloos et al. 2006)	
Acetylation	ubiquitous	p300	(Eckner et al. 1994)

Table 1: Common Histone modifying enzymes

1.4.3 Other chromatin features

1.4.3.1 DNA Topology

Apart from modifications on the two major parts of the chromatin complex, several other features play a role in gene regulation. One important factor is the DNA topology, the three dimensional structure and compaction that DNA assumes in the nucleus. The compaction can be visualized even by unspecific DNA stains like Dapi, and can impact gene regulation by tuning accessibility for e.g. transcription factors.

Regions of high compaction, so called heterochromatin, are generally less accessible and therefore correlate to low transcriptional activity, and are often marked by trimethylation of H3K9 and H3K27. In contrast, euchromatin specifies regions of low compaction usually associated to high activity of the genes located in these areas, and is marked by H3K4 tri-methylation and histone acetylation (Dame 2005, Bernstein et al. 2006).

Furthermore, compaction can bring distal DNA regions to spatial proximity. By that, interactions between regulatory elements can be created even when these elements are thousands of basepairs apart. These are called topologically associated domains (TADs (Pombo and Dillon 2015)), and new methods have arisen that can map these interactions on different scales (3C, 4C, HiC; (de Wit and de Laat 2012)). Already, proteins have been discovered that regulate DNA structure by maintaining and modifying interactions (e.g. Ctf and Cohesin, (Bell and Felsenfeld 2000, Hark et al. 2000, Merkschlager and Nora 2016)).

1.4.3.2 RNA modification and non-coding RNAs

Over 100 different modifications of different bases of RNA have been discovered (Cantara et al. 2011), of which at least 12 are also present in eukaryotic cells (Li and Mason 2014). For some of these modifications, functional roles have been implicated (Roundtree et al. 2017). For example, methylation of adenosine has been shown to regulate RNA stability (Mauer et al. 2017). Many techniques for characterization of various RNA modifications have been published recently (Helm and Motorin 2017), and new techniques are constantly being developed.

Apart from post-transcriptional modifications on the RNA itself that regulate e.g. its stability, RNA itself can bear a regulatory function. This is specifically the case for non-coding RNAs (ncRNA). These RNAs are not translated into proteins, but regulate transcription and translation of coding genes via e.g. interfering with polymerases or ribosomes (Lee 2012). The length of such RNAs varies strongly from very short (miRNAs, siRNAs) to long (lncRNAs, macro ncRNAs) RNA molecules (Guenzl and Barlow 2012).

1.5 Epigenetic engineering

The steadily increasing number of methods for the discovery and mapping of chromatin features, and their constant improvements has led to a plethora of data on such features. But even though they have been intensively examined, most studies are purely descriptive, generating many credent correlations between specific modifications and e.g. gene transcription. These correlations are however no clarification as to whether chromatin features play causal roles in defining a cellular identity. Furthermore, even when a correlation for a specific type of modification has been shown, it still remains unclear whether a function can be directed by a single mark, or only by a combination.

Recent advances in DNA binding proteins have made it possible to tackle these problems and investigate the causality of chromatin modifications with very high resolution and specificity. Particularly, these methods allow for the induction of endogenous transcription factors, which in turn enables the investigation of their behavior in the physiologic chromatin context (in contrast to previous methods like ectopic expression), and for the manipulation of chromatin features to directly infer their functional meaning in different contexts.

1.5.1 dCas9

The first DNA targeting systems were based on zinc finger proteins (ZFP), a class of transcription factors accounting for almost half of all mammalian transcription factors (Vaquerizas et al. 2009, Wolfe, Nekludova, and Pabo 2000). These proteins all have in common a DNA binding domain that recognizes distinct DNA motives of usually three basepairs (Wolfe, Nekludova, and Pabo 2000). Synthetic design of proteins combining several of these domains results in a DNA binding domain recognizing a unique DNA sequence of variable length (e.g. combination of 10 zinc finger domains would lead to a recognition sequence of 30 basepairs) (Kim, Lee, and Carroll 2010). A similar approach was later taken to develop TALE (Transcription Activation-Like Effector) domains, synthetic proteins derived from bacterial transcription factors (Reyon et al. 2013). While the length of DNA binding domains used in TALEs is similar to that of

ZNFs, the former only recognize one specific DNA base. By combining these domains, it is in principle possible to target any desirable DNA sequence, however the resulting proteins are far bigger than ZNFs. In addition, for each new target, a new protein has to be designed and synthesized when using ZNFs or TALEs, as the target site is encoded in their amino acid sequence. The discovery of RNA guided DNA targeting systems overcame this problem by replacing DNA binding based on protein-DNA-interaction with DNA-RNA hybridization.

The most commonly used system based on RNA guidance is the CRISPR/Cas9 (Clustered Regularly Interspaced Palindromic Repeats/CRISPR-associated proteins). Originally evolved as adaptive immune response against invading phages in bacteria and archaea, this system has been modified as a tool for genomic engineering by fusing the endogenous targeting RNAs into one single guide RNA (gRNA) (Jinek et al. 2012). The scaffold sequence of this RNA molecule binds the Cas9 protein, while the preceding 17-20bp sequence, called protospacer, defines the target DNA sequence. If the DNA features a specific three basepair sequence following the protospacer sequence of a gRNA (5'-NGG-3' in case of Cas9, but specific for each CRISPR/Cas variant), the gRNA/dCas9 complex binds to the target site. This sequence is called protospacer adjacent (PAM) motive and is compulsory for DNA binding (see Figure 3) (Sternberg et al. 2014). In its unmodified form, Cas9 then introduces a double strand break in the target sequence, enabling different approaches to genetic engineering. By exchanging one amino acid in each of the two nuclease domains of Cas9, Perez-Pinera et al. developed a nuclease deficient version of the protein (dCas9) in 2013 (see Figure 3) (Perez-Pinera et al. 2013). Fusion of different effector domains to this system facilitates directed transcriptional activation, repression, or even targeted manipulation of chromatin modifications.

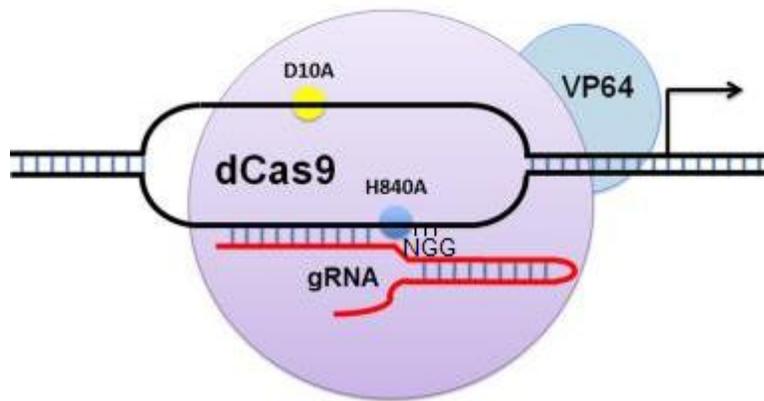


Figure 3: The dCas9/gRNA complex. A complex formed by dCas9 fused to the transcriptional activator VP64 and a gRNA is bound to the DNA double strand at the target site specified by the gRNA. The DNA target is followed by a PAM (NGG). Mutations D10A and H840A delete the nuclease activity. Modified from Perez-Pinera et al., 2013 (permission received from Springer Nature).

1.5.2 Transcription activating factors

By fusing trans-activating domains to the dCas9 protein, and targeting the system to gene promoters via specific gRNAs, strong induction of endogenous genes can be achieved. The majority of these so called transcription activating factors (TAFs) is based on the viral trans-activating domain VP16, that can be combined with itself or other activator domains to form potent tools for transcriptional induction. To this point the most commonly used TAF is dCas9-VP64, four copies of VP16 fused to the dCas9 targeting system (Maeder et al. 2013). Recently, even stronger versions have been developed, by increasing the number of VP16 copies (VP128, 8 copies (Li et al. 2017); VP160, 10 copies (Perrin, Rousseau, and Tremblay 2017); coupling VP64 to both ends of dCas9, butterfly dCas9 (Black et al. 2016)), or fusing additional trans-activators to VP64. By adding the p65 subunit of NF- κ B and the R trans-activator of the Epstein-Barr virus, Chavez et al. developed VPR, a new TAF with up 120-fold increased efficiency in gene induction, compared to VP64 alone (Chavez et al. 2015).

While the efficiency in gene induction of such TAFs differs strongly, some rules for the design of gRNAs to achieve the maximum possible effect are common to all of them:

- a) TAFs recruit the endogenous transcriptional machinery and therefore must be targeted to the promoter region of the target gene, or regions that are in its structural vicinity (see TADs, chapter 1.3.3.1). It has been shown that in general the efficiency of a TAF is highest when targeted to within 500 basepairs of the transcription start site (TSS) of a gene (Wang, La Russa, and Qi 2016). Binding of the dCas9-complex within a gene transcript can however interfere with the transcriptional machinery and by that even repress the target (CRISPRi, (Qi et al. 2013));
- b) Like the combination of several trans-activating domains into the same TAF, targeting the same promoter with several TAFs can severely enhance the effect on transcription. In general, a combination of two or three gRNAs is the most potent (Maeder et al. 2013, Chavez et al. 2015, Cheng et al. 2013). It is however important to design the gRNAs in a way that they do not inhibit each other, and a minimum distance of around 100 basepairs between each gRNA has been reported to be highly beneficial (Wang, La Russa, and Qi 2016);

1.5.3 Targeted epigenetic modifiers

The fusion of dCas9 and chromatin modifying enzymes enables the direct investigation of chromatin features and their functional relevance. By manipulating specific types of chromatin modifications in a defined region, the function of these modifications can be derived from the impact on a given phenotype of interest. This in turn could allow answering not only the question of cause or consequence of chromatin marks, but also of the relevance of their location for their functional impact.

To that end, several different enzymes have been fused to the targeting system and tested towards their potency to modify chromatin marks. In first efforts, catalytic domains of histone methyl transferases were tethered to dCas9, including PRDM9 (methylates H3K4 (Cano-Rodriguez et al. 2016)), G9a (H3K9 (O'Geen et al. 2017)), and EZH2 (H3K27 (O'Geen et al. 2017)). These can methylate histones at specific lysine

residues, to varying degrees of efficiency. In a similar way, targeted demethylation was achieved with histone demethylases (Kearns et al. 2015). Interestingly, when successful, the manipulation of chromatin marks in promoter or enhancer regions were sufficient to significantly change mRNA levels of related genes. Methylation of H3K9 or demethylation of H3K4 significantly reduced gene transcription (Kearns et al. 2015, Braun et al. 2017), while the inverse manipulation led to significant increase in gene transcription (Cano-Rodriguez et al. 2016).

Similar advances have been accomplished in the field of DNA-methylation. Several DNA-methyltransferases have been fused to dCas9, either as their full length version or catalytic domains, and achieved methylation of DNA up to 100% (Amabile et al. 2016, Vojta et al. 2016, Anton and Bultmann 2017, Ziller et al. 2018) with subsequent repression of transcription. Likewise, targeted DNA demethylation by dCas9-Tet1 has been employed successfully, and was in some cases reported to be accompanied by changes in gene transcription (Choudhury et al. 2016, Anton and Bultmann 2017, Gallego-Bartolomé et al. 2018) and even changes in a proliferation phenotype (Choudhury et al. 2016, Morita et al. 2016).

The most widely-used epigenetic modifier to date is dCas9-p300. This enzyme transfers acetyl-residues to the histone tail of H3, a chromatin feature strongly correlated with active genes (Eckner et al. 1994). In line with this function, a significant increase of mRNA was observed upon targeting of p300 to gene promoters (Hu et al. 2014). It is however not precisely clear to which amount the observed effect was facilitated by chromatin acetylation, because besides its modifying function, p300 also recruits several factors of the endogenous transcriptional machinery in a cell, a catalytic mutant of p300 is still lacking.

Of note, catalytic mutants have been developed for several other chromatin modifiers, which demonstrated the necessity of the chromatin manipulation to exert the observed transcriptional changes, proving a causal role of the chromatin marks therein (Hilton et al. 2015, Pflueger et al. 2018).

1.5.4 gRNA systems

1.5.4.1 gRNA design

Apart from the general rules for the layout of gRNA target sequences for a maximum effect in gene induction (discussed in 1.4.2), some major points have to be considered when designing gRNAs. For one, off-target binding is a common threat to the potency of targeting systems in general, and dCas9 is no exception. It has been shown that especially mismatches in the 5'-region of the protospacer do not interfere strongly with gRNA binding. Therefore, it is crucial to choose a targeting sequence that is unique for at least 15 3'-nucleotides, as mismatches in that part are detrimental to binding efficiency (Fu et al. 2013, Cho et al. 2014).

Furthermore, it has been suggested that certain motives and the presence of particular nucleotides can modulate the affinity of the gRNA/dCas9 complex to the DNA. Several algorithms have been developed that employ and unify these findings into easy to handle pipelines for gRNA design (Doench et al. 2014).

1.5.4.2 gRNA multiplexing

As indicated in chapter 1.4.2, many dCas9 approaches benefit from the combination of different gRNAs to increase the efficiency of the intended effect. The need to deliver multiple gRNAs is even more pronounced, when more than one target needs to be modified in an experimental setup (e.g. trans-activation of reprogramming factors). This can be readily achieved by co-transfection of gRNA expressing vectors, an approach which however bears significant draw-backs, as it cannot be controlled whether a cell received all relevant gRNAs (Weber et al. 2015, Maresch et al. 2016). The more gRNAs needed, the lower the efficiency of this strategy and it soon becomes unfeasible. This problem has been addressed by gRNA multiplexing, strategies that combine several different gRNA expressing cassettes into one plasmid, allowing efficient and homogenous delivery. Various such methods have been published, with a maximum of 14 gRNAs per plasmid (Kabadi et al. 2014, Peterson et al. 2016, Vazquez-Vilar et al. 2016, Breunig et al. 2018).

1.5.4.3 dCas9 screens

The bipartite nature of the dCas9 system makes it especially feasible for screening approaches that require the targeting of a vast number of sequences simultaneously in different cells. This is due to the fact that the target is solely defined by the gRNA – these are then usually delivered by lentiviruses, leading to stable integration of the gRNA cassette into the genome, and enabling easy identification of target sequences in hit cells by next generation sequencing (NGS) (Hartenian and Doench 2015). The variable nature of the targeted Cas9 protein (as wtCas9 or dCas9 fused to different effectors) allows for a broad variety of different screens, such as genetic or trans-activator screens.

Depending on the desired library complexity, different gRNA libraries can be employed. In arrayed libraries, each gRNA is present in a different well of a plate (usually 96-well or 384-well format). This allows for the investigation of complex phenotypes, since neither separation of hits from non-hits, nor retrieval of the gRNA sequence from hits is needed. The number of different gRNAs is however severely limited (Agrotis and Ketteler 2015). To circumvent this problem, pooled libraries can be used. In this case, gRNAs are synthesized together in one tube and delivered as a pool. This allows for much higher complexity of the library. On the downside, hits need to be identified from the pool of cells and gRNAs retrieved from the hits. This limits the possible phenotypes to screen for (Shang et al. 2017). By switching from synthesis of gRNAs to less-defined sources, e.g. genomic DNA, for the production of target sequences, the complexity can be further increased. Such libraries however contain many non-functional gRNAs (due to lack of PAM sites, length etc.) and are far less manageable during preparation and screening due to their complexity (Lane et al. 2015, Köferle et al. 2016).

1.5.5 Barriers to dCas9

The efficiency of DNA binding proteins is subject to chromatin features at their target site. This is apparent for endogenous proteins like transcription factors, but also applies to targeting systems like ZNFs, TALEs, and dCas9. Especially chromatin compaction has been investigated in this context, as the spatial requirements for dCas9 can be quite substantial (depending on the fused effector). It is an ongoing debate in the field, to which extent closed chromatin interferes with dCas9 binding. It has been shown to favor sites of open chromatin over those with high nucleosome occupancy (Hinz, Laughery, and Wyrick 2015); (Kuscu et al. 2014, O'Geen et al. 2015). It is however also clear that even though reduced, binding in heterochromatin is still possible to efficient levels (Perez-Pinera et al. 2013), and even a pioneering factor character has been suggested for dCas9 (Polstein et al. 2015).

Once bound, it is widely assumed that artificial trans-activators induce transcription independently of the chromatin landscape at the target locus, and indeed many genes have been efficiently upregulated by various dCas9-TAFs in methodological publications (Chavez et al. 2015). Targets used in such publications (e.g. *Actc1*, *Ttn*) however have a far less decisive effect on cell identity than reprogramming factors. Of note, in approaches where the gene of a reprogramming factor was induced by a dCas9 targeted trans-activator, the efficiency of reprogramming was often lower than by ectopic overexpression of the gene (Weltner et al. 2018, Liu et al. 2018). This hints towards barriers that are present specifically at the endogenous locus of a reprogramming gene, where they potentially interfere with the function of trans-activators and reduce their efficiency. While this observation has not been investigated in detail as of yet, it would fit to the fundamental of a cell to protect its identity by tightly regulating cell fate determining genes (see chapter 1.1).

1.6 Aim of study

Sox1 is the first specific marker of cells of the neural lineage and a potential master transcription factor. It exhibits a highly specified expression pattern and its transcription is tightly regulated throughout development, as its erroneous expression could have devastating consequences on neural development and subsequently lead to dysfunctions of the CNS. Although chromatin features such as histone and DNA modifications have been implicated in this regulatory process, it remains still largely unknown, a) which marks b) in which regions contribute to such barriers, and how they do so. Furthermore, the precise function of Sox1 is still largely understudied despite its fundamental role in neural development, especially compared to the other members of the SoxB1 family of transcription factors (Sox2 and Sox3).

Here I aimed firstly to investigate the potency of cells for neurogenic differentiation upon Sox1 induction. It has been shown *in vivo* and *in vitro* that the neurogenic potential of neural stem or progenitor cells correlates strongly with Sox1 expression. I therefore investigated whether induction of the neurogenic master transcription factor in NPCs using dCas9-VP64 would restore their neuronal differentiation potential and a phenotype resembling that of NSCs.

Secondly, I investigated chromatin features that act as barriers against induction of Sox1 transcription by identifying chromatin marks with different levels in Sox1 expressing versus non-expressing NPCs, and combining Sox1 induction via dCas9-VP64 with targeted removal of these marks.

Lastly, I investigated the chromatin landscape surrounding the Sox1 gene locus in large scale screening approaches, employing pooled gRNA libraries in combination with different chromatin modifying enzymes tethered to dCas9, in order to identify candidate distal regulatory elements.

2 MATERIAL AND METHODS

2.1 Materials

2.1.1 Chemicals

Chemicals that were routinely used are listed in Table M1.

Name	Catalogue #	Manufacturer
0.2µm Polyvinylidene membrane	LC2002	Thermo Fisher Scientific
10% APS	A3678-25G	Sigma
16% Formaldehyde (w/v), Methanol-free	10321714	Thermo Fisher Scientific
5x Laemmli Buffer		
6x DNA loading dye	R0611	Thermo Fisher Scientific
Agarose	870055	Biozym, Oldendorf
Aqua poly mount	18606	Polysciences
Bovine Serum Albumine	A2153-1KG	Sigma-Aldrich
DAPI Nuclear Staining Dye	1351303	Bio-Rad Laboratories
ECL Luminol Reagent	sc-2048	Santa Cruz Biotechnology
Ethanol, 99.9%	9065.2	Carl Roth
GeneRuler 1kb DNA ladder	SM0313	Thermo Fisher Scientific
Gibson Assembly Master Mix	E2611	NEB
HotStar Taq	203203	Qiagen
Methanol	34860-1L-R	Sigma-Aldrich
Methanol-free Formaldehyde	28906	ThermoFisher Scientific
Milk powder, blotting grade	T145.1	Carl Roth
Nuclease-free water	AM9932	Life Technologies
Paraformaldehyde	158127-5G	Sigma-Aldrich
Poly-D-Lysine	P6407-5MG	Sigma-Aldrich
Phusion DNA Polymerase Master Mix	M0531S	NEB
Potassium Perurethate	10378-50-4	Sigma-Aldrich
Protease Inhibitor Cocktail	P8340	Sigma-Aldrich

2 MATERIAL AND METHODS

PowerUp™ Sybr Green Master Mix	A25742	ThermoFisher Scientific
RIPA buffer	R0278	Sigma-Aldrich
Sodium Hydroxyde, 50% in H ₂ O	1310-73-2	Sigma-Aldrich
SYBR Safe DNA Gel Stain	5001208	Life Technologies
TEMED	T9281	Sigma
Triton-X 100	T8655.1	Biomol
Bovine serum albumin	A9418	Sigma-Aldrich

Table M1: Chemicals and reagents

2.1.2 Cell Culture Media and Supplements

Frequently used cell culture reagents are listed in Table M2.

Name	Catalogue #	Manufacturer
Accutase	A6964-100ML	Life Technologies
BDNF	PHC7074	Gibco
bFGF	78003	Stemcell Technologies
Blasticidin S	R21001	ThermoFisher Scientific
Bovine Serum Albumin	10773877	Thermo Fisher Scientific
cAMP	A6885	Sigma-Aldrich
Cryotubes	10577391	Thermo Fisher Scientific
DMSO	D5879-100ML	Sigma-Aldrich
eBioscience™ Fixable Viability dye eFluor™ 660	65-0864-14	ThermoFisher Scientific
EGF	78006	Stemcell Technologies
Fetal calf serum	C8056-500ML	Sigma-Aldrich
GDNF	PHC7041	Gibco
Hygromycin B	10687010	ThermoFisher Scientific
Laminin	L2020	Sigma
Lipofectamin® 2000	11668027	ThermoFisher Scientific

2 MATERIAL AND METHODS

NT3	PHC7036	ThermoFisher Scientific
NeuroCult™ Proliferation Kit	05702	Stemcell Technologies
Puromycin	A1113803	ThermoFisher Scientific

Table M2: Cell culture media and supplements

2.1.3 Kits

Frequently used kits are listed in Table M3.

Name	Catalogue #	Manufacturer
Agencourt AMPure XP magnetic beads	A63881	Beckman Coulter
Auto iDeal ChIP-Seq Kit for Histones	C01010171	Diagenode
Bradford Assay kit	5000201	Bio-Rad
Extractme total RNA kit	EM09.1	DNAGDANSK
EZ DNA Methylation Gold Kit	D5005	Zymo
Maxima first strand cDNA Synthesis kit	K1671	Thermo Fisher Scientific
MicroPlex Library Preparation Kit v2	C05010012	Diagenode
Micro-Bio-Spin P-6 SCC columns	7326201	Bio-Rad
Nextera DNA Library Preparation Kit	FC-121-1012	Illumina
PicoPure™ RNA Isolation Kit	KIT0204	Thermo Fisher Scientific
PureLink™ HiPure Plasmid Filter Maxiprep Kit	K210027	Invitrogen
QIAprep Spin Miniprep Kit	27104	Qiagen
Quant-iT™ PicoGreen™ dsDNA assay Kit	P7589	Thermo Fisher Scientific
SMART-Seq™ v4 Ultra™	634894	TaKaRa
RNeasy Mini Kit	74104	Qiagen

Table M3: Kits

2.1.4 Cell lines and bacterial strains

Embryonic stem cells containing a heterozygous GFP knock in in the open reading frame of Sox1 (Sox1^{wt/GFP}) were obtained from the Austin Smith lab and differentiated to neural stem cells by Dr. Stefan Stricker as described previously (Ying et al. 2003). Neural stem cells were sorted as described (2.2.2.3) to obtain neural rosettes, or differentiated further to obtain neural progenitor cells. Clonal NPC lines stably expressing dCas9-fusion proteins or gRNAs were established as described below (2.2.3.3).

2.1.5 Plasmids

Plasmid	Addgene #
PMLM3705	47754
pLKO.1	10878
P300 template plasmid	23252
Set7 template plasmid	24082
JMJD2a template plasmid	38846
Tet1 catalytic domain template plasmid	39454
Plenty-dCas9-VP64-Blast	61425

Table M4: Plasmids used in this study

2.1.6 Primers

Primers used for this study were designed using Primer3Plus (<http://primer3plus.com/cgi-bin/dev/primer3plus.cgi>) or MethPrimer 2.0 (Urogene, <http://www.urogene.org/cgi-bin/methprimer2/MethPrimer.cgi>, for Bisulfite Sequencing primers), and purchased from Metabion. Further specifications for the design of qPCR and bisulfate sequencing primers are described in the respective chapters (2.2.4.1 and 2.2.5.1) Primer sequences are listed in TableM5.

2 MATERIAL AND METHODS

Primer name	Sequence	Tm[°C]
Hygro_1fwd	AGTCAATAATCAATGTCAACCGGGTAGGGGAGGC G	57
Hygro_1rev	GGTGGGCGAAGAACTCTCGGCATCTACTCTATTC CTTTG	57
Hygro_2fwd	GAATAGAGTAGATGCCGAGAGTTCTTCGCCACC CC	57
Hygro_2rev	AAGTGCCACCTGACGTGACGGGTATACAGACAT GATAAGATACATTGATGA	57
P300_fwd	CGATGACAAGGCTGCAGGAGGCGGAGGTAGCAA AGAAAATAAGTTTTCTGCTAAAAGG	54
P300_rev	GCTGATCAGCGGGTTTTTCAGCATTTCATTGCAGGT GTAGACAAA	54
Set7_fwd	CGATGACAAGGCTGCAGGAGGCGGAGGTAGCTT CTTCTTTGATGGCAGCACC	58
Set7_rev	GCTGATCAGCGGGTTTTCACTTTTGCTGGGTGGC C	58
Tet1_fwd	CGATGACAAGGCTGCAGGAGGCGGAGGTAGCGA ACTGCCACCTGCAGCTG	61
Tet1_int_rev	GGCAGTGACGAAGGCTTACT	61
Tet1_int_fwd	AGTAAGCCTTCGTCACTGC	61
Tet1_rev	GCTGATCAGCGGGTTTTTCAGACCCAATGGTTATA GG	61
JMJD2a_fwd	CGATGACAAGGCTGCAGGAGGCGGAGGTAGCGC TTCTGAGTCTGAAACTCTGAATCC	59
JMJD2a_rev	GCTGATCAGCGGGTTTTTCATGCTTCTGGCGTGGG CAG	59
Tet1_mut_fwd	GTGCTCATCCCTACAGGGCCATTCACAACAT	61
Tet1_mut_rev	ATGTTGTGAATGGCCCTGTAGGGATGAGCAC	61
dCas9-lenti-T2A-puro_fwd	ATTCAGGTGTCGTGACGTACGGCCACCATGGAT AAAAAGTATTCTATTGGTTTAG	62

2 MATERIAL AND METHODS

dCas9-lenti-T2A- puro_rev	GCCCTCTCCACTGCCTGTACAGTTAATTAACATAT CGAGATCGAAATCG	62
amp_gRNA_fwd	CCATTTCGATTAGTGAACGGATC	62
amp_gRNA_rev	CGACTCGGTGCCACTTTTTTC	62
libgen_fwd	CTTGTGGAAAGGACGAAACA	63
libgen_rev	GCCTTATTTTAACTTGCTATTTCTAGC	63
Sox1_BiS_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGT TTTGGGTTTTTAATTTAAT	52
Sox1_BiS_rev	GCTGATCAGCGGGTTTTTCACTTTTGCTGGGTGGC C	52
Actcl_BiS_fwd	CGATGACAAGGCTGCAGGAGGCGGAGGTAGCGA ACTGCCACCTGCAGCTG	52
Actcl_BiS_rev	GGCAGTGACGAAGGCTTACT	52
Sox1_qPCR_fwd	AGACAGCGTGCCTTTGATTT	60
Sox1_qPCR_rev	TGGGATAAGACCTGGGTGAG	60
eGFP_qPCR_fwd	GAAGCAGCACGACTTCTTCAA	60
eGFP_qPCR_rev	AAGTCGATGCCCTTCAGCTC	60
Gapdh_qPCR_fwd	TTGCAGTGGCAAAGTGGAGA	60
Gapdh_qPCR_rev	CGTTGAATTTGCCGTGAGTG	60
Actcl_qPCR_fwd	ATGTGTGACGACGAGGAGAC	60
Actcl_qPCR_rev	CGGACAATTTACGTTTCAGCA	60
Actc1_ChIP_fwd	GGCCATATAGGGAGCTAGGG	60
Actc1_ChIP_rev	AGAGCAATAAGCCCACTCCA	60
Gapdh_ChIP_fwd	ACCAGGGAGGGCTGCAGTCC	60
Gapdh_ChIP_rev	TCAGTTCGGAGCCCACACGC	60
Oct4_ChIP_fwd	CCCCAGGGAGGTTGAGAGTT	60
Oct4_ChIP_rev	AAGGGCTAGGACGAGAGGGA	60
Sox1_ChIP_fwd	GCTGAGCTGAGTGCAAAGTG	60
Sox1_ChIP_rev	CCCTGGGTCGTGTTTAAATG	60
Cas9_qPCR_fwd	TCGTAGGGACCGCACTCATT	60
Cas9_qPCR_rev	TCGCTTTTCGCGATCATCTT	60

2 MATERIAL AND METHODS

Actc1_ChIP_fwd	GGCCATATAGGGAGCTAGGG	60
Actc1_ChIP_rev	AGAGCAATAAGCCCACTCCA	60
Oct4_BiS_1_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGG GAGTGGTTTTAGAAATAATTGG	52
Oct4_BiS_1_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG CACCCCTACCTTAAATCAC	52
Oct4_BiS_2_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGG GTGAGAGGATTTTGAAGGT	52
Oct4_BiS_2_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG AAAAACAAAACCTATAAAAATAAAAA	52
Nkx2-2_BiS_1_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGT TTTTAGAGTAAGATGAGAGGTG	52
Nkx2-2_BiS_1_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG ATATTAATAAAAAATTCTTTACCCCC	52
Nkx2-2BiS_2_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGA TAGAAAGGAGGGGGTAAAGAATTT	52
Nkx2-2_BiS_2_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG TAAATCTTATTTAAAAAACCAACCA	52
NeuroD4_BiS_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGG GAGGGGTTATTTTGTGGGTA	52
NeuroD4_BiS_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG CACTACCAAATAACCTTCATATCAATAC	52
Ngn2_BiS_1_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGT AATGAGTTGTTGAAAGGGAG	59
Ngn2_BiS_1_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG CAACTAACCAATCAATATTCC	59
Ngn2_BiS_2_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGT GATTAGATAAAGGGGGGA	59
Ngn2_BiS_2_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG ACCCCTCCTCACCTACCCTT	59
Il1rn_BiS_1_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGG	52

2 MATERIAL AND METHODS

	GTTTTAGGGTAGAGGTTAGTAAA	
Il1rn_BiS_1_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG	52
	AAACAATAAAACCTAATAAACAAAA	
Il1rn_BiS_2_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGG	52
	GATTTGTTATGTAAATGAGGGAG	
Il1rn_BiS_2_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG	52
	CACCAACCTATACTACTATCATTC	
MyoD_BiS_fwd	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGT	59
	TGAGGTTAGTATAGGTTGGAGGAG	
MyoD_BiS_rev	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG	59
	TATTTATCCAAAATAACCTAAAAACC	

Table M5: Primers used in this study

2.1.7 Antibodies

Primary and secondary antibodies for immunocytochemistry (ICC) and Western Blot, as well as ChIP Antibodies that were used in this study are listed in Table M6.

Antigen	Catalogue #	Application	Dilution	Company
a-Tubulin	T5168	Western Blot	1:2000	Sigma
Actcl	66125-1-1G	ICC	1:250	Proteintech
Calbindin	AB1778	ICC	1:1000	Merck/Millipore
Cas9	NBP2-36440	Western Blot	1:500	Novus
CD133	141201	ICC	1:300	Biologend
E-Cadherin	14-3249-82	ICC	1:300	Invitrogen
Flag-M2	F3165	ChIP	1:1000	Sigma
GFAP	G3893	ICC	1:500	Sigma
H3	ab1791	ChIP	1:1000	Abcam
H3K27me3	C15410195	ChIP	1:1000	Diagenode
H3K9me3	ab8898	ChIP	1:500	Abcam
IgG	C15400001-15	ChIP	1:1000	Diagenode

2 MATERIAL AND METHODS

Il1rn	HPA001482	ICC	1:500	Sigma-Aldrich
Map2	MAB378	ICC	1:300	Merck/Millipore
MyoD	MA5-12902	ICC	1:250	Invitrogen
Nestin	MAB5326	ICC	1:500	Millipore
NeuroD4	NBP2-13932	ICC	1:500	Novus
Ngn2	MAB3314	ICC	1:1000	R&D Systems
Nkx2-2	MAB8162	ICC	1:500	R&D Systems
Notch1	MA5-11961	ICC	1:500	Invitrogen
Occludin	NBP1-87402	ICC	1:250	Novus
Pou5F1	Ab181557	ICC	1:300	Abcam
Sox1	ab87775	Western Blot	1:1000	Abcam
Sox1	ab87775	ICC	1:400	Abcam
Tuj1	T8660	ICC	1:1000	Sigma
vGlut1	AB5905	ICC	1:300	Millipore
Zo-1	Sc-10804	ICC	1:500	Santa-Cruz

Table M6: Antibodies used throughout this study

2.1.8 gRNA Sequences

gRNA Sequences employed for promoter targeting are listed in Table M8.

gRNA name	gRNA construct	Sequence
Actcl_1	Actcl Stagr	GGCTCCAAGAATGGCCTCAG
Actcl_2	Actcl Stagr	GGGAGGGGCGAGGCCAGCAAG
Il1rn_1	Il1rn- Stagr	AGGAGCTGGATTTATGAGGT
Il1rn_2	Il1rn- Stagr	ACAAGTGCTCTCACAAACAC
MyoD1_1	MyoD1- Stagr	CCACCAATCACAATAGACAG
MyoD1_2	MyoD1- Stagr	ACTGACCTCAGTGCTCACTT
NeuroD4	NeuroD4- Stagr	TGACTCTACTACCCTACATG
Ngn2_1	Ngn2- Stagr	CCACCAATCACAATAGACAG
Ngn2_2	Ngn2- Stagr	AGTGTTCCCGGGACTCCGGG

2 MATERIAL AND METHODS

Nkx2-2_1	Nkx2-2- Stagr	GCCCTCTAGAGCAAGATGAG
Nkx2-2_2	Nkx2-2- Stagr	TCCTTTGTATGTAAATACTG
Oct4_1	Oct4- Stagr	GGATGTTTTAGGCTCTCCAG
Oct4_2	Oct4- Stagr	CATCGCACCAAAAGCCTGT
Sox1_1	S1-9 Stagr	GTTAATCATTCGGAGCGCGC
Sox1_9	S1-9 Stagr	GCGCGGGCGGGCGGAGCAAGG
Sox1_4	S4-7 Stagr	GAGGCAAAGGGGGCGAGCTC
Sox1_7	S4-7 Stagr	GGGGGGGAACAAGGGCAGGA
Sox1_1	SoxProm	GTTAATCATTCGGAGCGCGC
Sox1_2	SoxProm	GCGGGCGGGGAGAGGCAAAG
Sox1_3	SoxProm	GCGCGGGCGGGGAGAGGCAA
Sox1_4	SoxProm	GAGGCAAAGGGGGCGAGCTC
Sox1_5	SoxProm	GCCGCCGCGCGCGCGCGCTC
Sox1_7	SoxProm	GGGGGGGAACAAGGGCAGGA
Sox1_9	SoxProm	GCGCGGGCGGGCGGAGCAAGG

Table M7: gRNA Sequences for promoter targeting

2.1.9 Buffers

2.1.9.1 Western Blot

4 X Tris/SDS pH 6.8

60.5 g/l Tris base (Roth, 4855.2)

4g/l SDS (Roth, 0183.3)

adjust pH to 6.8

2 MATERIAL AND METHODS

4 X Tris/SDS pH 8.8

182g/l Tris base

4g/l SDS

adjust pH to 8.8

Blotting Buffer 10X

144,2g/l Glycin (Roth, 3790.3)

30.2g/l Tris base

2g/l SDS

PAGE Running buffer 10X

144,2g/l Glycin

30.2g/l Tris base

10g/l SDS

TBS 10X PH 7,4

24.22g/l Tris base

87.66g/l NaCl (Roth, P029.2)

TBST 10X PH 7,4

24.22g/l Tris base

87.66g/l NaCl

10ml/l Tween20 (Roth, 9127.1)

2.1.9.2 Bacterial culture

LB Medium

16g/l Trypon/Pepto from Casein (Roth, 8952.2)

5g/l yeast extract (Roth, 2363.2)

10g/l NaCl (Roth, P029.2)

LB Agar

15g/l agar (Roth, 5210.2) in LB-medium

2.1.9.3 Other

TAE 50X

242g/l Tris base

18.61g/l EDTA (Roth, 8043.3)

57.1 Glacial Acetic Acid (Roth, 6755.1)

2.1.10 Software and statistics

Software packages used for data analysis and gRNA design are referenced in the respective chapters. PCA plots, heatmaps, gRNA enrichment plots and hierarchical clustering plots were generated in RStudio v1.1.442 using the pheatmap (version 1.0.12), RColorBrewer (version 1.1-2), and ggplot2 (version 3.1.0) packages. All other graphs were generated in Microsoft® Office Excel® 2007 (12.0.6787.5000) SP3 MSO (12.0.6785.5000).

Values are depicted as mean±standard deviation and sample sizes are specified in the respective figures. For unpaired comparison of means, Mann-Whitney-U Test was

performed to test for statistically significant differences. Where $n > 10$, Shapiro-Wilk-Test was used to check for normal distribution of data, and if given, Student t-test was applied.

2.2 Methods

2.2.1 Plasmid generation

2.2.1.1 dCas9-Effectors

For the generation of dCas9 expression vectors, a hygromycin cassette was cloned into the pMLM3705 plasmid (Addgene plasmid 47754). The backbone was digested with restriction enzymes SgrDI (Thermo Scientific ER2031) and MluI (Thermo Scientific ER0561). The resistance cassette was amplified from Addgene plasmid 41721 with the primers Hygro_1fwd and Hygro_1rev. A SV40 polyadenylation sequence was amplified from Addgene plasmid 13820 with SV40_1fwd and SV40_1rev using the Phusion DNA Polymerase Master Mix according to the manufacturer's protocol. For Gibson assembly, inserts and backbone were combined in a molar ratio of 3:3:1, and Gibson Mastermix (NEB E2611S) was added. The reaction mix was incubated for 30 minutes at 50°C to generate dCas9-VP64-Hygro. For the exchange of VP64 with chromatin modifying domains, this plasmid was digested with PstI (NEB R3140S) and PmeI (NEB R0560S). Templates and primers used for the PCR of insert sequences are listed in Table M4 and M5 respectively. PCR was performed using the Phusion DNA Polymerase Master Mix according to the manufacturer's protocol. Gibson reaction was performed as described above with molar ratios of 3:1 (insert:backbone) to produce dCas9-Tet1-Hygro, dCas9-P300-Hygro, dCas9-JMJD2a-Hygro, and dCas9-Set7-Hygro respectively. Mutagenesis PCR was performed on dCas9-Tet1-Hygro to generate the catalytic mutant of Tet1 (Tahiliani et al. 2009) using the Phusion DNA Polymerase Master Mix according to the manufacturer's protocol. To generate lentiviral constructs, plenti-dCas9-VP64-Blast (Addgene Plasmid 61425) was digested with BsiWI (NEB R3553S) and BsrGI (NEB R0575S). The dCas9-fusion domains were amplified from the expression plasmids with the primers dCas9-lenti_fwd and dCas9-lenti_rev using the

Phusion DNA Polymerase Master Mix according to the manufacturer's protocol. Inserts and backbone were mixed in a ratio of 3:1 and Gibson assembly was performed as described above to produce the respective dCas9-Effector-T2A_Blast plasmids. Primers are listed in Table M5.

2.2.1.2 gRNA design and plasmid generation

Promoter gRNAs were designed manually. As target sequence, 250bp upstream of the transcription start site (TSS) of the targeted gene to the TSS were selected. To fit the requirements of the human U6 promoter on the gRNA expression plasmids, only gRNA Sequences with a 5'G were considered. To avoid spatial interference of targeted proteins, sequence pairs that were used for simultaneous targeting were selected to be at least 100pb apart. Lentiviral gRNA plasmids were generated by modification of the pLKO.1 plasmid as described in Kofler et al., 2016. Single gRNAs were ordered as strings (Table M7) with a 5' and a 3' overhang for Gibson cloning (see below) and amplified with libgen_fwd and libge_rev using the Phusion DNA Polymerase Master Mix according to the manufacturer's protocol. The modified pLKO.1 backbone was digested with AgeI and combined with the insert at a molar ratio of 1:3. Gibson assembly was performed as described above.

Overhangs:

5'overhang: TCTTGTGGAAAGGACGAAACACC

3'overhang: GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT

STAgR cloning was performed according to the protocol published by Breunig et al., 2018 in order to combine two gRNAs for simultaneous targeting in one vector. For subcloning of these constructs into the modified lentiviral gRNA backbone, the gRNA expression cassettes were amplified using amp_gRNA_fwd and amp_gRNA_rev and the Phusion DNA Polymerase Master Mix according to the manufacturer's protocol. The modified pLKO.1backbone was digested with AgeI (NEB M0531S) and combined with the insert at a molar ratio of 1:3. Gibson assembly was performed as described above.

2 MATERIAL AND METHODS

For the design of the Sox20000 gRNA library, the genomic locus chr8:12,295,501-12,495,500 of the mouse genome mm10 was obtained from the UCSC genome browser (genome.ucsc.edu/cgi-bin/hgTracks?db=mm10). Potential gRNA sequences of the form GN₂₀GG (first G necessary as transcription start nucleotide for the human U6 promoter, last NGG is the PAM sequence) were extracted from both DNA strands using RStudio v1.1.442 and the Biostrings (version 2.50.2) package. Bowtie aligner (Bowtie 2.3.4.3, John Hopkins University) was used to identify all gRNAs with a unique target sequence in the mouse genome. To increase specificity, 5 5'-nucleotides were cropped for this analysis. This left a total of ca. 6000 gRNA sequences. In addition, 150 gRNAs were generated using the rice genome (RGAP 7) that do not bind in mm10, 100 random gRNAs that bind in the promoter region of Igf2R and Airn each, and 150 gRNAs in the Sox1 locus, that do not have a PAM. These served as negative control gRNAs, and were generated similarly.

gRNA sequences were ordered with a 5' and a 3' overhang for Gibson cloning as a pool from arbor biosciences and cloned in a pool as described above for single gRNAs. The pool was transformed in 50 reactions into 50µl chemocompetent 5-alpha *E.coli* (NEB, C2987) each according to the manufacturer's protocol, and plated on a 25cmx25cm LB Agar plate (SigmaAldrich, D8679-1CS). After incubation over night at 37°C, the bacteria lawn was washed off with LB medium and collected. The plasmid DNA from the collected bacterial suspension was extracted using the PureLink HiPure Plasmid Filter Kit according to the manufacturer's protocol. Plasmids were resuspended in sterile H₂O and DNA concentration measured at a Nanodrop ND-1000 (NanoDrop Technologies, Inc.).

2.2.1.3 Gel electrophoresis and DNA isolation

PCR products for cloning steps were analyzed on agarose gels to confirm specific amplification. DNA Gel loading dye was added to PCR products in a proportion of 1:5, and DNA was loaded on 1% agarose gels (in TAE buffer, containing 1:20000 Sybr[®] Safe DNA Gel Stain). DNA was separated at 120V and visualized on a ChemiDoc[™] MP System (Bio-Rad). If specific amplification was confirmed, DNA was

purified using Agencourt AMPure XP magnetic beads according to the manufacturer's protocol.

2.2.1.4 Transformation, plasmid isolation, and Sanger Sequencing

Chemocompetent 5-alpha *E.coli* (NEB, C2987) were transformed with DNA according to the manufacturer's protocol. 50µl transformed bacteria were plated on an LB agar plate supplemented with 50µg/ml ampicillin and incubated at 37°C over night. Single colonies were resuspended in 2ml LB medium containing 50µg/ml ampicillin and incubated over night at 37°C and 200rpm. Plasmid DNA was isolated using the QIAprep Spin Miniprep Kit according to the manufacturer's protocol. For quality control, plasmids were sequenced at Eurofins Genomics (Ebersberg, Germany). Cultures of validated plasmids were inoculated in 200ml LB medium supplemented with 50µg/ml ampicillin and incubated over night at 37°C and 200rpm. Plasmids were isolated using the PureLink HiPure Plasmid Filter Kit according to the manufacturer's protocol. Plasmids were resuspended in sterile H₂O and DNA concentration measured at a Nanodrop ND-1000 (NanoDrop Technologies, Inc.). DNA was stored at -20°C.

2.2.2 Cell culture

2.2.2.1 Generation and cultivation of murine Sox1-GFP cells

Neural progenitor cells (NPCs), neural stem cells (NSCs), and neural rosettes (NRs) were derived from embryonic stem cells with a heterozygous GFP knock-in in the open reading frame of Sox1 (Figure 5), Sox1^{wt/GFP} cells, gifted from Austin Smith lab) as described (Ying et al. 2003). NPCs were cultured in NeuroCultTM Proliferation kit, supplemented with 10ng/ml human recombinant bFGF, 20ng/ml human recombinant EGF, and 1µg/ml Laminin, and grown in a monolayer on cell culture dishes at 37°C and 5% CO₂.

Cells were passaged at around 95% confluency to avoid overgrowth. Medium was aspirated and the cells were washed with sterile PBS. PBS was aspirated and Accutase was added. Cells were incubated at 37°C for 5 minutes and Accutase stopped

by addition of 4x the volume of wash medium. Cells were transferred to 15ml centrifuge tubes and centrifuged for 5 minutes at 350 rcf. Supernatant was aspirated and the cell pellet resuspended in growth medium and plated on fresh cell culture dishes.

2.2.2.2 Transfection

450000 cells were plated per well on 6-well plates one day prior to transfection. 2µg of STAgR plasmid or 3.2µg of dCas9 expression plasmid respectively were transfected with Lipofectamin 2000 according to the manufacturer's protocol. Cells successfully transfected with dCas9 expression plasmids were selected with 150µg/ml Hygromycin B, starting 2 days after transfection.

2.2.2.3 Flow cytometry and fluorescence activated cell sorting (FACS)

Cells were detached using Accutase as described in chapter 2.2.1. After centrifugation, cells were twice washed with PBS by resuspending the pellet in 1ml and centrifugation for 5 minutes at 350rcf. The cells were then resuspended in 1ml PBS containing 1µl eBioscience™ Fixable viability dye eFluor™ 660 and incubated at 37°C for 10 minutes. After incubation, 9ml PBS were added to each sample, and the samples were centrifuged for 5 minutes at 350rcf. The supernatant was then aspirated and cells resuspended in PBS containing 10% BSA at a maximum of 10×10^6 cells/ml. Cells were sorted into cell culture medium containing 5% BSA and cultured immediately after sorting as described in chapter 2.2.1. For flow analysis and sorting, the FACSAriaIII™ flow cytometer (Becton Dickinson) was used. Analysis was performed with the FlowJo X 10.0.7r2 software.

2.2.3 Lentiviral work

2.2.3.1 Lentivirus generation

5×10^6 HEK 293T cells were plated for transfection of the lentiviral vector cocktail 1 day before transfection on 3 10cm cell culture dishes and cultivated at 37°C and 5% CO₂. Each dish was then transfected with a mix of 15µg packaging plasmid

pCMVDR8.9, 20µg pseudotyping plasmid pVSVG and 10µg of the lentiviral expression plasmid using TransIT according to manufacturer's protocol. To harvest the virus, the supernatant was transferred into a 50ml centrifuge tube 96 hours after transfection, centrifuged for 10 minutes at 4500rcf, and filtered through a 0.45µm filter. To pellet viral particles, the supernatant was then centrifuged for 2 hours at 27000rcf and resuspended in 50µl TBS-5.

2.2.3.2 Titer measurement

For titer measurement of gRNA lentiviruses, 100000 Sox1-GFP NPCs were plated on each of 6 wells of a 12-well cell culture plate. 1 day later, a series of 6 dilutions of the lentivirus was prepared as follows: 1µl of virus was added to 1.5ml of medium and mixed by pipetting up and down several times. Of that dilution, 0.5ml were transferred to 1ml of medium and mixed similarly. This step was repeated 4 times, each time transferring 0.5ml of the new dilution to 1ml medium. The medium on the plated cells was aspirated, and 0.9ml of each dilution added to 1 well. 24 hours later the cells were detached and the cells from each well were plated on two new wells of a 12-well cell culture plate. 0.08µl Puromycin was added to one of the wells of each dilution, while the other wells were kept without selection. 3 days after the start of selection, the cells were detached from each well, pelleted, and resuspended in 10µl medium. 10µl of Trypan Blue were added and living cells were counted with a counting chamber. For each dilution the ratio of infected cells was calculated as follows:

$$TU = \frac{x * s}{y * v}$$

with

TU = transduction units [/ml]

x = cellcount in puromycin treated cells

y = cellcount in non-treated cells

s = number of cells at transduction

v = volume of virus in this dilution [ml]

2.2.3.3 Establishment of stable cell lines

To establish clonal dCas9-effector expressing cell lines, 300000 Sox1-GFP NPCs were plated per well on a 6-well cell culture plate. After 1 day, 2.5 μ l of the respective lentivirus were added to the different wells, and 24 hours later the medium was exchanged. Selection for dCas9 expression was started 3 days after transduction with 8 μ l/ml Blastidicin S. Resistant cells were seeded as single cells on a 96-well cell culture plate 7 days later, and clones were expanded. dCas9 expression was verified by qPCR and Western Blot. Stable cell lines were kept under Blastidicin S throughout all experiments.

To establish clonal NPC lines stably expressing gRNAs targeting the Sox1 promoter, 300000 NPCs were seeded on a 6 well cell culture plate and transduced the next day with the SoxProm lentivirus pool at an MOI of 4. The next day medium was changed and Puromycin added to the Medium at 0.6 μ g/ml. Cells were kept under selection for the duration of the experiments. 5 days after transduction, cells were seeded on 96 well cell culture plates at a density of one cell per well, and expanded.

2.2.3.4 Transduction of gRNA lentivirus

For candidate approaches, 250000 Sox1-GFP NPCs were plated on T25 cell culture flasks one day prior to transduction. An equivalent of 1×10^6 viral particles (equals multiplicity of infection (MOI) of 4) were added the next day. For screens, 1.5×10^6 Sox1-GFP NPCs were plated on T75 cell culture flasks one day prior to transduction. An equivalent of 600000 viral particles (equals an MOI of 0.4) was added the next day. Medium was changed 24 hours after transduction and selection with 0.08 μ l/ml Puromycin was started 48 hours after transduction. Cells were kept under selection for the remainder of the respective experiments.

2.2.4 Molecular methods

2.2.4.1 Reverse transcription and qPCR

For extraction of RNA, cells were detached using Accutase as described in chapter 2.2.1. Cells were centrifuged for 5 minutes at 350rcf and the supernatant was aspirated. RNA was extracted using the RNeasy Mini Kit according to the manufacturer's protocol. RNA content was measured using the NanoDrop ND-1000 and RNA was stored at -80°C or directly proceeded with. Maxima first strand cDNA synthesis kit was employed for DNase treatment and reverse transcription of mRNA. 100ng RNA per sample were employed, and reverse transcription performed according to the manufacturer's protocol. For quantification of targets on the QuantStudio™ 6 Flex Real-Time PCR System (Life Technologies), the PowerUP™ Sybr Green Master Mix was used. Each sample was analyzed in triplicates and quantified as mRNA content normalized to Gapdh. qPCR primers are listed in Table M7. Primers were designed to produce amplicons of 80 – 120 basepairs and efficiency was tested on a dilution of genomic DNA (reaching from 100ng/μl to 1.5625ng/μl in dilutions of 1:4. mRNA levels were quantified as relative levels compared to Gapdh.

2.2.4.2 Western Blotting

For western blot analysis, cells were detached using Accutase as described in chapter 2.2.1 and centrifuged at 350rcf for 5 minutes. The supernatant was aspirated and the cells resuspended in RIPA buffer supplemented with 1x protease inhibitor at density of 100000 cells per 50μl for optimal protein extraction. Samples were used immediately. Protein content of each sample was measured in duplicates using a Bradford assay according to manufacturer's protocol. Protein extract was stored at -80°C. Sterile H₂O was added to 40μg protein to a total of 16μl per sample, and 4μl 5x Laemmli buffer was added. Samples were boiled for 5 minutes at 95°C and separated using SDS PAGE on a 10% gel (see below) in PAGE running buffer at 100V for 90 minutes. After separation, proteins were blotted onto a 0.2μm nitrocellulose membrane (ThermoFisher, LC2009) in blotting buffer, using the Trans-Blot® Turbo™ Transfer

2 MATERIAL AND METHODS

System (Bio-Rad) according to the manufacturer's protocol. Membranes were blocked at room temperature in 5% milk powder in TBS-T for 60 minutes and incubated in primary antibodies (at working concentration) in TBS-T over night at 4°C. Membranes were washed 3 times for 5 minutes in TBS-T and incubated at room temperature for 2 hours in secondary antibodies (linked to HRP, at working concentration) in TBS-T. Membranes were washed three times for 5 minutes and analyzed on a ChemiDoc™ MP System (Bio-Rad) using ECL solution. Antibodies and gel concentrations are listed in Table M6.

Reagent	Volume	Reagent	Volume
Bis/Acrylamid	0,75ml	Bis/Acrylamid	3.3ml
4xTris/SDS pH 6.8	1.25ml	4xTris/SDS pH 8.8	2.5ml
H ₂ O	3ml	H ₂ O	4.2ml
10% APS	0.05ml	10% APS	0.05ml
Temed	0.01ml	Temed	0.01ml

Table M8: Stacking and Running Gel for PAGE

2.2.4.3 Differentiation assay

40000 Sox1-GFP NPCs were plated per well on poly-D-lysine coated coverslips in 24-well cell culture plates. Cells were cultured in NeuroCult™ Proliferation kit supplemented with 1% FCS (v/v) and 1µg/ml Laminin, but without growth factors. Every other day, the medium was changed. For short analysis, cells were fixed in 4% Paraformaldehyde as described in chapter 2.4.4 after 7 days. For neuronal subtype analysis, cells were differentiated for an additional 14 days. Medium was changed to 1ml NeuroCult™ supplemented with 1% FCS (v/v), 1µg Laminin, 10ng/ml NT3, BDNF, GDNF, and cAMP at day 7 and was exchanged every other day thereafter. Cells were fixed after 21 days.

2.2.4.4 Immunofluorescence staining and microscopic analysis

To fix cells, medium was aspirated and cells were washed two times in PBS. 4% PFA were added and the coverslips were incubated for 10 minutes at room temperature. After 15 minutes of permeabilization in 0.5% TritonX-100 in PBS (PBS-T), cells were treated for 30 minutes in 3% BSA (Sigma-Aldrich A2153-1KG) in PBS-T at room temperature. Cells were incubated overnight at 4°C in primary antibodies in 3%BSA in PBS-T. Coverslips were washed three times with PBS-T and incubated in the respective secondary antibodies at room temperature. From this step on, stainings were covered whenever possible to avoid quenching of fluorophores. Coverslips were washed three times and treated with DAPI. Coverslips were then mounted on glass slides using Aqua-Poly/Mount. Stainings were analyzed on an epifluorescence microscope (Zeiss). Per coverslip, several randomly selected pictures were taken and counted using Fiji software (National Institute of Health, USA). Pictures were treated the same way across all samples from a biological replicate and with linear processing techniques only. Antibodies are listed in Table M6.

2.2.4.5 Chromatin Immunoprecipitation (ChIP)-qPCR

Cells were detached using Accutase as described in chapter 2.2.1, aliquoted in 500000 cells, and washed twice in PBS. Cells were then fixed in 500µl 1% methanol free formaldehyde in PBS for 8 minutes at room temperature. The formaldehyde was then quenched using 57µl 1.25M glycine for 5 minutes at room temperature. Cells were centrifuged for 5 minutes at 500rcf and the supernatant was aspirated. After two washes with 4°C cold PBS, the pellet was resuspended in 500µl hypotonic buffer containing 1x protease inhibitor for 10 minutes on ice. Cells were centrifuged for 5 minutes at 500rcf, the supernatant aspirated, and the pellet resuspended in 500µl lysis buffer containing 1x protease inhibitor. After 30 minutes incubation on ice, the chromatin was transferred to sonication tubes and sheared with 5 on/off cycles of 30 seconds each with a Bioruptor™ (Diagenode). Each aliquot of sheared Chromatin, containing 500000 cells, was used for 4 ChIP reactions and 1 input control (equals 100000 cells per reaction) on the IP-Star® Compact (Diagenode). The Auto iDeal ChIP-seq kit for Histones was employed

according to manufacturer's protocol. The chromatin enrichment at target loci was quantified via qPCR in technical triplicates to control for variance, and quantified as % enrichment of the input sample. qPCR was performed as described in chapter 2.4.1. Primers and Antibodies used for ChIP-qPCR are listed in Tables M6 and M7.

2.2.5 Sequencing

2.2.5.1 Bisulfite and oxidative bisulfite sequencing

Primers for Bisulfite sequencing were designed to produce amplicons of 70 – 300 basepairs, which contain at least 4 CpGs. CpGs in the primer sequence itself were not admitted, and a minimum of 4 Cs in non-CpG context was applied. This guarantees the amplification of completely converted DNA only.

Prior to extraction of genomic DNA for bisulfite conversion, cells were detached with Accutase as described in chapter 2.2.1, aliquoted to 200000 cells per aliquot, and centrifuged at 350rcf for 5 minutes. After aspirating the supernatant, genomic DNA was extracted with the DNeasy Blood and Tissue kit. DNA was stored at -20°C. For oxidative bisulfite conversion, 1µg of DNA was denatured in 0.05M NaOH at 37°C for 30 minutes. Micro-Bio-Spin P-6 SCC columns were centrifuged at 1000 x rcf for 1 minute and washed four times with 500µl H₂O. After incubation, DNA was loaded on washed columns and centrifuged for 8 minutes at 1000 x rcf, and then cooled on ice for 5 minutes. 1µl of 15mM K₂Cr₂O₇ (in 0.05M NaOH) was added to the DNA and DNA oxidized for 1 hour on ice. After Oxidation, DNA was loaded on prewashed columns and centrifuged for 8 minutes at 1000 x rcf. After oxidation, DNA was immediately converted as follows.

DNA was converted with the EZ DNA Methylation-Gold Kit directly after extraction of genomic DNA to measure the amount of Cytosine methylation (5mC) and hydroxymethylation (5hmC) combined, or after oxidation to measure the amount of 5mC specifically. The converted DNA was used as template for the sequencing library preparation. Converted DNA was stored at -20°C.

2 MATERIAL AND METHODS

To amplify target loci and to introduce overhangs for the Illumina indexing adapters, a primary PCR was performed. 1µl of bisulfite converted DNA was mixed with 2.5µl 10x buffer, 1µl of 10nM dNTPs, 0.25µl of each, the forward and reverse primer targeting the locus of interest, 0.5µl HotStarTaq Polymerase, and 19.5µl H₂O. Reagents were taken from the HotStarTaq kit. PCR products were purified with Agencourt AMPure XP magnetic beads according to Manufacturer's specifications. In a second PCR, purified products were amplified with Illumina Sequencing adapters from the Nextera DNA Library Preparation Kit to introduce sequencing adapters. 1µl of DNA was mixed with 12.5µl of Phusion® High-Fidelity PCR Master Mix, 10µl H₂O, and 0,75µl of each, forward and reverse primer. Each sample was amplified with a different combination of barcodes to ensure simple identification of reads after sequencing. The PCR products were purified with Agencourt AMPure XP magnetic beads according to Manufacturer's specifications.

The DNA content in the purified samples was measured on the Infinite® M1000 plate reader (Tecan), using the Quant-iT™ PicoGreen™ dsDNA Assay Kit according to the manufacturer's protocol.

Samples were pooled in an equimolar ratio, and the quality was analyzed with an Agilent Bioanalyzer 2100 on a DNA High Sensitivity chip according to the manufacturer's protocol.

For Sequencing of the Bisulfite Pool, a Miseq kit v2 was used. The pool was diluted to 4nM with H₂O and 5µl of the pool were added to 5µl of 0.2M NaOH. The mix was vortexed briefly, centrifuged for 1 minute at 280rcf, and incubated at room temperature (20-25°C) for 5 minutes to allow denaturation. 990µl of Hyb Buffer were added, and 210µl of the resulting dilution were mixed with 390µl Hyb Buffer to achieve a final library concentration of 7pM. 2µl PhiX control library were mixed with 3µl Tris x HCl and 5µl 0.2M NaOH, vortexed briefly, centrifuged at 280rcf for 1 minute, and incubated at room temperature for 5 minutes. PhiX was diluted with 990µl Hyb Buffer. 540µl of the 7pM library were mixed with 60µl PhiX library and loaded on the sequencing cartridge. The sequencing was performed on a MiSeq (Illumina).

2 MATERIAL AND METHODS

Demultiplexed sequencing reads were retrieved. Trim Galore (Babraham Institute) was used with default settings to trim the adapter sequences and filter reads of insufficient quality. Trimmed sequences were aligned versus a bisulfite converted mouse genome (version mm10) using the Bowtie aligner (Bowtie 2.3.4.3, John Hopkins University), the methylation count calculated, and the bed file converted to bedGraph using the Bismark methylation extractor with default settings (Babraham Institute). The methylation levels of CpGs of interest were extracted in R Studio v1.1.442 using the RSamtools (version 134.1) and the BiSeq (version 1.22.0) packages.

2.2.5.2 RNA sequencing

The RNA of 100000 cells per sample was isolated with the PicoPure™ RNA Isolation Kit according to the manufacturer's protocol. RNA was stored at -80°C. The RNA quality was assessed on the Agilent Bioanalyzer 2100. RNA samples with an integrity of <8 were dismissed. cDNA synthesis, amplification, and purification were performed with 4.5ng RNA per sample, using the SMART-Seq™ v4 Ultra Low Input RNA Kit according to the manufacturer's protocol. cDNA was amplified with 8 cycles. Shearing of the DNA was performed on the Covaris S220 sonicator, and the library was prepared using the MicroPlex Library Preparation Kit v2. The quality and quantity of the library were assessed on the Agilent Bioanalyzer 2100. Samples were pooled in an equimolar ratio and the pool was diluted to a final concentration of 5nM. The library was sequenced on a HiSeq4000 (Illumina) by the Sequencing Core Facility of the Helmholtz Zentrum München.

The demultiplexed reads were retrieved and Sequences aligned to the mouse transcriptome (mm10) using the STAR v2.6 RNA-seq aligner. Expression was calculated using the RSEM tool (v1.1.14), and analyzed in R Studio v1.1.442 using the DESeq2 package (Love, Huber, and Anders 2014).

2.2.5.3 Sequencing of gRNA amplicons

For DNA extraction of screened cells, cells were sorted directly into the Lysis buffer of the DNeasy Blood and Tissue kit. DNA was extracted according to

2 MATERIAL AND METHODS

manufacturer's specifications, and stored at -20°C. To amplify the gRNA locus from the genomic DNA, a PCR with 1µl of each, Nextera_fwd and Nextera_rev primers in Phusion® High-Fidelity PCR Master Mix was performed. PCR products were purified with Agencourt AMPure XP magnetic beads to Manufacturer's specifications. To introduce Illumina sequencing adapters, DNA amplicons were amplified with Illumina Sequencing adapters from the Nextera DNA Library Preparation Kit using Phusion® Master Mix. Samples were pooled at equimolar ratio and quantity and quality assessed on the Agilent Bioanalyzer 2100. The library was sequenced on a HiSeq4000 (Illumina) by the Sequencing Core Facility of the Helmholtz Zentrum München.

Demultiplexed sequencing reads were retrieved. Trim Galore (Babraham Institute) was used with default settings to trim the adapter sequences and filter reads of insufficient quality. Trimmed sequences were aligned against the gRNA library sequences using the Bowtie aligner (Bowtie 2.3.4.3, John Hopkins University). Sequence enrichment was calculated in R using the Rsamtools, dplyr (version 0.8.0.1), and the DESeq2 packages.

3 Results

3.1 Generation and characterization of dCas9-expressing cell lines

Deactivated Cas9 lacking the nuclease activity of its wild type form has been successfully used as tool for targeting of various effector domains to specific genomic loci. In this study, I made use of different dCas9 fusion proteins to target a variety of genes in mouse neural progenitor cells (NPCs). In order to guarantee expression of dCas9-fusion proteins in all experiments, clonal cell lines of cultured NPCs stably expressing the respective fusion protein were established for different effector domains fused to dCas9 as described in chapter 2.2.3.3. Transgene expression was confirmed via qPCR, where high levels of dCas9 mRNA were detected in all tested clonal lines (Figure 4). Clonal cell lines were kept under Blasticidin S selection for all experiments to avoid silencing of the trans-genic construct.

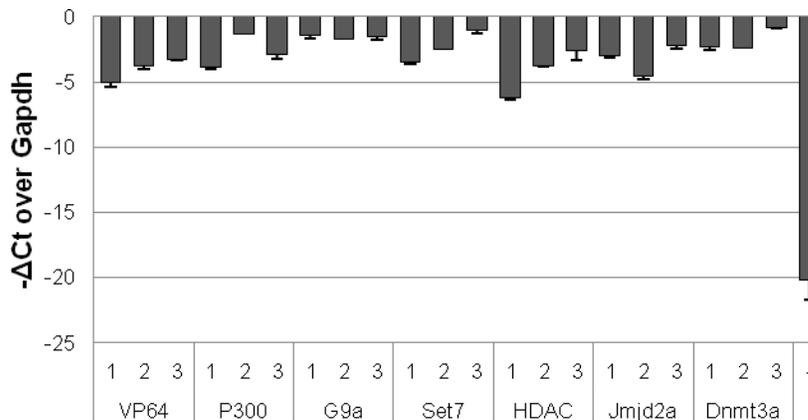


Figure 4: Clonal lines express high levels of dCas9 mRNA. The established clonal lines were tested for dCas9 expression with qPCR. For each dCas9 modifier, three cell lines were tested, and strong expression could be confirmed in each of them.

3.2 Targeted activation of Sox1

Chapter 3.2 of this thesis was prepared in parallel with the manuscript "Targeted removal of epigenetic barriers during transcriptional reprogramming", published in Nature Communications (Baumann et al. 2019). Therefore, the text and data presented in this thesis may overlap with the manuscript.

Neural stem cells (NSCs) that are cultured in proliferative conditions differentiate to NPCs. These differ not only in morphology, but also lose the potency for neuronal differentiation. The neural transcription factor Sox1 that is strongly expressed in NSCs but not in NPCs, has been shown to be neurogenic *in vivo* and *in vitro*, which makes it a potential cell fate determining factor for NSCs. To test this hypothesis we aimed to induce transcription of the endogenous Sox1 in NPCs by transcriptional engineering, and further investigate whether reactivation of this gene would restore the neurogenic differentiation potential. To this end we made use of NPCs stably expressing dCas9-VP64, a targetable transcriptional activator.

3.2.1 Sox1 upregulation following VP64 targeting

To test the capacity of the transcriptional activator VP64 to induce the expression of Sox1 in NPCs, I used the three generated clonal NPC lines expressing dCas9-VP64, and transfected an expression plasmid containing two gRNAs targeting the promoter of Sox1 (S1-9). As positive control, I transfected an expression plasmid with two gRNAs targeting Actc1 (A1-9), an actin gene expressed in heart and muscle tissue, which has been shown to be easily induced by this tool (Chavez et al. 2015) (see Figure 5A+B for gRNA position). The effect of VP64 targeting was assessed by qPCR. Strong induction of Actc1 mRNA (>100-fold) to expression levels comparable to those found in heart tissue (as quantified by qPCR, Figure 5C) confirmed the functionality of dCas9-VP64 in the NPCs. In contrast, upregulation of Sox1 mRNA was, albeit statistically significant, almost forty fold lower (Figure 5C).

3 Results

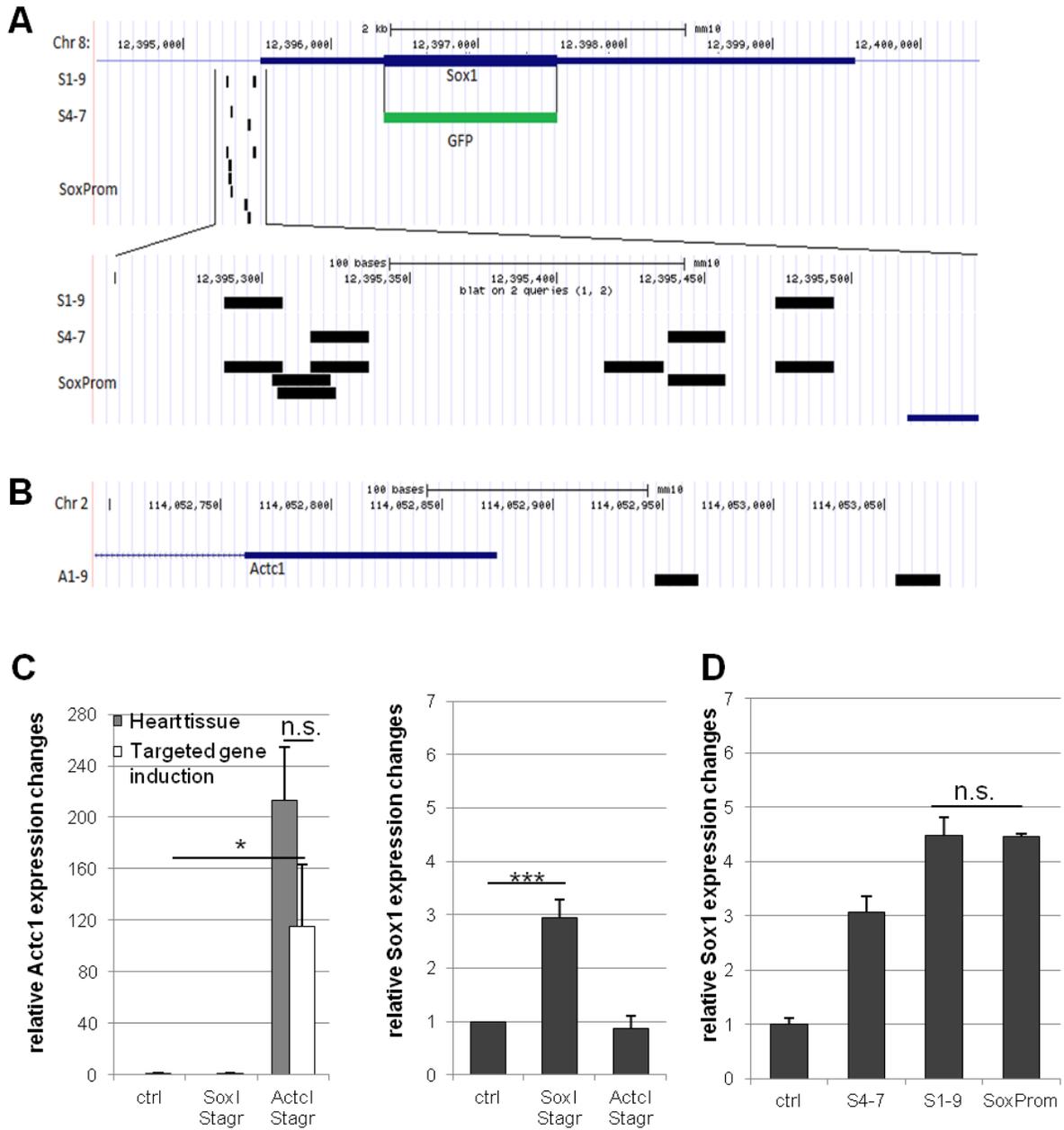


Figure 5: Targeted induction of Sox1 in NPCs is significant, but minor. **A** gRNA position of all gRNAs used for Sox1 targeting. Sox1 gene shown in blue, heterozygous GFP knock-in in the ORF of Sox1 shown in green. **B** gRNAs targeting Actc1. **C** Targeting dCas9-VP64 to the promoter of Sox1 and Actc1 leads to significant increase of the respective mRNA. Increase ~120-fold for Actc1 and ~3-fold for Sox1. No off-target effects can be observed. **D** Variation of the employed Sox1 targeting gRNAs does not significantly increase the efficiency of gene induction. Data shown as mean + standard deviation of three biological replicates (performed on different days in different clonal NPC lines). N.s. not significant; * p-value < 0.05; *** p-value < 0.005;

3 Results

Several publications have shown that the amount of gRNAs employed and the position of targeting relative to the gene's transcription start site (TSS) can strongly impact the efficiency of TAFs. I therefore designed an alternative STAgR construct containing two different gRNAs (S4-7), as well as a pool of 7 individual gRNAs (SoxProm, Figure 5A), each targeting the Sox1 promoter. To furthermore rule out the possibility that differences in transfection efficiency affect the efficiency of target activation, we exchanged gRNA expression plasmids for lentiviral particles. All lentiviruses were transduced into the dCas9-VP64 expressing NPCs at an MOI of 4 to guarantee high gRNA levels in all cells, and cells were selected for gRNA expression with Puromycin as described in chapter 2.2.3.4. Even though the induction of Sox1 mRNA was slightly stronger compared to transfection of gRNA plasmids without selection, it still was over 10-fold weaker than that of Actc1. In addition, exchanging gRNAs did not lead to increased efficiency (Figure 5D), indicating that neither the exact target site or sequence, nor the mode of gRNA delivery were responsible for the minor response.

To test whether the limited response was due to minor, but uniform induction of expression, or rather strong induction in only a subpopulation of the NPCs, I performed immunocytochemistry (ICC) on cells co-transfected with A1-9 and S1-9 gRNA expression plasmids, co-staining for both targets. While control cells without gRNAs were largely negative for both markers, Actc1 was expressed in a majority of cells, whereas only around 10% of the cells were positive for Sox1 after targeting VP64 to the promoter of the respective gene (Figure 6A+B). To confirm this result, I made use of NPCs derived from ESCs carrying a heterozygous GFP knock-in in the ORF of Sox1, earlier described by Aubert et al. in 2003 (see page 14 and Figure 5A). In these cells, GFP expression correlates to Sox1 expression and can be used as a direct read-out thereof. For clarity, throughout this thesis cells expressing GFP will be termed Sox1^{GFP} positive, while others will be termed Sox1^{GFP} negative. I transduced dCas9-VP64 expressing NPCs with the S1-9 lentivirus at an MOI of 4 and analyzed GFP expression at 7 days after transduction via flow cytometry. Cells without gRNAs or transduced with control gRNAs (non-targeting) were almost exclusively Sox1^{GFP} negative. However, even when transduced with Sox1 targeting gRNAs, only a minority of NPCs reacted with GFP

3 Results

upregulation (on average 2.9% cells in GFP positive gate, p-value < 0.0001, n = 18 biological replicates performed in three different clonal lines and on different days; Figure 6C+D). These results underline the findings of the ICC, in that only a small part of NPCs react to the activating stimulus of VP64 at the Sox1 promoter, while the majority remains unresponsive.

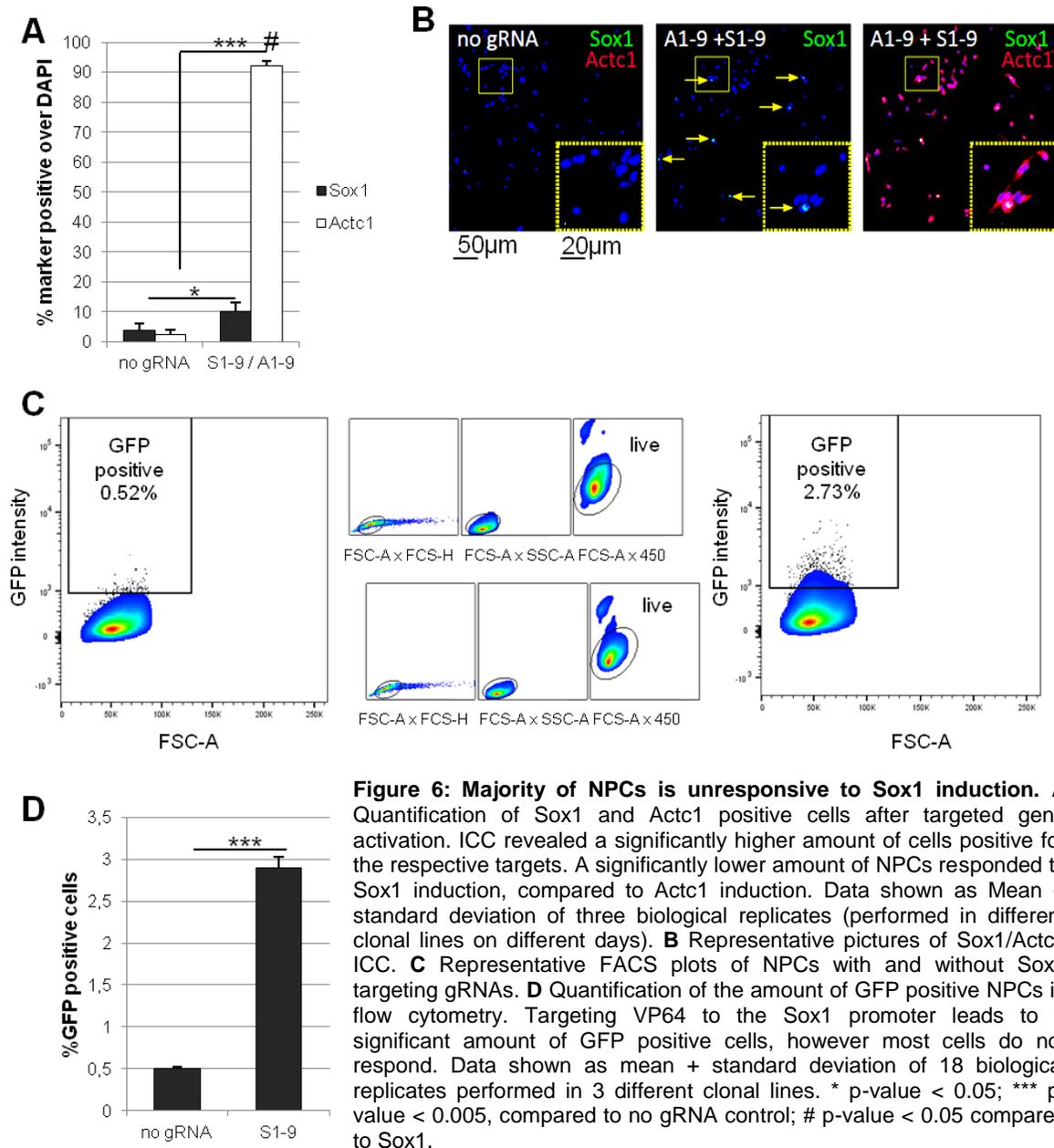


Figure 6: Majority of NPCs is unresponsive to Sox1 induction. A Quantification of Sox1 and Actc1 positive cells after targeted gene activation. ICC revealed a significantly higher amount of cells positive for the respective targets. A significantly lower amount of NPCs responded to Sox1 induction, compared to Actc1 induction. Data shown as Mean + standard deviation of three biological replicates (performed in different clonal lines on different days). **B** Representative pictures of Sox1/Actc1 ICC. **C** Representative FACS plots of NPCs with and without Sox1 targeting gRNAs. **D** Quantification of the amount of GFP positive NPCs in flow cytometry. Targeting VP64 to the Sox1 promoter leads to a significant amount of GFP positive cells, however most cells do not respond. Data shown as mean + standard deviation of 18 biological replicates performed in 3 different clonal lines. * p-value < 0.05; *** p-value < 0.005, compared to no gRNA control; # p-value < 0.05 compared to Sox1.

3 Results

To confirm that GFP expression is representative of Sox1 levels, I sorted Sox1^{GFP} positive and negative cells from samples with S1-9 gRNAs and control samples (without gRNAs or with non-targeting gRNAs) and performed qPCR and Western Blot analysis for Sox1. All controls expressed Sox1 at very low levels, both in the Sox1^{GFP} positive as well as the Sox1^{GFP} negative population; the same was observed in the Sox1^{GFP} negative population from NPCs expressing S1-9 gRNAs. This indicated that in control samples, the events that are recorded in the GFP gate were false-positive and did not represent Sox1 expression (Figure 7A). Sox1^{GFP} positive cells that were sorted from the sample transduced with S1-9 gRNAs did however express significantly higher levels of Sox1 mRNA (Figure 7A) and, even more pronounced, Sox1 protein. The upregulation reached levels of around 30-fold compared to control samples (Figure 7B).

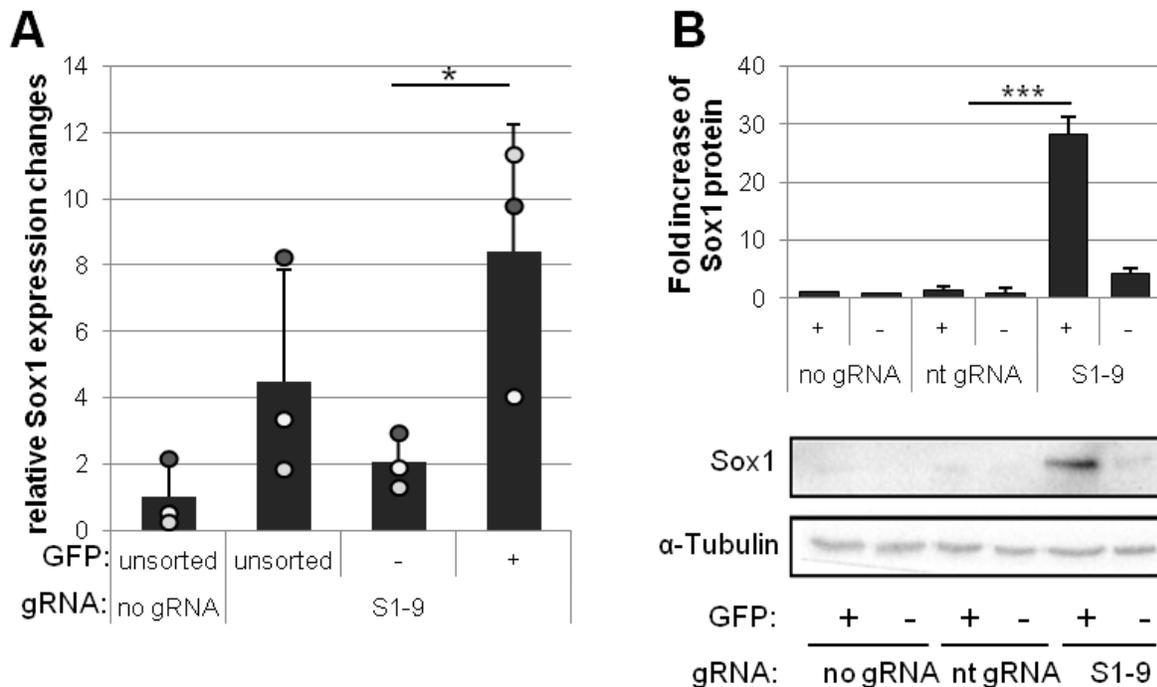


Figure 7: GFP and Sox1 levels correlate in Sox1^{wt/GFP} NPCs. **A** qPCR performed on unsorted and sorted NPCs with and without Sox1 targeting gRNAs reveals that cells sorted from the GFP positive gate do indeed express significantly higher levels of Sox1 mRNA. **B** Western Blot for Sox1 on sorted NPCs verifies this finding on protein level. Data shown as mean + standard deviation of three biological replicates (performed in different clonal lines on different days). * p-value < 0.05; *** p-value < 0.005;

3 Results

To further assess the stability of induced Sox1 expression in NPCs, I sorted 5000 GFP positive cells from the control (no gRNAs) and the sample expressing S1-9 gRNAs, and expanded them. Flow cytometry performed 14d after FACS showed that cells from the control resembled the initial population, corroborating the finding that these cells are false positives with no stable Sox1 expression (Figure 8). Cells sorted from the gRNA expressing sample on the other hand were strongly GFP positive (up 45% cells in the GFP gate, see Figure 8). This indicates that expression of the target gene is stable, once induced.

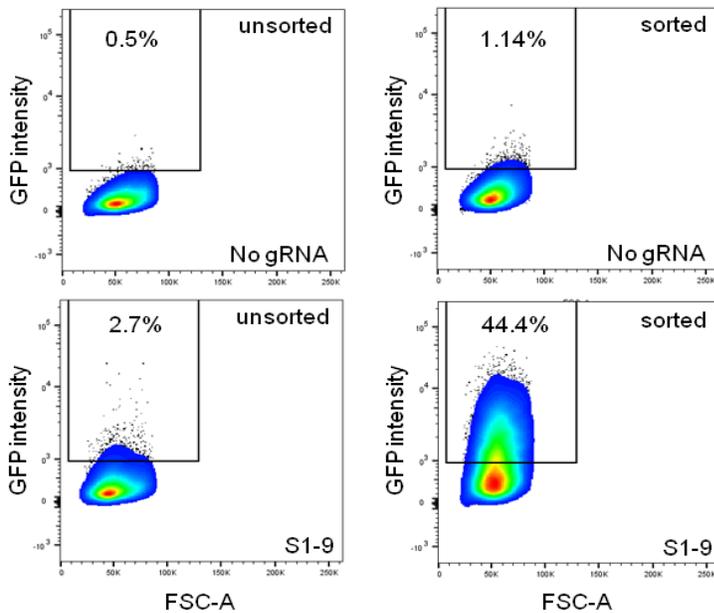


Figure 8: GFP induction is stable in induced cells. Sorted GFP positive cells from the control sample cluster comparable to the starting population after expansion. Cells sorted from S1-9 expressing NPCs show a high proportion of GFP positive cells after expansion, indicating that gene induction is at least in part stable.

Taken together, these results confirm the functionality of targeted induction of *Actc1* by dCas9-VP64 in NPCs. Transcriptional activation of the developmental transcription factor Sox1 however proved difficult, as the majority of cells were unresponsive to the stimulus. In those cells that did however induce Sox1 expression, this effect was long-lasting.

3.2.2 Characterization of Sox1 positive cells

To test whether transcriptional activation of Sox1 in NPCs would restore the identity of NSCs, I investigated the transcriptome of these cells. I performed RNA sequencing on sorted Sox1^{GFP} positive cells from Sox1^{wt/GFP} NPCs transduced with the S1-9 gRNA lentivirus, untransduced NPCs as control, and NSCs derived from Sox1^{wt/GFP} embryonic stem cells (ESCs). As these cells display heterogeneous levels of Sox1, I sorted Sox1^{GFP} positive cells to obtain a strictly Sox1 positive control (Neural Rosettes, NR).

Principal component analysis (PCA) of these samples revealed two clusters of cells, one containing the developmentally Sox1 positive controls (NSCs and NR), and the other containing the NPCs (both untransduced and Sox1^{GFP} positive). Within this latter cluster however, the two states of NPCs clustered clearly apart from each other (Figure 9A). This was confirmed by hierarchical clustering: analysis of all four cell types showed a strong separation of NSCs and NR, and NPCs (untransduced and GFP positive), however within the NPC cluster, cells showed only a mild tendency to cluster according to their Sox1 expression status. This effect was however more pronounced when these two methods were applied to only untransduced and Sox1^{GFP} positive NPCs (Figure 9B), strongly indicating that even though the global changes induced by Sox1 upregulation are seemingly small compared to the difference of the two developmental stages, they are enough to clearly distinguish Sox1^{GFP} positive NPCs from the negative baseline.

I further analyzed differentially regulated genes to investigate in more detail the nature of the propagated change. Between the Sox1^{GFP} positive NPCs and the untransduced control, 1060 genes were significantly upregulated (>4fold, p-value<0.05) and 482 genes were significantly downregulated (>4fold, p-value<0.05), as shown in Figure 10A. Most of the significantly upregulated genes were also found to be significantly increased in the transcriptomes of NSCs and NR, when compared to Sox1^{GFP} negative NPCs (Figure 10C), indicating a progression of NPCs towards a NSC identity upon Sox1 induction. To confirm this hypothesis, I investigated the mRNA levels of different neuroepithelial markers that should be strongly expressed in NSCs. 3 of the

3 Results

9 genes that I analyzed, namely E-Cadherin, Prominin, and Sox10, were not differentially regulated between control NPCs and Sox1^{GFP} positive NPCs. The other 6 displayed significantly different levels after the induction of Sox1, even though none of them reached the levels present in NSCs (Figure 10B). This further indicated a clear, albeit not fully pronounced change in cell identity towards NSCs.

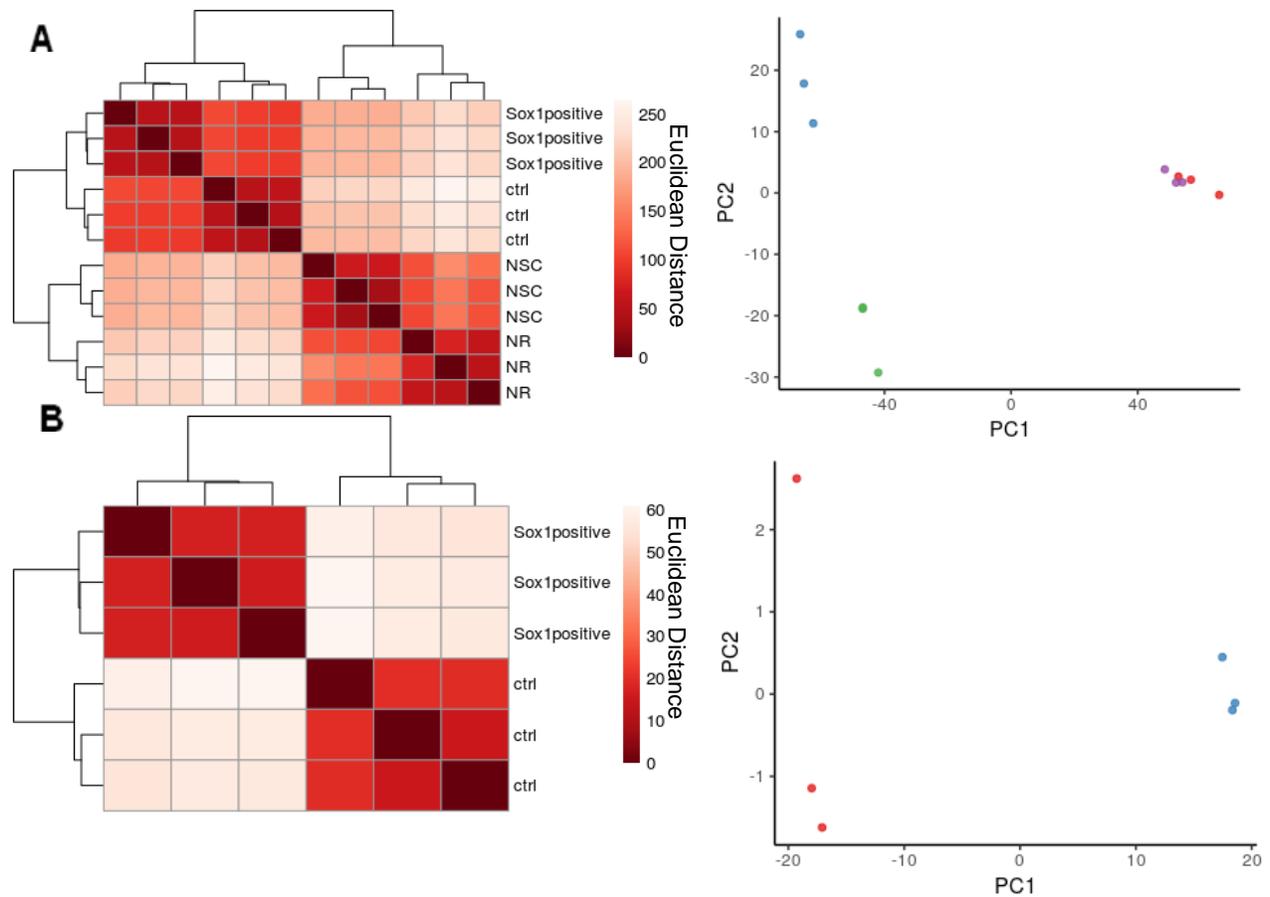


Figure 9: Transcriptomes of Sox1 positive and negative cells cluster apart. Heatmaps and PCA of either NSCs, NR, Sox1 positive NPCs, and control NPCs, or only Sox1 positive NPCs and control NPCs displays a strong clustering of the two developmental stages (A, red: control; blue; neural rosettes; green: NSCs; purple: Sox1 positive NPCs), but also the two Sox1 stages in NPCs (B, red: control; blue; Sox1 positive NPCs), indicating cell identity changes.

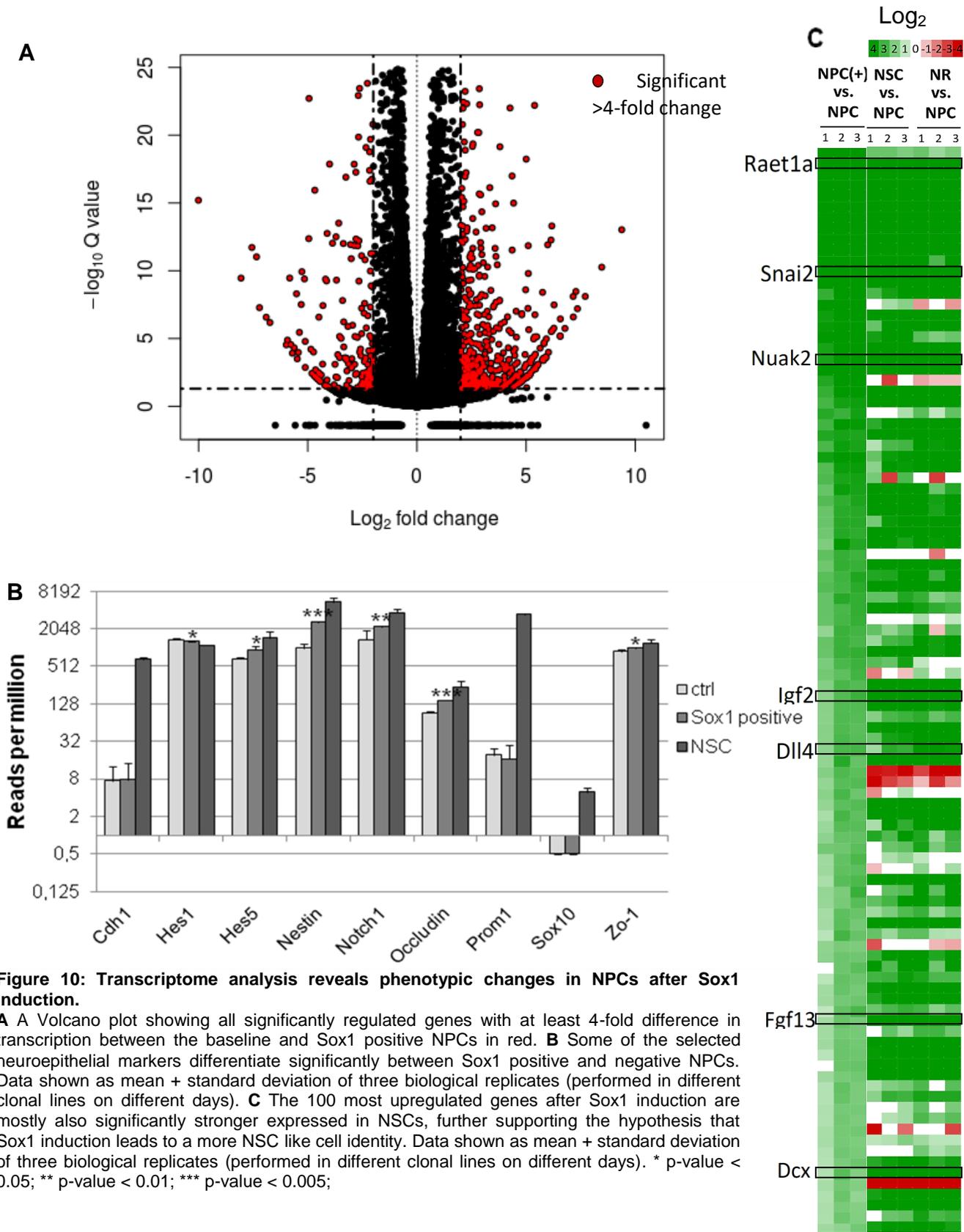


Figure 10: Transcriptome analysis reveals phenotypic changes in NPCs after Sox1 induction.

A A Volcano plot showing all significantly regulated genes with at least 4-fold difference in transcription between the baseline and Sox1 positive NPCs in red. **B** Some of the selected neuroepithelial markers differentiate significantly between Sox1 positive and negative NPCs. Data shown as mean + standard deviation of three biological replicates (performed in different clonal lines on different days). **C** The 100 most upregulated genes after Sox1 induction are mostly also significantly stronger expressed in NSCs, further supporting the hypothesis that Sox1 induction leads to a more NSC like cell identity. Data shown as mean + standard deviation of three biological replicates (performed in different clonal lines on different days). * p-value < 0.05; ** p-value < 0.01; *** p-value < 0.005;

3 Results

To further test the neuroepithelial character of NPCs after induction of Sox1, I performed immunocytochemistry (ICC) on sorted Sox1^{GFP} positive cells from NPCs transduced with the S1-9 lentivirus, and untransduced NPCs as control. Sox1 positive NPCs showed a significantly higher proportion of cells positive for 4 out of 6 neuroepithelial markers. More specifically, Nestin, Notch1, and the tight junction proteins Occludin and zonula occludens 1 (Zo1) were expressed in a significantly higher number of cells. Of note, the amount of cells expressing the surface protein E-Cadherin and the early NSC marker CD133 were not significantly increased (Figure 11). These findings confirmed the results of RNA sequencing, in that Sox1 induction partially restores a neuroepithelial character in NPCs.

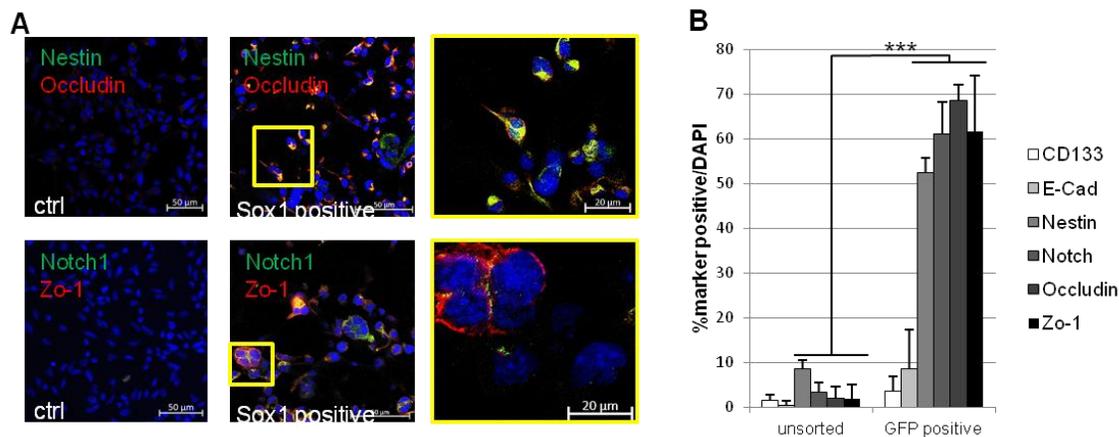


Figure 11: Sox1 induction leads to expression of neuroepithelial markers in NPCs. **A** Representative image of ICC for different neuroepithelial markers in Sox1 positive and negative NPCs. **B** Quantification of cells expressing the different tested marker genes reveals a significantly higher amount of marker positive NPCs for 4 out of six markers after Sox1 induction. Data shown as mean + standard deviation of 3 biological replicates (performed in different clonal lines on different days). *** p-value < 0.005;

Sox1 has been shown to have strong neurogenic potential *in vivo* and *in vitro* (Kan et al. 2004, Kan et al. 2007). It is furthermore known that NSCs, which bear the potential to generate not only glial cells but also cells of the neuronal lineage, strongly express Sox1, while more committed NPCs are negative for the transcription factor. I therefore aimed to test, whether the lost neurogenic potential can be restored by upregulation of Sox1 in NPCs. To that end I sorted Sox1^{GFP} positive and negative cells from an untransduced control sample and a sample transduced with the S1-9 lentivirus,

3 Results

and differentiated the cells for 7 days by removing growth factors from the culture medium. I then performed an immunostaining for the astrocytic marker S100 β and the early neuronal marker Tuj1 to investigate possible changes in the differentiation potential of NPCs. As expected, cells differentiated from both, Sox1^{GFP} positive and negative untransduced control samples, as well as Sox1^{GFP} negative cells isolated from the transduced NPCs, were almost exclusively positive for S100 β , while almost no Tuj1 positive cells were observed. The amount of cells positive for the neuronal marker were however significantly increased in the NPC population with induced Sox1 expression, where significantly fewer cells expressed the astrocytic marker (Figure 12A+B).

To further analyze the generated neurons, cells were stained for the markers of mature neurons Map2 and NeuN. However, after 7 days of differentiation, none of the Tuj1 positive cells were positive for the mature neuronal markers (Figure 12). I therefore extended the period of differentiation to 21 days, adding neuronal survival medium every other day, from day 7 of differentiation. After prolonged differentiation, generated neurons were positive for Map2. To investigate their subtypes, I stained differentiated neurons for the subtype specific markers vGlut1 (glutamatergic neurons (Bellocchio et al. 2000)), Calbindin (gabaergic neurons (Hendry et al. 1989)), and tryptophan hydroxylase (TH, serotonergic neurons (Walther et al. 2003)). Over 70% of Map2 positive neurons were also positive for vGlut1, while only 15% expressed Calbindin. None of the neurons were positive for TH (Figure 12C+D).

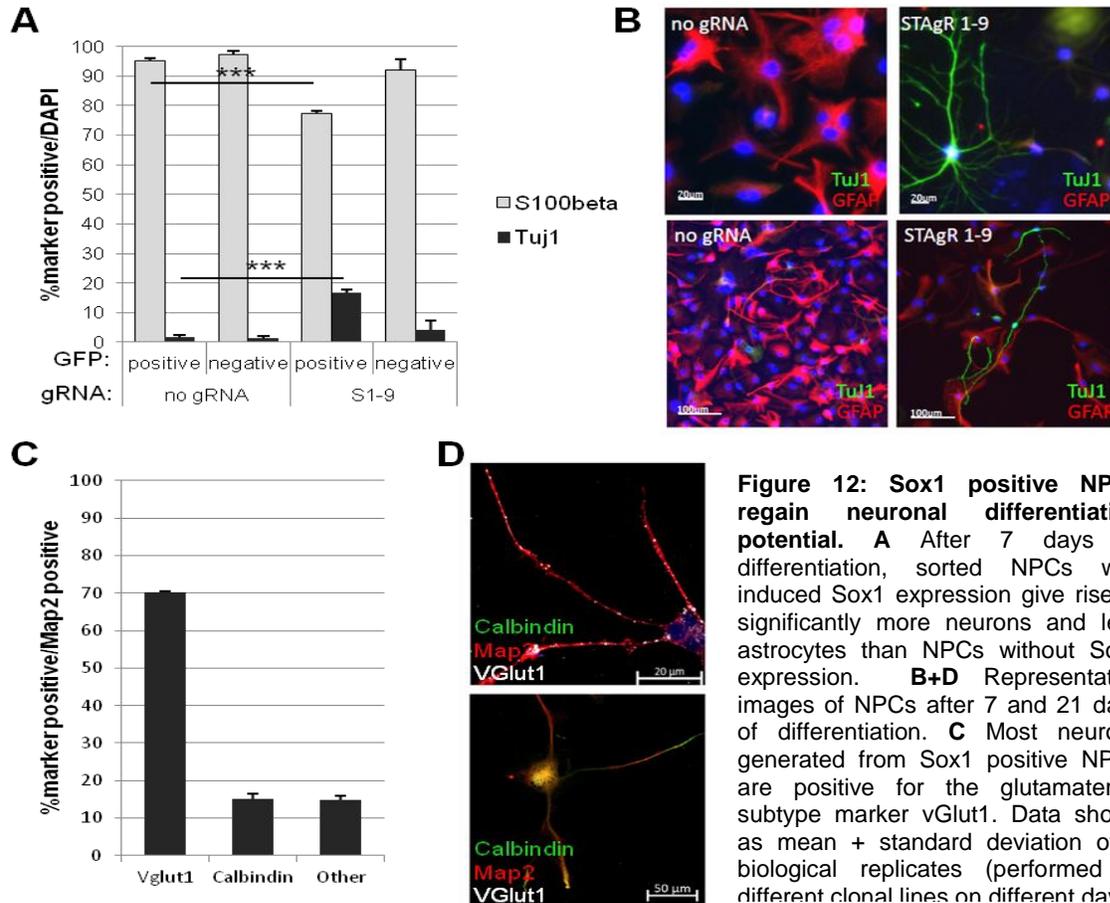


Figure 12: Sox1 positive NPCs regain neuronal differentiation potential. **A** After 7 days of differentiation, sorted NPCs with induced Sox1 expression give rise to significantly more neurons and less astrocytes than NPCs without Sox1 expression. **B+D** Representative images of NPCs after 7 and 21 days of differentiation. **C** Most neurons generated from Sox1 positive NPCs are positive for the glutamatergic subtype marker vGlut1. Data shown as mean + standard deviation of 3 biological replicates (performed in different clonal lines on different days).

3.3 Investigation of technical barriers to targeted gene activation

The analysis of Sox1^{GFP} positive NPCs showed a substantial change in phenotype, as indicated by changes in their transcriptome, expression of neuroepithelial markers, and regeneration of neuronal differentiation potential. Given the potency of the dCas9-VP64 tool to invoke such phenotypic changes in NPCs, it is striking that despite the strong stimulus only a small subpopulation was responsive to the Sox1 induction. I therefore aimed to test whether technical limitations of the dCas9 targeting tool could be the reason for the disparate cellular response.

To first investigate whether variations in gRNA levels due to different copy numbers or integration sites could be limiting to the activating potential of dCas9-VP64, I then generated clonal NPC lines expressing gRNAs from the SoxProm pool (see

3 Results

chapter 2.2.3.3). I transfected a dCas9-VP64 expression plasmid into four different gRNA expressing cell lines and compared the amount of responsive cells in the different clones 8 days after transfection. Flow cytometry revealed that no clone failed to respond completely, with at least 1.5% Sox1^{GFP} positive cells in each clone. Furthermore, there were no clones with a significantly higher amount of GFP positive cells than any other clone or the clonal dCas9-VP64 expressing NPCs transduced with Sox1 targeting gRNAs (chapter 3.2.1, Figure 13). It is therefore unlikely that differences in gRNA levels after transduction are responsible for the discrepant response of NPCs to Sox1 induction.

Next, to rule out differences of dCas9 levels in responsive and unresponsive NPCs, I sorted Sox1^{GFP} positive and negative cells from NPCs transduced with S1-9 gRNAs, and the corresponding populations from two control samples (without gRNAs, or with non-targeting gRNAs), and performed western blot analysis for dCas9. All six populations expressed similar levels of dCas9 protein, with no significant differences (Figure 13A).

Several recent publications on Cas9 suggest that chromatin compaction and high histone occupancy at the target site can influence the binding efficiency of the nuclease. This could also affect its modified catalytic mutant, dCas9, and subsequently the activating effect of VP64 and lead to a heterogeneous response to the stimulus even in cells with comparable levels of dCas9 protein and gRNAs. I therefore aimed to quantify the amount of bound dCas9 protein at the Sox1 promoter of responsive and unresponsive NPCs using chromatin immunoprecipitation (ChIP)-qPCR. As shown in Figure 13B, dCas9 is specifically enriched at the Sox1 promoter, but not at control loci. Furthermore, levels of bound dCas9 at the target site are comparable between GFP positive and negative populations, excluding dCas9 as reason for the discrepancy in Sox1 activation.

The possibility however remained that VP64 was too limited in its efficiency and that a stronger transcriptional activator could overcome the barriers in NPCs that are unresponsive to Sox1 induction by VP64. Several of such TAFs have been published over the last years with varying degrees of potency (see chapter 1.5.2). VPR is a VP64

3 Results

domain additionally fused to two trans-activating domains from the NF-kappa-B p65 subunit and the Epstein-Barr-Virus (Rta). This modified VP64 has been shown to activate target genes to up to 120 fold compared to the traditional VP64, when targeted in the same fashion (Chavez et al. 2015). To compare the effect of the two TAFs on Sox1 expression in Sox1-GFP NPCs, clonal NPC lines expressing Sox1 targeting gRNAs were transfused with dCas9-VP64 and dCas9-VPR respectively, and selected for the expression of fusion protein as described in chapter 2.2.2.2. 8 days after transfection Sox1 expression was analyzed on mRNA level via qPCR, and on protein level as GFP expression via flow cytometry. qPCR confirmed that VPR is indeed a stronger trans-activator than VP64, as Sox1 mRNA levels were significantly higher after targeting with the former than the latter. However, the amount of GFP positive cells as determined by flow cytometry was not significantly different between the two TAFs (Figure 13D). This indicated that even though induction of Sox1 is indeed stronger with VPR than with VP64, no additional cells were responsive to this stronger stimulus.

Taken together, these results show that not only do all NPCs express dCas9-VP64 to comparable levels, independent of their responsiveness to Sox1 induction, and that gRNA levels are likely not responsible for the heterogeneity in the response, but also that similar amounts of dCas9 are bound to the Sox1 promoter in both the responsive and unresponsive population. This strongly suggests that the targeting tool itself is working properly, and that cell intrinsic differences in NPCs impair the induction of Sox1 downstream of dCas9-VP64 binding.

3 Results

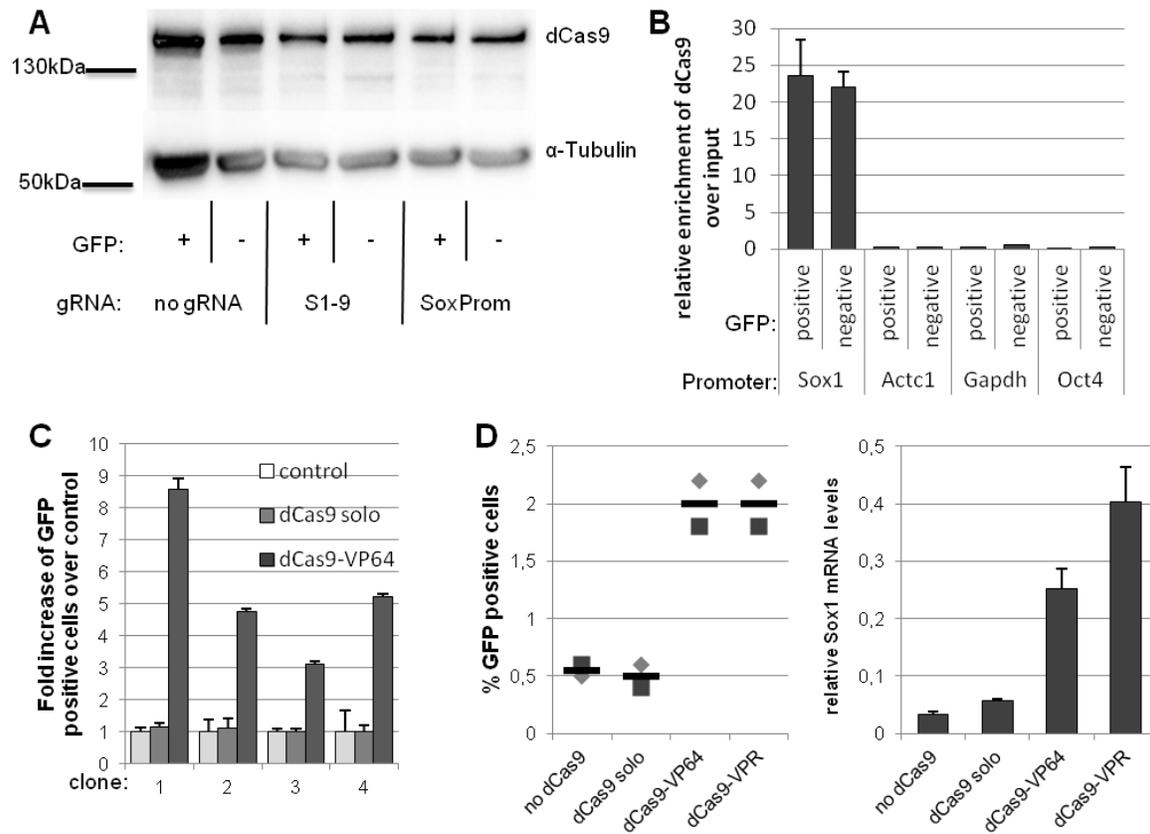


Figure 13: Discrepancy in response to induction not based on technical issues. **A** Western Blot of dCas9 on sorted NPCs with and without S1-9 gRNAs shows comparable levels of the targeting protein between all samples. **B** Quantification of dCas9 via ChIP qPCR reveals strong enrichment at the Sox1 promoter in the positive and negative population. **C** Analysis of GFP positive cells after Sox1 induction in different gRNA clonal lines. Comparable amount of activation in all tested lines. Data shown as mean + standard deviation of technical replicates. **D** VPR leads to a significantly stronger induction of Sox1 mRNA compared to VP64, the amount of responsive cells however stays the same. Data shown as single data points of biological replicates (performed on different days; flow cytometry) or mean + standard deviation of technical replicates.

3.4 Investigation of chromatin barriers to targeted gene activation

3.4.1 Characterization of chromatin features at the Sox1 locus

VP64 can induce gene transcription by direct recruitment of Polymerase II or indirectly by recruiting other transcription factors (Wysocka and Herr 2003). It is therefore not unlikely that transcriptional barriers that safeguard a cellular identity during development impair the function of a trans-activator like VP64 even when it is properly targeted to a gene's promoter. To investigate the chromatin landscape at the promoter of Sox1 and identify such potential barriers, I sorted GFP positive and negative cells from NPCs expressing Sox1-targeting gRNAs. I then performed ChIP-qPCR to quantify regulatory chromatin marks at the target locus, as well as the promoters of an active gene (Gapdh) and a repressed gene (Oct4) as controls.

Tri-methylation of the Lysine 9 and Lysine 27 residue of Histone 3 (H3K9me3 and H3K27me3) is strongly correlated to gene repression when present in the promoter of a gene (Zhou, Goren, and Bernstein 2011). In line with this correlation, both marks were significantly enriched at the promoter of Oct4, compared to the promoter of Gapdh. However, the enrichment of the two marks at the Sox1 promoter differed strongly. While H3K27 tri-methylation levels were comparable to those at the active Gapdh promoter (Figure 14A) and did not vary between responsive and unresponsive populations, tri-methylation of H3K9 was enriched at the Sox1 promoter to levels comparable to those at the inactive Oct4 promoter. This however only held true for the Sox1^{GFP} negative cells, while levels in the Sox1^{GFP} positive cells were significantly reduced, and not anymore significantly different from those at the Gapdh promoter (Figure 14B).

3 Results

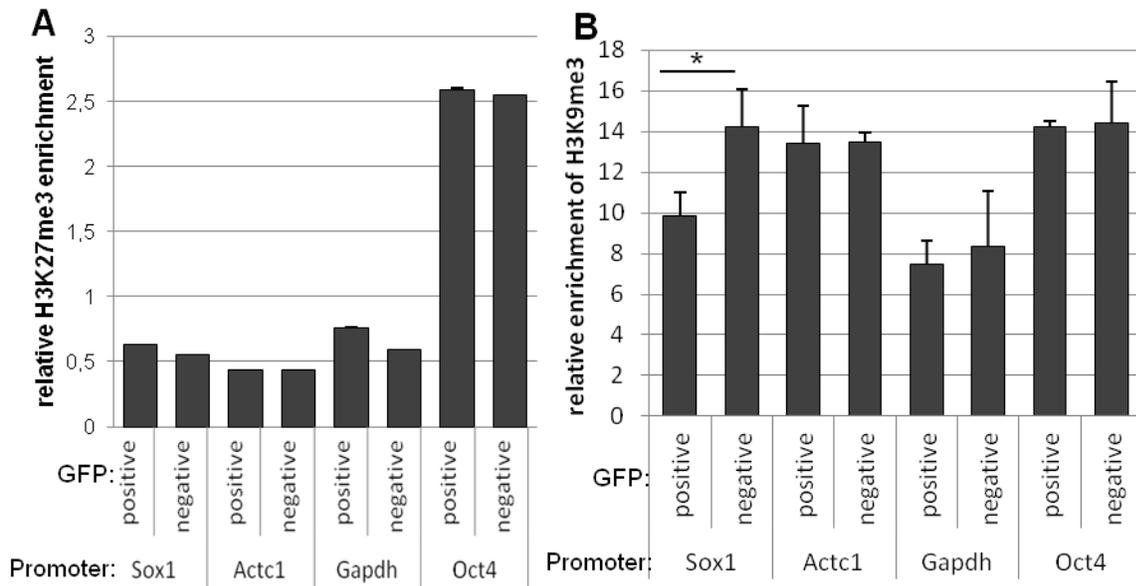


Figure 14: Quantification of histone modifications at the Sox1 promoter. **A** H3K27me3 was quantified via ChIP-qPCR. Levels at the Sox1 promoter were comparable to levels at promoters of actively transcribed genes. Data shown as mean + standard deviation of 3 technical replicates. **B** H3K9me3 was quantified via ChIP-qPCR. Sox1 negative NPCs exhibit levels at the Sox1 promoter comparable to those at the promoter of repressed genes. Levels at the Sox1 promoter in Sox1 positive NPCs are significantly lower. Data shown as mean + standard deviation of 3 biological replicates performed in different clonal lines on different days. * p-value < 0.05;

3 Results

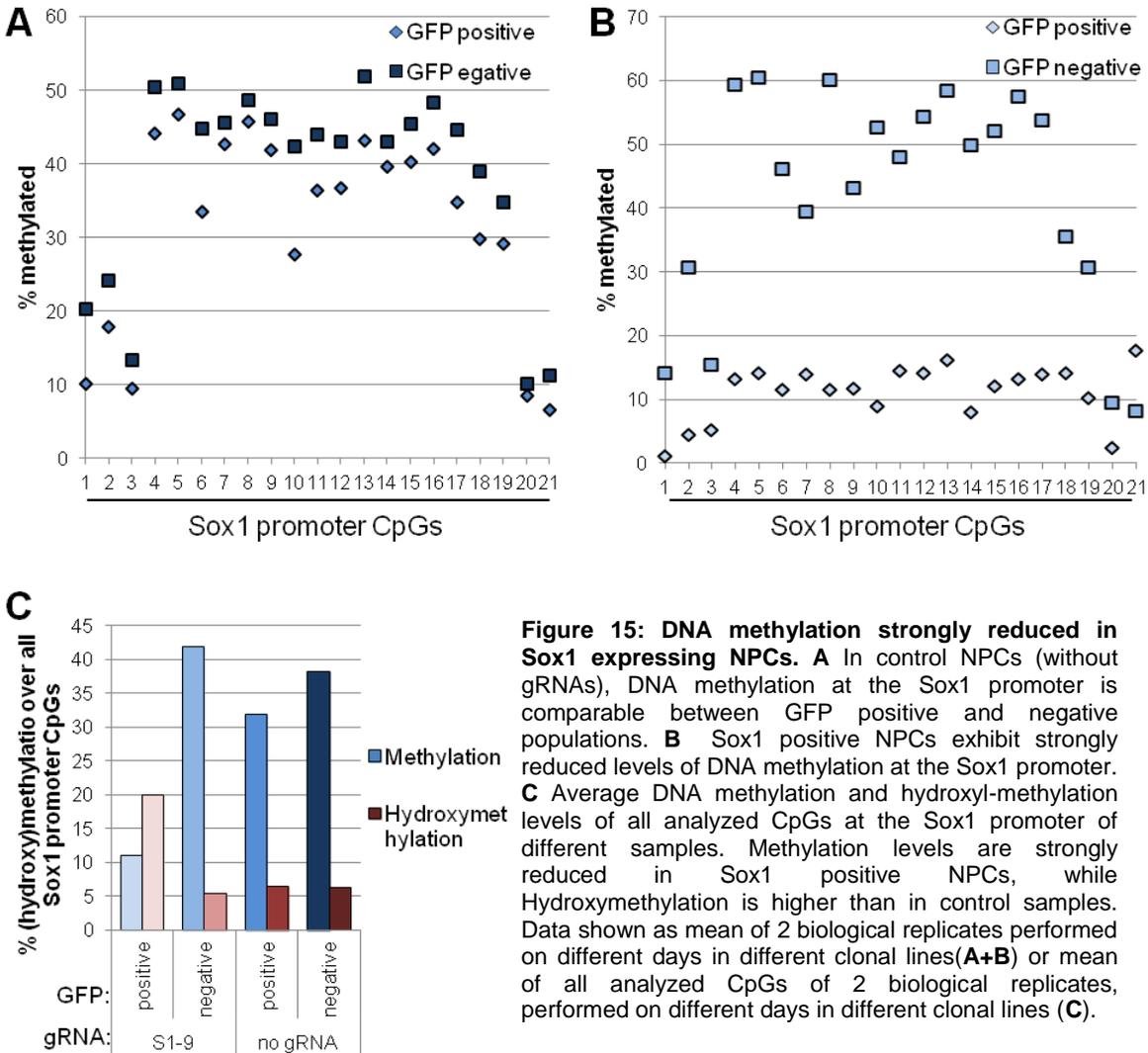


Figure 15: DNA methylation strongly reduced in Sox1 expressing NPCs. **A** In control NPCs (without gRNAs), DNA methylation at the Sox1 promoter is comparable between GFP positive and negative populations. **B** Sox1 positive NPCs exhibit strongly reduced levels of DNA methylation at the Sox1 promoter. **C** Average DNA methylation and hydroxyl-methylation levels of all analyzed CpGs at the Sox1 promoter of different samples. Methylation levels are strongly reduced in Sox1 positive NPCs, while Hydroxymethylation is higher than in control samples. Data shown as mean of 2 biological replicates performed on different days in different clonal lines (**A+B**) or mean of all analyzed CpGs of 2 biological replicates, performed on different days in different clonal lines (**C**).

A similar, inverse correlation between responsiveness to Sox1 induction and abundance of a repressive mark at the promoter was found for DNA methylation. I performed bisulfite and oxidative bisulfite sequencing to quantify levels of 5mC and its metabolite 5hmC, a modification that correlates to active transcription. Control NPCs without gRNA expression sorted for positive and negative GFP expression respectively exhibit strong and comparable levels of cytosine methylation (Figure 15A+C). As cells transduced with gRNA lentivirus and sorted for no GFP expression show similar levels of DNA methylation, VP64 targeting to the promoter alone does not seem to modify this particular mark (Figure 15C). However, methylation levels were strongly reduced at the promoter of NPCs with active Sox1 expression (sorted for GFP expression from NPCs

3 Results

transduced with S1-9 lentivirus), with all CpGs at the core of the promoter exhibiting methylation levels below 20% (between 14 and 16 CpGs were methylated to over 40% in the different control samples; Figure 15B+C).

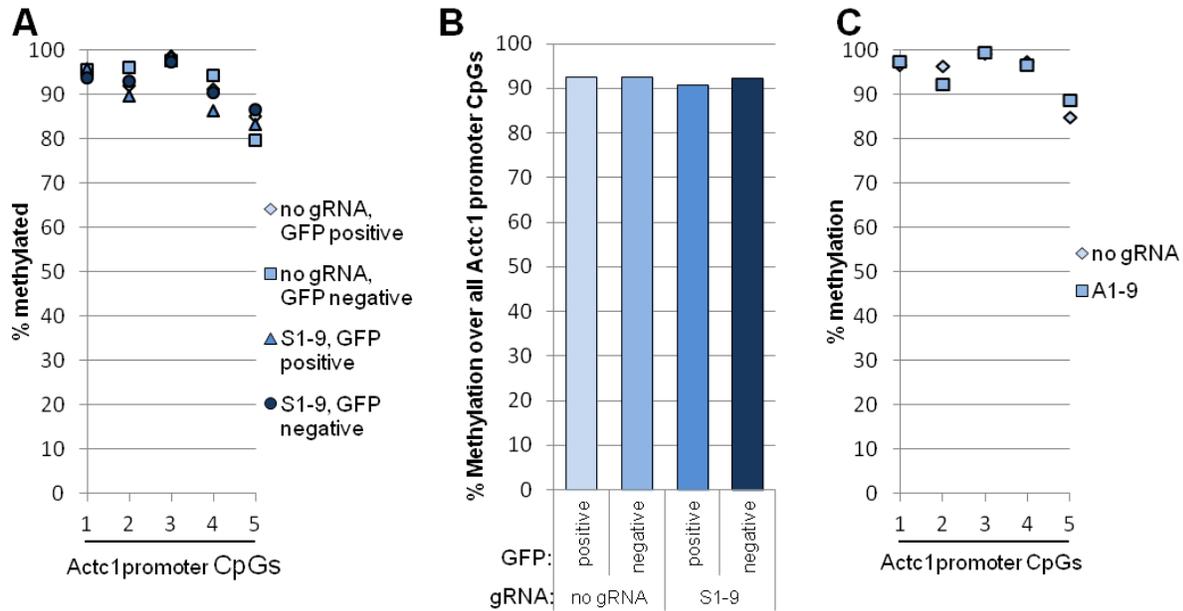


Figure 16: DNA Methylation at a control locus does not vary. **A+B** Methylation levels at single CpGs or averaged over all CpGs of the *Actc1* promoter do not vary between *Sox1* positive and negative NPCs. **C** Inducing *Actc1* does not lead to reduced DNA methylation levels at its promoter. Data shown as mean of 2 biological replicates performed on different days in different clonal lines;

Analysis of methylation levels at the *Actc1* promoter as control revealed similar and very high methylation levels, confirming that the effect observed at the *Sox1* promoter was indeed not an artifact of the sequencing method itself, but linked to the transcriptional state of *Sox1*. Surprisingly, even though the *Actc1* promoter is strongly methylated, activation with dCas9-VP64 was not compromised at this locus (Figure 16A+B). To compare the effect of VP64-induced gene transcription on methylation-levels at the two loci, I performed bisulfite sequencing at the *Actc1* locus in NPCs transfected with the A1-9 *Stagr* (see chapter 3.2.1). In contrast to *Sox1* activation, the induction of *Actc1* did not lead to reduction of methylation at its promoter (Figure 16C). This opposed effect could originate from the different densities of CpGs at the promoters of *Actc1* and *Sox1*.

3.4.2 Combining transcriptional editing with epigenome engineering enhances efficiency of gene induction

Next, I aimed to test whether the observed inverse correlation between responsiveness to Sox1 induction on the one hand, and H3K9me3 and DNA methylation on the other hand, might be the cause of the heterogeneous response to Sox1 induction, or merely the consequence of gene activation. To that end, chromatin modifying enzymes were fused to dCas9 in a similar manner as VP64 (see chapter 2.2.1.1). I transfected them into NPCs stably expressing Sox1-targeting gRNAs. Chromatin modifying enzymes used were Tet1, a DNA demethylase that catalyzes the oxidation of 5mC to 5hmC, Jmjd2a, a demethylase that removes trimethylation from the Lysine 9 residue of histone 3, and P300, an acetyltransferase that has been shown to bear activating potential (Eckner et al. 1994). By themselves, these enzymes activated Sox1 in a significant amount of NPCs when targeted with the S1-9 gRNAs, but their efficiency was significantly lower than the effect of VP64 (Figure 17A). I then transfected these constructs into NPCs stably expressing dCas9-VP64 and transduced Sox1 targeting gRNAs (S1-9). 8 days after transfection, I analyzed the amount of GFP positive cells via flow cytometry to test whether chromatin modulation in combination with a transcriptional activator would enhance gene induction.

Transfection of an expression plasmid for dCas9 without a fused effector domain did not have a significant effect on the amount of GFP positive cells after induction with VP64. While demethylation of K9 with Jmjd2a only increased the amount of responsive cells slightly (Figure 17B, not statistically significant as tested with Kruskal-Wallis-Test), this effect was more pronounced for P300 and most strongly Tet1. Demethylation of DNA at the Sox1 promoter simultaneous to VP64 targeting increased the amount of responsive cells on average 3-fold compared to VP64 alone. Furthermore, this effect was dependent on the catalytic activity of Tet1, as the catalytically dead mutant of Tet1 (dTet1) did not increase the amount of responsive cells significantly (tested with Kruskal-Wallis-Test) when combined with VP64. Furthermore, addition of an inhibitor of the DNA methyltransferase Dnmt3, Zebularine (Zhou et al. 2002), to NPCs expressing dCas9-VP64 and Sox1 targeting gRNAs had a significant effect on the amount of GFP positive cells (tested with Kruskal-Wallis-Test), corroborating the hypothesis that this

3 Results

effect is dependent on DNA demethylation. When added to NPCs without gRNA expression, this inhibitor did not lead to GFP positive cells, confirming that VP64 targeting to the Sox1 promoter is necessary to conduct this effect (Figure 17B).

To further confirm that in fact the promoter of Sox1 is demethylated by Tet1, but not by dTet1 targeting, I performed bisulfite and oxidative bisulfite sequencing on different NPC samples: a) as control, I used NPCs stably expressing dCas9-VP64, but no gRNAs; b) the same NPCs were transduced with S1-9 expressing lentivirus to target VP64 to the promoter of Sox1; c) dCas9-dTet1 and d) dCas9-Tet1 expression plasmids were additionally transfected and selected for. To investigate the demethylation efficiency independently of responsiveness to Sox1 induction by VP64, NPCs were not sorted, but analyzed as a bulk. As before, methylation levels at the Actc1 promoter did not differ between the different samples, indicating that possible changes at the Sox1 locus would not be an artifact of library preparation or sequencing. Methylation levels at the Sox1 promoter did not vary between the control sample (a) and the sample with VP64 targeting (b), and were at around 40-50%. This was expected, as the Sox1 positive population, exhibiting significantly lower methylation levels, only accounts for a small amount of NPCs in this sample (see Figure 6). Similar methylation levels were observed in NPCs with dCas9-dTet1 targeted to the Sox1 promoter, in addition to VP64; when however Tet1 instead of dTet1 was targeted to the locus, strong demethylation occurred (Figure 17C). This finding, combined with the significantly enhanced amount of responsive cells when Tet1 is added to VP64, supports the claim that methylation of DNA impairs gene induction by VP64, and that this effect can be reversed by targeted demethylation.

Ultimately, I analyzed whether cells that are initially unresponsive to Sox1 induction by VP64, but upregulate the target after additional targeting of Tet1 to the promoter, would exhibit a similar change in phenotype as cells that are responsive to the VP64 stimulus alone. To that end I performed a differentiation assay on NPCs, NPCs with VP64 targeted to the Sox1 promoter, and NPCs with additional Tet1 targeting the same promoter. After 7 days, not only was the amount of neurons significantly higher after Sox1 induction with VP64 alone, compared to the control, but further increased significantly when VP64 was combined with Tet1 (Figure 17D).

3 Results

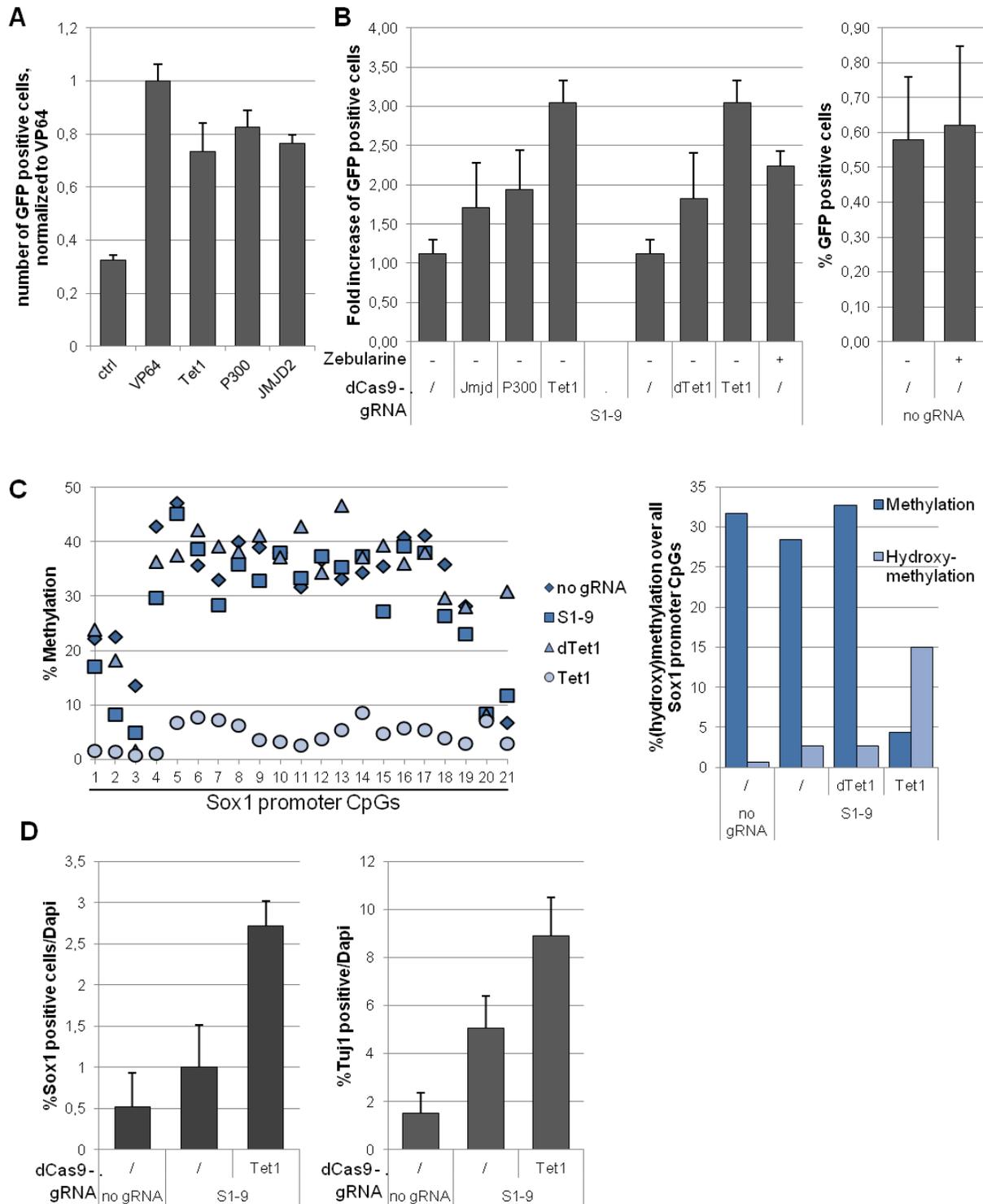


Figure 17: Combination of dCas9-VP64 and dCas9-Tet1 leads to an increase in the amount of responsive NPCs. **A** Targeting Chromatin modifying enzymes to the promoter of Sox1 in NPCs with dCas9 leads to induction of Sox1 in a minor proportion of NPCs. **B** Targeting VP64 in combination with chromatin modifying enzymes leads to a significant increase in the number of Sox1 positive cells, specifically for Tet1. This effect appears to be DNA demethylation dependent. **C** The combination of VP64 and Tet1 targeting leads to strong demethylation at the Sox1 promoter even in unsorted cells. **D** The number of Sox1 positive cells as assessed by ICC, as well as the number of neurons generated from unsorted NPCs is significantly higher when VP64 and Tet1 are targeted to the Sox1 promoter together. Data shown as mean + standard deviation of 3 (**A+D**) or 4 (**B+C**) biological replicates performed on different days in different clonal lines. * p-value < 0.05;

Taken together, these results suggest that DNA methylation is a primary barrier to targeted activation of Sox1 expression in NPCs. By demethylation of DNA at the target site, this barrier can be overcome, as significantly more cells respond to the activation, when Tet1, but not dTet1 is targeted to the Sox1 promoter in addition to VP64. Moreover, these additional responders exhibit a similar change in phenotype as assessed by neuronal differentiation.

3.4.3 DNA methylation as barrier to transcriptional engineering is not exclusive to Sox1

Investigating the response of NPCs towards the activating stimulus of a TAF, I revealed DNA methylation to be a strong transcriptional barrier at the Sox1 promoter. It is however not clear, whether this phenomenon is specific for Sox1, or whether the same chromatin modifications are also relevant at other promoters. I therefore aimed to analyze potential epigenetic barriers at other developmental transcription factors that would be detrimental to the NPCs' identity if activated mistakenly.

I analyzed DNA methylation levels at the promoter of six different transcription factors associated with cell fate determination or used in reprogramming. These were Il1rn (a marker of progenitors from the hematopoietic lineage (Arend et al. 1998)), MyoD1, for reprogramming to muscle cells (Davis, Weintraub, and Lassar 1987), NeuroD4 and Neurogenin2 (Ngn2), both used for direct reprogramming to neurons (Berninger et al. 2007, Masserdotti et al. 2015), Nkx2-2, which is necessary for differentiation into pancreatic beta cells (Sussel et al. 1998), and Oct4, used amongst other factors for the generation of induced pluripotent stem cells (Takahashi and Yamanaka 2006). Four out of six of the analyzed promoters were strongly methylated, with varying densities of CpGs in the proximity to the TSS (Il1rn, MyoD1, Nkx2-2, Oct4), while NeuroD4 and Ngn2 exhibited very low levels of the DNA modification (Figure 18B).

I designed two gRNAs for each of the six promoters, each pair at least 100bp apart, and cloned them pair wise into gRNA expression vectors as described in chapter 2.2.1.2. I then transfected these plasmids into NPCs expressing dCas9-VP64. To

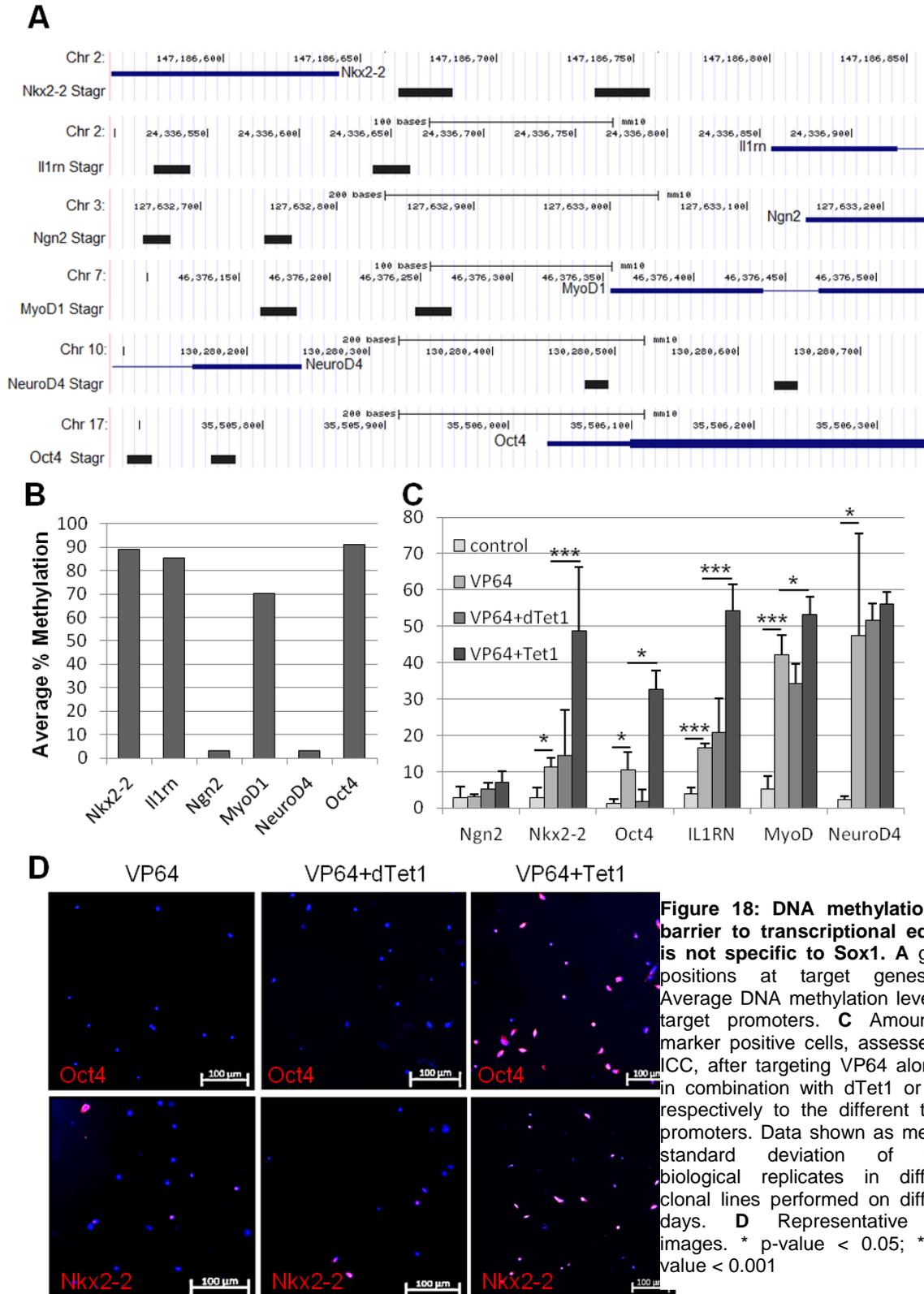
3 Results

analyze the amount of cells that upregulated transcription of the respective target, I performed ICC and counted the amount of NPCs positive for each target. The number of responsive cells varied strongly for the different targets, ranging from below 5% positive cells (Ngn2) to around 50% positive cells (NeuroD4), as depicted in Figure 18. However, the amount of responsive cells did not seem to correlate with the levels of DNA methylation, as the two targets that were unmethylated produced the strongest and weakest effect respectively.

To test whether DNA demethylation could increase the amount of cells upregulating the different targets, I transfected dCas9-Tet1 and dCas9-dTet1 as control, in addition to the respective gRNA expression constructs, to NPCs expressing dCas9-VP64. Cells were selected for expression of the transfected dCas9 fusion proteins and analyzed 8 days after transfection. Additional targeting of dCas9-dTet1 did not have a significant effect on the number of responsive cells, as determined by ICC, for any of the targets (Figure 18C+D). Targeting dCas9-Tet1 to the different promoters did however significantly increase the number of positive cells for some, but not all of the targets. Interestingly, the number of NPCs that upregulated the expression of NeuroD4 and Ngn2 respectively, did not significantly increase with additional targeting of the DNA demethylase (Figure 18C+D). The reason for this effect most likely lies in the low levels of methylation at those promoters. Gene induction efficiency by targeting the other four, strongly methylated promoters, did however change upon presence of Tet1. In each case, the amount of responsive NPCs increased significantly when the promoter of the target was demethylated with Tet1 (Figure 18C+D).

Taken together, these results show that DNA methylation is acting as a barrier to transcriptional activation not only at the Sox1 promoter, but also at the promoters of other developmental transcription factors. This regulatory mechanism however presupposes DNA methylation at the respective target region. In case of Ngn2 and NeuroD4, demethylation of the target DNA did not increase the response of NPCs – it is nevertheless noteworthy that in the case of Ngn2, there seemed to be a different barrier to activation, since even in the absence of DNA methylation, only a small subset of NPCs reacted to transcriptional activation at this promoter.

3 Results



3.5 Identification of regulatory domains at the Sox1 locus

It is clearly established that gene promoters and distal regulatory elements like enhancers play important roles in the regulation and activation of genes. These regions can be predicted based on the chromatin state and certain sequence motifs, but clear definitions of what constitutes such regulatory elements of genes are still lacking. We therefore aimed to characterize the Sox1 locus in a broad range and identify potential enhancers and repressors by epigenetic screens.

For this approach, I designed 6000 gRNA sequences targeting the Sox1 locus in a range of 200kb in the mouse genome (mm10, genomic coordinates chr8:12,295,501-12,495,500; Figure 19). As quality control for screening approaches, I included 500 negative gRNA sequences, targeting either a different genome or genes unrelated to Sox1, or lacking a PAM. These sequences were purchased from a company and cloned as a pool into lentiviral gRNA backbones (as described in chapter 2.2.1.2), and a lentiviral library, from here on denoted as Sox20000. I employed this library in two different approaches using either VP64 or different chromatin modifying enzymes tethered to dCas9 as effector domains.

3.5.1 VP64 Screen: Identification of distal regulatory elements

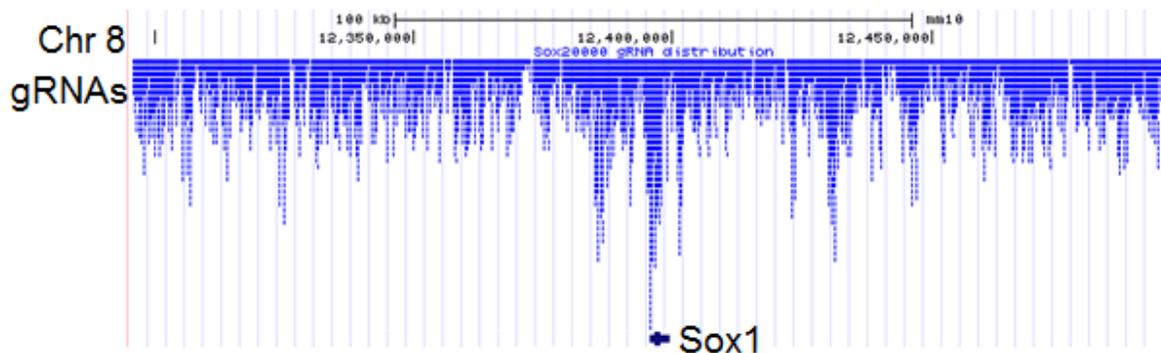


Figure 19: Sox20000 library positions. Distribution of gRNAs in the Sox20000 gRNA library, covering the Sox1-locus over 200000bp.

3 Results

In a first screen, VP64 was employed to validate the reliability of the screening approach and to identify candidate distal regulatory domains for Sox1. Sox1^{wt/GFP} NPCs stably expressing dCas9-VP64 were transduced with the Sox20000 library lentivirus at an MOI of 0.4. This approach ensured a high amount of cells receiving only one, not multiple gRNAs, limiting potential false positive reads by contamination of cells with negative gRNAs. To prevent biases in the delivery of gRNA sequences, 1.5 million cells were transduced, so that at the given MOI around 100-fold representation of each gRNA sequence in the library would be achieved. Cells were selected for gRNA expression as described in chapter 2.2.1.2, expanded, and GFP expression was analyzed via Flow cytometry 7 days after gRNA transduction (Figure 20A). At this stage, transduced NPCs did not differ strongly from untransduced control samples (Figure 20B). To enrich GFP expressing cells, 10 million NPCs were sorted with the sorting gate set to the highest 5% cells in GFP expression of the control sample. Sorted cells were expanded again to 10 million cells and subsequently resorted to further enrich for positive cells and deplete false positives. In the second sort, the sorting gate was set to the highest 2.5% GFP-positive cells in the untransduced control. NPCs were again expanded and sorted a third time, with the sorting gate was set to the highest 0.5% GFP-positive cells in the control sample (Figure 20A). The continuous narrowing of the sorting gates allowed enrichment of strictly Sox1^{GFP}-positive NPCs (Figure 20A). From each sorting step, including unsorted NPCs, 1 million transduced cells were subjected to sequencing. DNA was isolated, the gRNA locus amplified, and the sequencing library prepared as described in chapter 2.2.5.3. Figure 21A shows the bioanalyzer trace of an exemplary sequencing pool.

3 Results

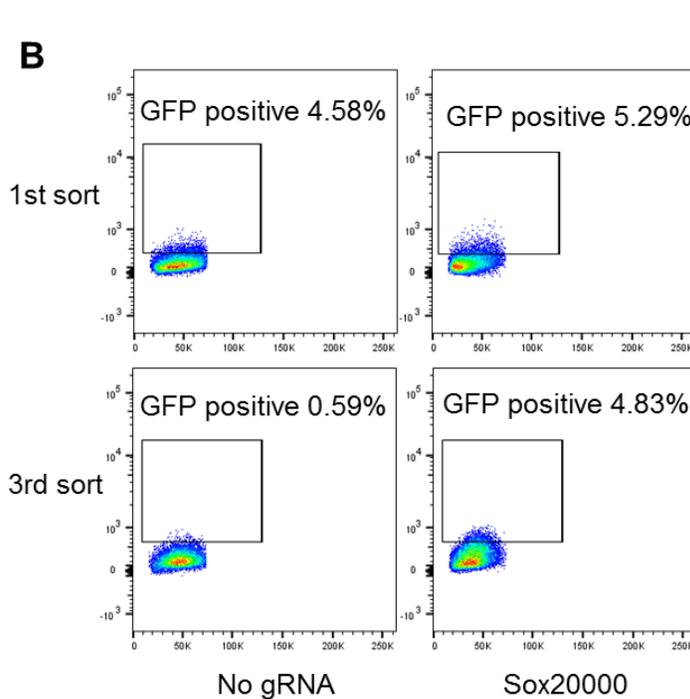
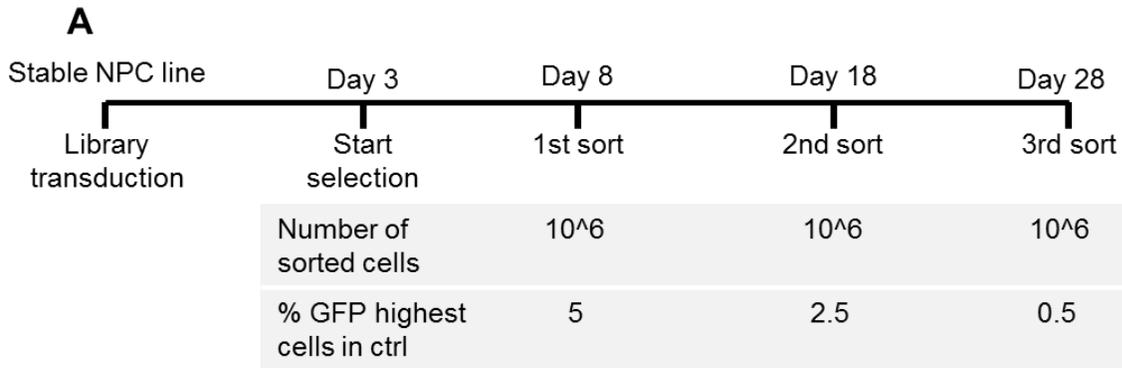


Figure 20: Enrichment of Sox1^{GFP}-positive cells over three subsequent sorts. **A** Experimental setup for all screens. NPCs stably expressing a dCas9-effector fusion protein were transduced with the gRNA library, and sorted three times with increasingly stringent gates. **B** At the first sort, transduced and untransduced samples were comparable, while transduced samples exhibited a clear enrichment of GFP-positive cells at the time point of the third sort.

To assess the quality of the gRNA library, the reads from unsorted NPCs were demultiplexed, aligned to the mouse genome, and the relative abundance of each gRNA sequence calculated. Only one sequence from the *in silico* designed gRNA pool was absent in all samples. Furthermore, the reads exhibited distribution over the locus comparable to the theoretical design. This indicated that no large biases were introduced to the library during cloning, preparation of the virus, or transduction of NPCs (data not shown).

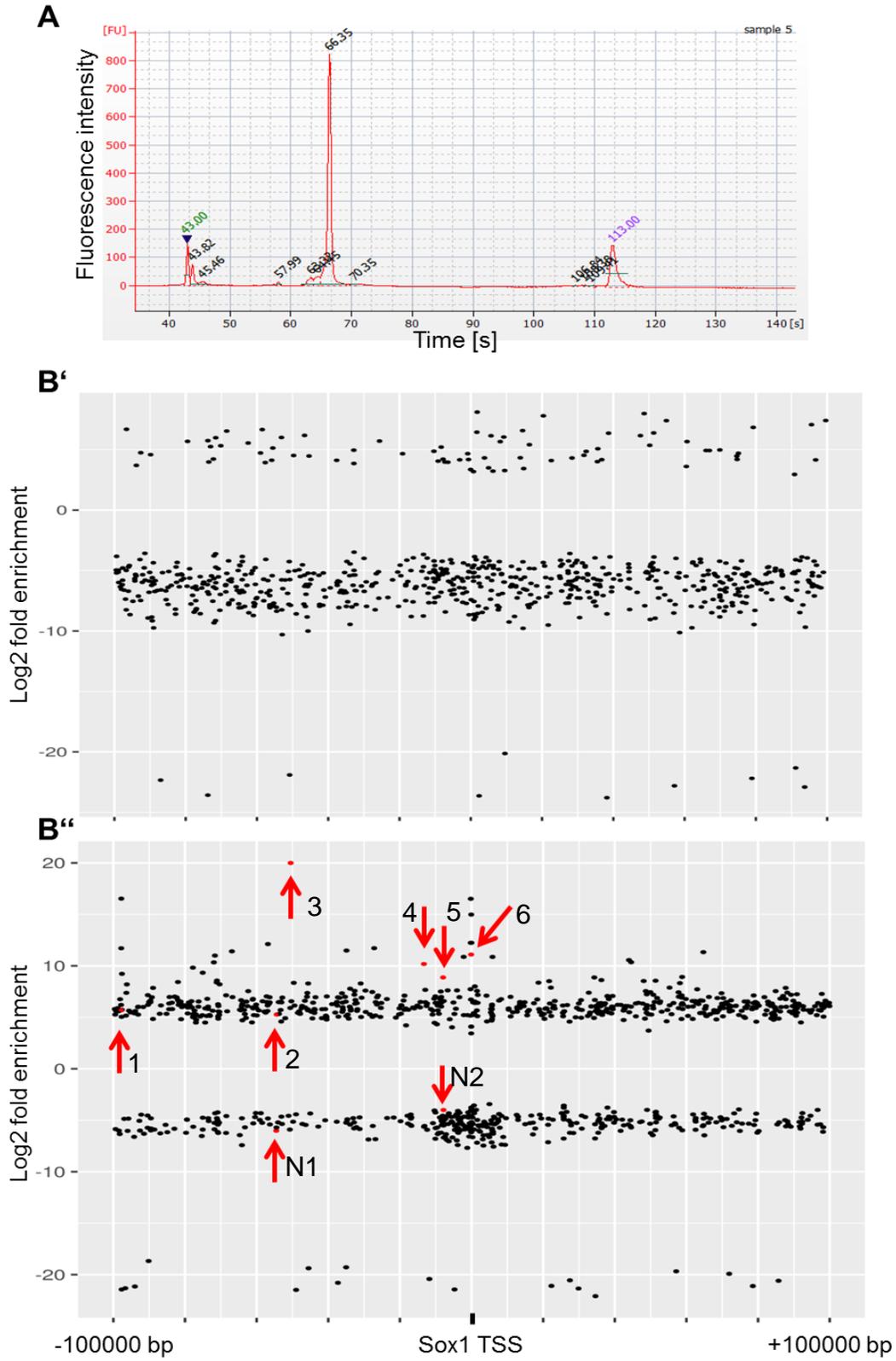
3 Results

To further investigate the quality of the FACS setup for the screen, reads from triple sorted samples were analyzed for the presence of negative control sequences. Only one negative gRNA sequences was significantly enriched over the control samples, while all others were virtually absent after three sorting steps. Interestingly, the enriched gRNA sequence was a Sox1 targeting sequence with a missing PAM (see library design, chapter 2.2.1.2). In all, this supported the enrichment of truly GFP-positive cells by subsequent sorts.

To determine potential candidate gRNAs activating transcription of Sox1, gRNA sequences with a significant increase in their relative abundance between the unsorted samples and one or three sorting steps respectively were identified. The enrichment of each significantly regulated gRNA was plotted versus the relative location to the Sox1 TSS (Figure 21B). This revealed a uniform enrichment of gRNAs over the whole analyzed region, with no obvious candidate regulatory regions. This included the promoter region, a target that has been shown to induce Sox1 expression (chapter 3.2).

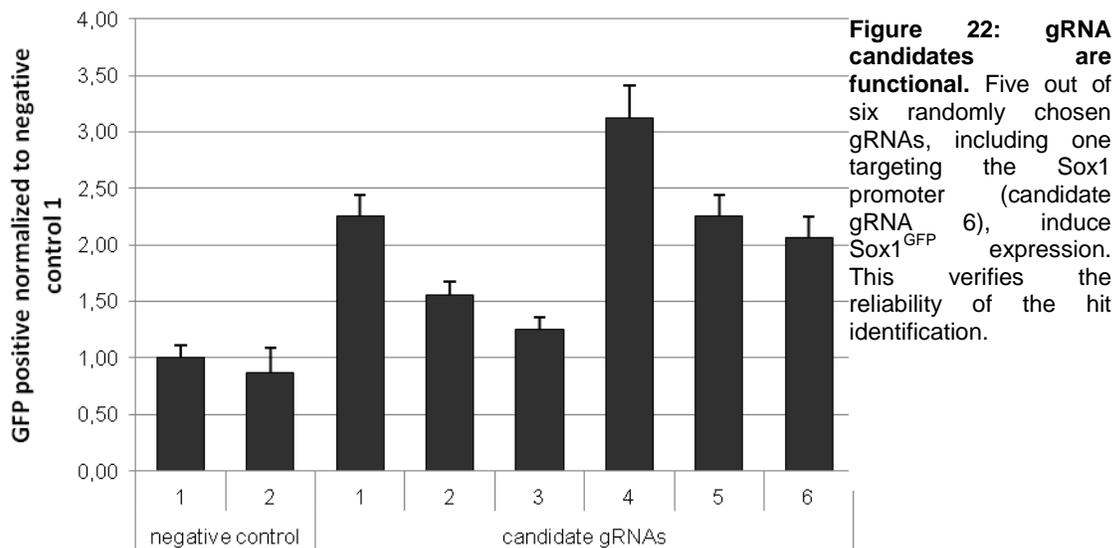
Figure 21 (next page): gRNAs are equally enriched over the Sox1 locus. **A** The bioanalyzer file of the sequenced pool shows a single peak at the expected size (261bp). **B** Enrichment of significantly regulated gRNAs after one sort (**B'**) or three sorts (**B''**) did not reveal obvious regulatory elements at the Sox1-locus. Candidate gRNA sequences for hit verification are highlighted in red and with red arrows.

3 Results



3 Results

Five gRNA sequences from the triple sorted, significantly enriched sequences were chosen randomly in addition to one significantly enriched gRNA at the Sox1 promoter. These sequences were cloned as single gRNAs into expression plasmids as described in chapter 2.2.1.2. In addition, two gRNA sequences that were significantly depleted after three sorts were cloned as negative controls. To verify the functionality of the gRNA hits, each of the plasmids was transfected into dCas9-VP64 expressing NPCs, and the amount of GFP positive cells measured after 8 days via flow cytometry. Five out of six gRNAs led to higher numbers of GFP positive cells compared to the negative control gRNAs. This indicated that even though the enriched gRNAs seemed to be randomly distributed, enriched sequences were in fact functional.



Taken together, these results show that by triple sorting of transduced NPCs for GFP expression, false positive hits can be depleted, and functional sequences identified reliably. Based on the chosen way to analyze sequencing data, an identification of the most promising hits or candidate regulatory regions is difficult, because no obvious candidate regions appear in the pattern of significantly enriched gRNAs. The analysis could therefore profit from being refined; a more stringent selection of candidate gRNAs could potentially reveal clusters by depletion of hits in less potent regions. In addition, the investigation of unsorted NPCs reveals that almost all sequences designed *in silico*

are present in the actual lentiviral pool, and that no strong biases were introduced during library preparation or transduction.

3.5.2 Future Experiments

To identify distal elements that regulate Sox1 expression through chromatin features, I performed the screen with the same library in NPCs expressing dCas9 tethered to one of three chromatin modifying enzymes. These can manipulate chromatin features and change modifications to a state correlated to active gene transcription by either demethylation of H3K9 by Jmjd2a, methylation of H3K4 by Set7 or acetylation by p300. These enzymes have been shown to induce transcription when targeted to the promoter of a gene (Figure 16), but it is not yet clear whether the same manipulations in distal regulatory elements can achieve the same effect.

To that end I performed the screens in the same way as described for dCas9-VP64. Briefly, NPCs were transduced with the gRNA library virus at an MOI of 0.4 and sorted three subsequent times to enrich for GFP positive and deplete negative cells. From each sorting step, 1 million cells were subjected to sequencing of the integrated gRNA locus to identify targets of the chromatin modifiers that lead to gene induction.

While the analysis of the sequencing data is still pending, this allows to identify regulatory chromatin features in the locus of Sox1 and might ultimately even help to further the understanding and definition of regulatory elements.

4 Discussion

4.1 Sox1 is instructive for neuroepithelial cell identity

Many members of the Sox family of transcription factors have been shown to be potent master transcription factors playing fundamental roles during development. While Sox10 for example is important for neural crest development (Kelsh 2006), recent studies on Sox2 have described its capability to induce pluripotency, even on its own (Liu et al. 2018), when induced in somatic cells. Furthermore, some of them are pioneering factors and as such able to bind even tightly compacted chromatin to induce gene transcription (Kamachi and Kondoh 2013). Sox1 on the other hand, another member of the same subfamily as Sox2 (SoxB1), is relatively understudied: It is clear that its expression pattern during embryonic development is highly specific for the neural lineage, and that it marks a population of neurogenic progenitor cells in the adult subventricular zone and hippocampus (Aubert et al. 2003, Venere et al. 2012). Furthermore, overexpression of Sox1 in NPCs *in vitro* significantly increases neuronal differentiation potential (Kan et al. 2004). Based on these findings, and considering the potential of other Sox transcription factors, an instructive role for Sox1 during development seems likely. Direct proof for its pioneering potential or reprogramming capabilities is however still lacking, and it is not clear, whether it classifies as master transcription factor, able to impose a NSC identity on a cell by overexpression or other means.

Here, I investigated the impact of Sox1 induction in strictly Sox1 negative NPCs, corresponding to a developmental stage of glial progenitor cells. These cells differ from Sox1 expressing NSCs in their transcriptional patterns, their morphology, and their differentiation potential. qPCR, flow cytometry, ICC, and Western Blot analysis confirmed strong expression of Sox1 in NPCs after targeting of the trans-activating domain VP64 to its promoter, and that the majority of cells that activated Sox1 expression did not lose it again over time. This is reminiscent of reprogramming factors that have been shown to introduce a new cell identity stably, even when the expression plasmid of the initial reprogramming factor is lost again (Hobert 2008). Reprogramming

approaches based on transcriptional induction of reprogramming factors rather than ectopic overexpression moreover showed a fundamental remodeling of chromatin features at the promoters of the respective reprogramming genes (i.e. H3K27 acetylation and H3K4me3)(Black et al. 2016, Liu et al. 2018). I observed a similar remodeling also at the Sox1 promoter, albeit for other features, indicating a substantial intrusion in the cell intrinsic transcriptional regulation. Despite these similarities, I can however not conclude the behavior of NPCs upon loss of targeted Sox1 induction from my experiments, because I used stable expression of all components of the TAF during all long-term experiments.

Furthermore, when differentiated for seven days in an undirected manner, NPCs expressing Sox1 gave rise to a significantly larger amount of Tuj1 positive immature neurons, compared to control NPCs. This is in line with the overexpression phenotype that was earlier reported for Sox1 and confirmed the functionality of the employed TAF. Further analysis of the subtype of the generated neurons after prolonged differentiation revealed that around 70% are positive for the marker of glutamatergic excitatory neurons, vGlut1, while ca. 15% expressed Calbindin, a marker of gabaergic inhibitory neurons. During embryonic development, both of these subtypes originate in the ventricular zone, and are generated from Sox1 positive aRGCs during the second wave of neurogenesis, direct descendants of NSCs (Jiang and Nardelli 2016). In line with these findings, Zhu and colleagues have shown that Sox1 positive neuroepithelial cells derived from human ESCs give rise to a similar proportion of glutamatergic to gabaergic neurons (76%±3.1% and 21.5%±8.9% respectively) (Zhu et al. 2016). Furthermore, the proportion of produced neurons of the two subtypes corresponds to neural development, where a majority of newborn neurons are radially migrating excitatory glutamatergic neurons (Marín and Müller 2014). The resemblance of early born neurons indicates that the Sox1 expressing NPCs that gave rise to those neurons were indeed, at least partially neuroepithelial in character. This in turn underlines the potential cell fate instructing role of Sox1.

The results of long term differentiation of NPCs were corroborated by the results of the transcriptome analysis following Sox1 induction. Over 1500 genes exhibited significant differential expression with a more than 4-fold change in transcript levels,

indicating a pronounced shift in cell identity induced by Sox1 expression. A more detailed analysis of the evoked transcriptional changes revealed a resemblance to the differences between truly Sox1 positive NSCs and Sox1 negative NPCs. This effect was in particular observed for significantly upregulated genes following Sox1 induction. Most of these transcripts were also significantly higher in NSCs and NRs, the *in vitro* correlate of the cells lining the neural tube, NECs, compared to NPCs. Many of the upregulated genes have been shown to be specific for Sox1^{GFP} positive in the E10.5 embryo (e.g. Nestin, Vimentin (Aubert et al. 2003)). Moreover, four out of six neuroepithelial markers that are specifically expressed in neural stem cells during early embryonic development (Nocht1, Nestin (Higuchi et al. 1995, Lendahl, Zimmerman, and McKay 1990)), or different epithelial cells throughout development, including neuroepithelial tissue (Occludin, Zo-1 (Hirase et al. 1997, Aaku-Saraste, Hellwig, and Huttner 1996)), were significantly enriched on protein level in the Sox1 positive NPC population. Taken together, these results reveal a pronounced change in transcriptional patterns of NPCs elicited by Sox1 expression. This change indicates a shift in the cells' identity, and this newly instructed identity exhibits attributes that are characteristic for NSCs. This in turn implies a cell fate determining role for Sox1 and suggests that it indeed acts as a developmental master transcription factor.

It is however noteworthy that the shift of NPCs towards a NSC-like identity was only partial. Firstly, even though four neuroepithelial markers were significantly upregulated following Sox1 induction, as assessed by ICC, two other markers were still not expressed at detectable levels. These markers were E-Cadherin, a cell-cell adhesion glycoprotein that is specifically expressed in epithelial cells, and Cd133, which in the neural lineage is specific for early neuroepithelial cells (Uchida et al. 2000). Furthermore, the transcriptome data reveal that even though the majority of the 100 most upregulated genes in NPCs after Sox1 induction resembles those of NSCs, this similarity was far less pronounced in the 100 most downregulated genes. In addition, levels of upregulated genes in NPCs did often still not reach the levels in NSCs. Hierarchical clustering and principal component analysis of transcriptomes exhibited that the developmental Sox1 expressing stage (i.e. NSCs and NRs) clustered apart from NPCs with induced Sox1 expression.

Taken together, I was able to substantiate the hypothesis that Sox1 is a master transcription factor with the potential to instruct a shift in the identity of NPCs. This shift was, albeit directed towards NSCs, only partial. The reasons for this could be manifold. It has for example been shown to be beneficial for reprogramming efficiency to combine various transcription factors for optimized efficiency. The efficiency of inducing pluripotency for example is lower when using only Sox2 or Oct4, compared to the original combination of four factors (Takahashi and Yamanaka 2006, Liu et al. 2018, Kim et al. 2008). Likewise, first approaches in neuronal reprogramming relied on the combination of Ascl1, Brn2, and Myt1l to reliably generate neurons from fibroblasts, an approach that is still being used up-to-date (Vierbuchen et al. 2010, Black et al. 2016). Similarly, it could be that Sox1 alone is not sufficient to completely convert Sox1-negative NPCs to NSCs. Furthermore, even though stronger induction using VPR did not help in increasing the number of responsive cells (as shown in Figure 13D), it did induce Sox1 to higher levels. RNAseq revealed that Sox1 levels in the Sox1^{GFP} positive is significantly higher than in the baseline, but also still significantly lower than in Sox1 positive neural stem cells (data not shown). Higher levels that could be reached by a stronger tool could therefore also help to increase the extent of the phenotypic shift. It seems however obvious that the strong potential of Sox1 can be detrimental to neural development if errors in its expression arise. This is in line with a study that linked high levels of Sox1 with different neurological disorders. Even though the prevalence of patients with increased Sox1 levels was low, it indicates a role of Sox1 not only in neural development, but also in disease (Berger et al. 2016). It therefore seems likely that Sox1 expression is tightly regulated to interfere with erroneous expression.

4.2 DNA methylation serves as transcriptional barrier at several master transcription factors

Since the development of dCas9 as a shuttle for effector domains in 2013, the system has been widely used to deliver trans-activators to gene promoters for targeted gene induction (Black et al. 2016, Chavez et al. 2015, Liu et al. 2018, Perez-Pinera et al. 2013). The corresponding studies soon revealed that the effect of targeted activation varies strongly (to about 4 magnitudes between very inefficient and highly efficient

trans-activation) between different target genes, but the molecular mechanisms leading to the observed discrepancies have not been investigated so far. And despite recent studies using TAFs to induce transcription of endogenous reprogramming factor genes, most approaches to date employ ectopic over-expression of reprogramming factors, thusly circumventing barriers at the respective promoters. This work is therefore one of the first studies to examine the effect of targeted gene induction on single cell level, and to analyze the cell-intrinsic barriers that are the basis for the heterogeneous response to the targeted induction of a master transcription factor.

Strikingly, my results demonstrated a strong reluctance of NPCs to targeted Sox1 induction by dCas9-VP64. This came as a surprise, because the extent of the evoked changes in responsive NPCs demonstrates that the employed TAF possesses high activating potency. Still, VP64, and even the more potent VPR were not able to overcome the transcriptional barriers present in unresponsive cells. On the one hand, these findings make low efficacy as reason for the unresponsiveness in the majority of NPCs unlikely (as discussed below in more detail). On the other hand, they further substantiate the hypothesis that as a master transcription factor, Sox1 needs to be tightly regulated to withstand erroneous activation in the wrong settings.

By separating Sox1^{GFP} positive and negative cells I was able to investigate the chromatin landscape at the promoter of Sox1 and deduct potential causes for the discrepancy in the response to Sox1 induction. I analyzed three different chromatin modifications that are by and large correlated to repressed genes: H3K27me3 is a mark that is regulated in a rather dynamic manner and is set at gene promoters already shortly after silencing (Ferrari et al. 2014). It can furthermore be found at the promoter of poised genes; these genes carry bivalent chromatin features at their promoters (most commonly H3K4me3 and H3K27me3), a condition that allows for quick gene induction (Vastenhouw and Schier 2012). In contrast, H3K9me3 is a mark for highly compacted and silenced heterochromatin and correlates to a more substantial gene repression (Schotta et al. 2004). The relation between DNA methylation and gene transcription is more complex; in general DNA-demethylation at a gene promoter is correlated with expression of the respective locus (Feng et al. 2010, Zemach et al. 2010). Accordingly, low levels of H3K27me3, paired with higher levels of H3K9me3 at the Sox1 promoter

indicate a robust repression of this gene. This finding is further underlined by the fact that Sox1^{GFP} positive NPCs have significantly lower levels of the latter mark at the Sox1 locus.

I further focused on DNA methylation as a potential barrier to gene induction, because there was a pronounced difference between responsive and unresponsive cells in DNA methylation levels at the Sox1 promoter. The comparable levels of DNA methylation at the Actc1 promoter before and after gene induction by VP64 suggests that the difference at the Sox1 locus is not simply the consequence of the presence of VP64 at a gene promoter, or of active transcription, but might indeed be causal for the response in NPCs with low DNA methylation levels. Moreover, combining the targeting of dCas9-VP64 with dCas9-Tet1, a DNA demethylating enzyme, led to a significantly increased proportion of Sox1^{GFP} positive NPCs, an effect that was a) dependent on the catalytic activity of Tet1, as dTet1 did not elicit any effect, and b) was reproduced by the inhibition of DNA methyltransferases Dnmt1 and 3 by application of Zebularine, underlining the necessity of DNA demethylation for the effect of Tet1. While DNA methylation has been correlated to gene repression, and targeted demethylation can de-repress genes (Liu et al. 2016), this study is the first to show direct proof that DNA methylation interferes with targeted gene activation. These findings indicate that DNA methylation plays an important role in the upkeep of specific Sox1 expression patterns during development, a finding that corroborates recent implications of this particular chromatin mark in neural cell fate decisions (reviewed in (Stricker and Götz 2018)). Several studies discussed in this review implicate that temporal change of DNA methylation in NSCs during embryonic development lead to changes in the expression of cellular programs (Takizawa et al. 2001, Sanosaka, Namihira, and Nakashima 2009, Lee, Hore, and Reik 2014). This in turn leads to a temporal resolution of lineage commitment and has been suggested to be the reason for e.g. the late onset of gliogenesis (Takizawa et al. 2001).

Importantly, I discovered a similar role of DNA methylation in the regulation of additional master transcription factors, two of which have been successfully used in reprogramming, namely MyoD1 and Oct4 (Davis, Weintraub, and Lassar 1987, Takahashi and Yamanaka 2006). It is therefore obvious that this principle is not specific

for Sox1 or the neural lineage. My findings could thus be of broad interest in the field of reprogramming, because an increasing amount of reprogramming approaches relies on targeted induction of reprogramming genes by trans-activators tethered to dCas9, rather than ectopic expression of the reprogramming factors. DNA methylation, but also other chromatin features (see also next chapter) can potentially interfere with the transcriptional activation of reprogramming factors and thereby be detrimental to the efficiency of reprogramming in a similar manner as observed in this study. In such cases, combination of transcriptional engineering with epigenetic editing would be a major benefit, and might even render the need for exceedingly potent TAFs unnecessary (as discussed below). The result of targeted trans-activation of NeuroD4 respectively shows however that this combination is not necessarily needed (as NeuroD4 can readily be activated by VP64 alone). In any case, DNA methylation at the target region is a substantial factor for the impact of Tet1 on gene induction.

Of note, the mechanism that underlies the ability to resist targeted gene induction by high DNA methylation levels still remains unknown. It is however clear that VP64 recruits cell intrinsic proteins, such as other transcription factors, to facilitate transcriptional activation (Wysocka and Herr 2003). I therefore searched the Sox1 promoter region for predicted transcription factor binding sites and indeed binding motives of several transcription factors are present. Further analysis revealed that of these transcription factors some have been shown to exhibit DNA methylation sensitive binding, and to play important roles during neural development, namely YY1, Sp1, and E2F-1 (Campanero, Armstrong, and Flemington 2000, Cooper-Kuhn et al. 2002, Zhu et al. 2003, He et al. 2011, He and Casaccia-Bonnel 2008). Interestingly, the promoters of Nkx2-2 and Pou5F1 also feature binding motives for YY1 or YY1, Sp1, and E2F-1 respectively. RNAseq revealed that these factors are indeed expressed in all investigated states of NPCs and NSCs. Furthermore, while YY1 and Sp1 are significantly lower in NSCs, NRs, and Sox1^{GFP}-positive NPCs compared to baseline NPCs, E2F-1 behaved the other way around (Data not shown). Taken together, this suggests potential common actors that mediate trans-activation dependent on DNA methylation levels. Even though other factors might be relevant in the context of DNA methylation barriers, these transcription factors might be good candidates to further

investigate; ChIP-qPCR for example could give important cues on whether their recruitment is dependent on demethylation their respective binding sites.

Another point worth noting in regard of DNA methylation as transcriptional barrier is the density of CpGs at the different targets analyzed in this thesis. NeuroD4 and Actc1 both have a low number of CpGs in the investigated region in specific, and the promoter locus in general. Both genes could be readily activated, and interestingly independently of the methylation levels. Furthermore, gene activation did not entail Demethylation at the Actc1 locus. This is in line with studies that insinuate only little functional relevance to DNA methylation in CG-poor promoters (Schübeler 2015). The promoters of Sox1, Oct4, Ngn2, and Nkx2-2 on the other hand did exhibit a higher density of CpGs. These four loci displayed a significantly restricted response to targeted gene induction, again independent of DNA methylation levels, a finding that has so far not been shown in the literature. Noteworthy, two of the investigated loci even have annotated CpG-islands, which are DNA regions with a minimum C+G content (e.g. 50%) and CpG density over a certain DNA length (e.g. 400-500bp) (Illingworth and Bird 2009). These thresholds are however arbitrarily defined, and even though Oct4 and Nkx2-2 do not possess annotated CpG islands, they exhibited similar behavior to targeted activation as those promoters with annotated CpG islands (Sox1 and Ngn2). It might therefore be necessary to refine the definition of these regulatory elements to class them according to their function. It needs to be stressed however that based on the small number of analyzed promoters, no significant pattern can be deducted, and further investigations are required to support this claim.

4.3 Additional epigenetic barriers against activation of transcription

Interestingly, even though demethylation at the Sox1 promoter led to a significant increase in trans-activation efficiency, a major amount of NPCs still remained unresponsive to Sox1 induction, even though DNA methylation level dropped to below 5% in all cells (Figure 17). This implies that even though DNA methylation potentially regulates Sox1 transcription, and furthermore safeguards the gene against unphysiologic activation, it seems to be not the only barrier present at the Sox1

promoter. This is further supported by the behavior of Oct4, MyoD1, Il1rn, and Nkx2-2 upon combined targeting of dCas9-VP64 and dCas9-Tet1 to their respective promoters. Even though DNA demethylation of the highly methylated, CpG-rich promoters did increase gene induction facilitated by VP64, in all cases a significant amount of NPCs did not react with activation of the respective target, underlining the probability of other barriers being present. Furthermore, targeting VP64 to the promoter of Ngn2 was not successful, as NPCs did not respond with target induction, even though the CpG rich promoter exhibits only very low levels of DNA methylation. Collectively, it is therefore evident that even though important, DNA methylation is not the only barrier safeguarding cell identity.

Of the analyzed set of chromatin features at the Sox1 promoter, H3K9me3 could potentially act as a transcriptional barrier, as it was present in high levels, comparable to those at the promoter of a repressed control gene. Furthermore, its levels were significantly lower in Sox1^{GFP} positive NPCs. Nevertheless, combining targeted gene induction via dCas9-VP64 with the H3K9-demethylase Jmjd2a fused to dCas9 did not lead to a significant increase in activation efficiency. Whether the lack of effect is due to incomplete histone demethylation, or due to a lack of functionality of the chromatin mark, at least in this context, cannot be inferred from this data. ChIP-qPCR on NPCs following targeting of Jmjd2a to the Sox1 promoter to investigate whether H3K9me3 levels are decreased could shed light on this specific question.

H3K27me3 is an important repressive chromatin mark in the context of poised gene promoters and is correlated to dynamic changes in gene transcription (Ferrari et al. 2014, Vastenhouw and Schier 2012)). It does however not seem to play an important role in the regulation of Sox1, as levels were almost not detectable at the gene locus. The findings on H3K9me3 and H3K27me3 collectively indicate a substantial, rather than dynamic repression of Sox1 (as discussed above).

There are many chromatin features, for which the investigation was beyond the scope of this thesis, but which could be relevant in the regulation of master transcription factors, and thereby the upkeep of cell identities. H3K9me3 for example has been correlated to highly compacted and silenced heterochromatin (Schotta et al. 2004), and

it has been shown to perform lineage-specifying functions during development, where it represses lineage specific genes as long as they are not in need and silence pluripotency or stem cell genes as soon as they are not longer transcribed to allow the cell further differentiation (Nicetto et al. 2019). In line with these results, the levels of this mark at the Sox1 promoter of Sox1^{GFP}-positive and -negative NPCs suggest that H3K9me3 might in fact play a role in silencing and compacting the proneural gene during the transition of NSCs to glial precursors. Therefore it could be interesting to further look into chromatin accessibility at the Sox1 promoter between Sox1^{GFP} positive and negative NPCs. In addition, its silencing function could take over when the Sox1 promoter loses other barriers like DNA-methylation to keep the integrity of the cells intact. This could explain the substantial number of unresponsive cells that are present even after complete demethylation of the Sox1 promoter (Figure 17A+C).

It is furthermore likely that not only the promoter region of the gene locus possesses a regulatory function on the transcription. Distal regulatory elements might regulate a gene in a similar manner as barriers at its promoter. I therefore performed screens with VP64 and different chromatin modifying enzymes tethered to dCas9 in order to identify candidate regulatory regions around the Sox1-gene. While the analysis did not reveal explicit, clearly defined regions with regulatory functions (Figure 21), the verification of the activating potential of six randomly chosen gRNA candidates from the VP64 screen indicates that I was indeed able to isolate functional gRNAs from the library. By refining the method of analysis it will likely be possible to identify regulatory regions from the four different modifiers employed, which in turn will give valuable information on functional chromatin networks and gene regulation.

4.4 The chromatin model of epigenetic gene regulation

To date, many publications have shown correlations between certain chromatin modifications and different transcriptional states. This is also the case for DNA methylation, which is associated with repressed genes when present in CpG-rich regions in the promoter of the respective gene. Low levels of DNA methylation in gene promoters on the other hand have been implied in active gene transcription (Grosjean

2009). In conflict with these correlations are studies in ESCs lacking all DNA methylation and hydroxymethylation by triple Dnmt knock-out. These cells did not exhibit any changes in their transcriptome, indicating that the expected de-repression by lack of DNA methylation did not take place. Upon differentiation however, the knock-out cells died (Domcke et al. 2015, Verma et al. 2018). It appears thusly that only when the transcriptional patterns in a cell need to undergo profound alterations, in this case when ESCs exit pluripotency and commit to embryonic lineages, missing DNA methylation will lead to gene misregulation. This is further supported by my findings upon inhibition of Dnmts with Zebularine, where the resulting reduction of DNA methylation on a global level did not affect NPCs in any obvious way. Upon targeted induction of Sox1, global demethylation did however lead to a significant increase in the amount of responsive NPCs. This indicates that on the one hand, DNA methylation is a potent barrier of transcription, confirming previously described correlations. On the other hand, without an external stimulus (i.e. a stimulus from outside the cellular context, as it also appears during differentiation) the removal of that barrier does not induce any phenotype. This is furthermore shown by the NeuroD4 promoter, where the lack of DNA methylation only comes into effect when an external stimulus activates transcription of the gene. This suggests that even though a causal role for DNA methylation seems obvious, the correlation is not as linear as might be expected, and a complex connection exists between chromatin features and gene transcription.

In line with this complex relationship between DNA methylation and transcription, my findings do not support a direct correlation of the methylation status and trans-activation at the Actc1, Ngn2 and NeuroD4 loci. Despite the lack of the putative repressive chromatin modification while CpGs are found in the respective promoters, Ngn2 and NeuroD4 are not expressed in NPCs, as assessed by transcriptome analysis and ICC, an observation that is in line with earlier studies that showed that many CpG islands in gene promoters stay unmethylated independently of the transcriptional state of the respective gene (Weber et al. 2007, Mohn et al. 2008). Similar to the observations made by Domcke et al., the lack of DNA methylation at the NeuroD4 promoter only comes into effect, when the gene is targeted by a trans-activator, while the Ngn2 locus is unresponsive even to such a manipulation. On the other hand, at the

Actc1 promoter, high DNA methylation levels (over 85% methylation at all analyzed CpGs) did not interfere with trans-activation of the gene via dCas9-VP64. Furthermore, induction of transcription in turn did not change the DNA methylation levels significantly. This promoter therefore seems to be independent of any functional potential that DNA methylation might exert at other loci, probably due to the low density of CpGs that have earlier been linked to a lack of function (Schübeler 2015).

Taken together, the results shown in this thesis generally confirm functional implications of DNA methylation. In line with my findings, DNA methylation has been shown to play a substantial role in cell fate decision and survival in the neural lineage, and misregulation of this chromatin feature is related to psychiatric diseases like schizophrenia (reviewed in (Symmank and Zimmer 2017)). At the same time, I clearly demonstrate that regulatory functions of chromatin features are part of a much more complex mechanism than simple correlations might suggest, and is locus-specific and dependent on external stimuli. This is in line with the mentioned functional studies from the Schübeler Lab (Weber et al. 2007, Mohn et al. 2008, Domcke et al. 2015).

4.5 Implications for the dCas9 tool

Recent publications on the binding behavior of Cas9 sparked discussions in the field on whether or not nucleosome occupancy at the target site could interfere and thus be detrimental to the binding efficiency of the nuclease (Hinz, Laughery, and Wyrick 2015, Kuscu et al. 2014, O'Geen et al. 2015). The relevant studies focused on the wild type protein, while the mutant version (dCas9) has commonly been neglected. Therefore, the kinetics downstream of Cas9 binding has not yet been investigated in depth. Nevertheless, when used as a shuttle for an effector domain, knowledge of subsequent processes becomes substantial in order to estimate the efficiency of the desired effect.

Herein I was able to show that even though NPCs react very heterogeneously to a trans-activating stimulus at the Sox1 promoter, all cells have comparable dCas9 occupancy at the target site, independently of their phenotype. While this rules out defective binding as a reason for differential responsiveness, it also underlines the

possibility of additional factors modulating the potency of targeted effector domains downstream of their binding (e.g. chromatin barriers, see above). It is therefore substantial for the field to not only focus on optimization of binding efficiencies (e.g. by new or re-worked algorithms or analysis of Cas9 kinetics (Doench et al. 2014)) but also to investigate in more detail the mechanisms that interfere with trans-activation, and how these could be circumvented.

My findings not only address potential obstacles of trans-activation, but also point out a profound divergence between the potency of a TAF to induce a target gene to high levels, and its efficiency to do so in a high amount of cells. Comparing the efficacies of VP64 and VPR on Sox1 induction, I was able to show that even though VPR leads to higher levels of mRNA, the number of responsive NPCs did not change. A number of recent methodical papers continually improved TAFs, with VPR being one of the most potent domains this far (La Russa and Qi 2015). However, the magnitude of trans-activation is mainly determined by qPCR and does not take single cell analysis into consideration. Even though this method appears laborious for verification of a trans-activating domain, my results conclusively show confirmation on the single-cell level would indeed be an important aspect to take into account. This is especially relevant for experimental approaches that rely not only on the level of gene induction, but also on the number of responsive cells. Since the results presented in this work indicate that this issue is mostly relevant for master transcription factors, reprogramming approaches in particular will benefit from further advancement of dCas9-tools.

Considering the functional proof for a gene regulatory role of DNA methylation and other chromatin modifications presented here and elsewhere, it would be of high interest to combine potent trans-activating domains like dCas9-VPR with chromatin modifying enzymes. Recent advances in the development of orthogonal methods would allow for elegant targeting of several effector domains with the same dCas9, and by that maximize spatial efficiency of the targeting machinery. For example, by tagging dCas9 with a specific peptide sequence that contains 10 copies of an antibody epitope and linking VP64 to an antibody light-chain that recognizes this epitope, it is possible to target 10 trans-activating domains with only one dCas9 protein (Tanenbaum et al. 2014). In a different study, a SAM loop was introduced into the gRNA stem loop. This

loop protrudes from the gRNA-dCas9 complex and can be bound by an MS2 binding protein, enabling the combination of three different effector domains to produce a potent TAF (Zhang et al. 2015). However, my findings suggest that merely enhancing the potency of targeted trans-activators further might not be the best approach to overcome barriers at gene promoters. The mentioned approaches would however allow for the combination of up to four different effector domains on one dCas9 protein when fused to the N- and C-terminal tail of dCas9 as well as the respective tags. This would open up the potential for highly intricate modifying complexes, which could manipulate the chromatin on several levels into a more admissive state and at the same time induce transcription of a target gene with high efficiency.

4.6 Conclusion

Taken together, my study underlines the importance of Sox1 as cell fate determining factor during neural development that has been suggested previously. Its potency to change the identity of NPCs at least partially to NSCs requires a tight transcriptional regulation. I identified DNA methylation at the promoter of such master transcription factors as potential barrier against trans-activation, a finding that does not only shed new light on the functional role of chromatin features, but also emphasizes the relevance of new approaches in transcriptional engineering (Figure 23). I showed that combination with epigenetic editing is a worthwhile method on this context.

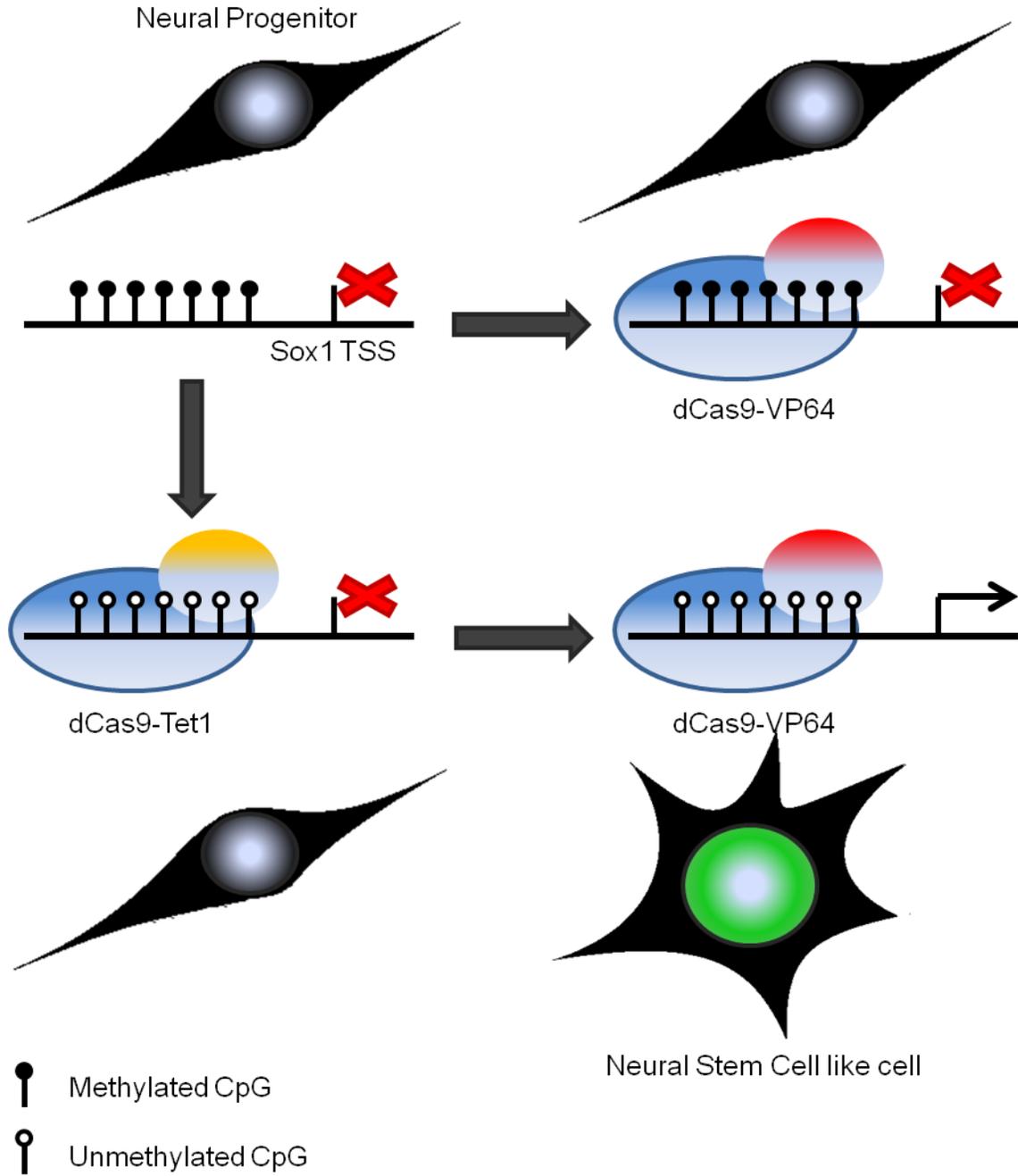


Figure 23: Sox1 transcription is blocked by promoter DNA methylation in NPCs. I found that the induction of Sox1 and other master transcription factors by dCas9-VP64 is prevented by DNA methylation. Targeted Demethylation allows for gene induction with subsequent changes in cell identity (in case of Sox1 to NSCs).

Bibliography

- Biological theory*. Cambridge, MA : MIT Press, c2006-.
- Aaku-Saraste, E., A. Hellwig, and W. B. Huttner. 1996. "Loss of occludin and functional tight junctions, but not ZO-1, during neural tube closure--remodeling of the neuroepithelium prior to neurogenesis." *Dev Biol* 180 (2):664-79. doi: 10.1006/dbio.1996.0336.
- Agrotis, A., and R. Ketteler. 2015. "A new age in functional genomics using CRISPR/Cas9 in arrayed library screening." *Front Genet* 6:300. doi: 10.3389/fgene.2015.00300.
- Amabile, A., A. Migliara, P. Capasso, M. Biffi, D. Cittaro, L. Naldini, and A. Lombardo. 2016. "Inheritable Silencing of Endogenous Genes by Hit-and-Run Targeted Epigenetic Editing." *Cell* 167 (1):219-232.e14. doi: 10.1016/j.cell.2016.09.006.
- Anton, T., and S. Bultmann. 2017. "Site-specific recruitment of epigenetic factors with a modular CRISPR/Cas system." *Nucleus* 8 (3):279-286. doi: 10.1080/19491034.2017.1292194.
- Arend, W. P., M. Malyak, C. J. Guthridge, and C. Gabay. 1998. "Interleukin-1 receptor antagonist: role in biology." *Annu Rev Immunol* 16:27-55. doi: 10.1146/annurev.immunol.16.1.27.
- Arendt, D., J. M. Musser, C. V. H. Baker, A. Bergman, C. Cepko, D. H. Erwin, M. Pavlicev, G. Schlosser, S. Widder, M. D. Laubichler, and G. P. Wagner. 2016. "The origin and evolution of cell types." *Nat Rev Genet* 17 (12):744-757. doi: 10.1038/nrg.2016.127.
- Aubert, J., M. P. Stavridis, S. Tweedie, M. O'Reilly, K. Vierlinger, M. Li, P. Ghazal, T. Pratt, J. O. Mason, D. Roy, and A. Smith. 2003. "Screening for mammalian neural genes via fluorescence-activated cell sorter purification of neural precursors from Sox1-gfp knock-in mice." *Proc Natl Acad Sci U S A* 100 Suppl 1:11836-41. doi: 10.1073/pnas.1734197100.
- Bard, J. 2008. *Biological Theory*. Springer Netherlands.
- Baumann, V., M. Wiesbeck, C. T. Breunig, J. M. Braun, A. Kofler, J. Ninkovic, M. Gotz, and S. H. Stricker. 2019. "Targeted removal of epigenetic barriers during transcriptional reprogramming." *Nat Commun* 10 (1):2119. doi: 10.1038/s41467-019-10146-8.
- Bell, A. C., and G. Felsenfeld. 2000. "Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene." *Nature* 405 (6785):482-5. doi: 10.1038/35013100.
- Bellocchio, E. E., R. J. Reimer, R. T. Freneau, and R. H. Edwards. 2000. "Uptake of glutamate into synaptic vesicles by an inorganic phosphate transporter." *Science* 289 (5481):957-60.
- Berger, B., R. Dersch, E. Ruthardt, C. Rasiah, S. Rauer, and O. Stich. 2016. "Prevalence of anti-SOX1 reactivity in various neurological disorders." *J Neurol Sci* 369:342-346. doi: 10.1016/j.jns.2016.09.002.
- Berninger, B., M. R. Costa, U. Koch, T. Schroeder, B. Sutor, B. Grothe, and M. Götz. 2007. "Functional properties of neurons derived from in vitro reprogrammed postnatal astroglia." *J Neurosci* 27 (32):8654-64. doi: 10.1523/JNEUROSCI.1615-07.2007.

Bibliography

- Bernstein, B. E., T. S. Mikkelsen, X. Xie, M. Kamal, D. J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S. L. Schreiber, and E. S. Lander. 2006. "A bivalent chromatin structure marks key developmental genes in embryonic stem cells." *Cell* 125 (2):315-26. doi: 10.1016/j.cell.2006.02.041.
- Bianconi, E., A. Piovesan, F. Facchin, A. Beraudi, R. Casadei, F. Frabetti, L. Vitale, M. C. Pelleri, S. Tassani, F. Piva, S. Perez-Amodio, P. Strippoli, and S. Canaider. 2013. "An estimation of the number of cells in the human body." *Ann Hum Biol* 40 (6):463-71. doi: 10.3109/03014460.2013.807878.
- Black, J. B., A. F. Adler, H. G. Wang, A. M. D'Ippolito, H. A. Hutchinson, T. E. Reddy, G. S. Pitt, K. W. Leong, and C. A. Gersbach. 2016. "Targeted Epigenetic Remodeling of Endogenous Loci by CRISPR/Cas9-Based Transcriptional Activators Directly Converts Fibroblasts to Neuronal Cells." *Cell Stem Cell* 19 (3):406-14. doi: 10.1016/j.stem.2016.07.001.
- Braun, S. M. G., J. G. Kirkland, E. J. Chory, D. Husmann, J. P. Calarco, and G. R. Crabtree. 2017. "Rapid and reversible epigenome editing by endogenous chromatin regulators." *Nat Commun* 8 (1):560. doi: 10.1038/s41467-017-00644-y.
- Breunig, C. T., T. Durovic, A. M. Neuner, V. Baumann, M. F. Wiesbeck, A. Köferle, M. Götz, J. Ninkovic, and S. H. Stricker. 2018. "One step generation of customizable gRNA vectors for multiplex CRISPR approaches through string assembly gRNA cloning (STAgR)." *PLoS One* 13 (4):e0196015. doi: 10.1371/journal.pone.0196015.
- Buecker, C., and J. Wysocka. 2012. "Enhancers as information integration hubs in development: lessons from genomics." *Trends Genet* 28 (6):276-84. doi: 10.1016/j.tig.2012.02.008.
- Bultmann, S., and S. H. Stricker. 2018. "Entering the post-epigenomic age: back to epigenetics." *Open Biol* 8 (3). doi: 10.1098/rsob.180013.
- Campanero, M. R., M. I. Armstrong, and E. K. Flemington. 2000. "CpG methylation as a mechanism for the regulation of E2F activity." *Proc Natl Acad Sci U S A* 97 (12):6481-6. doi: 10.1073/pnas.100340697.
- Cano-Rodriguez, D., R. A. Gjaltema, L. J. Jilderda, P. Jellema, J. Dokter-Fokkens, M. H. Ruiters, and M. G. Rots. 2016. "Writing of H3K4Me3 overcomes epigenetic silencing in a sustained but context-dependent manner." *Nat Commun* 7:12284. doi: 10.1038/ncomms12284.
- Cantara, W. A., P. F. Crain, J. Rozenski, J. A. McCloskey, K. A. Harris, X. Zhang, F. A. Vendeix, D. Fabris, and P. F. Agris. 2011. "The RNA Modification Database, RNAMDB: 2011 update." *Nucleic Acids Res* 39 (Database issue):D195-201. doi: 10.1093/nar/gkq1028.
- Chavez, A., J. Scheiman, S. Vora, B. W. Pruitt, M. Tuttle, E. P R Iyer, S. Lin, S. Kiani, C. D. Guzman, D. J. Wiegand, D. Ter-Ovanesyan, J. L. Braff, N. Davidsohn, B. E. Housden, N. Perrimon, R. Weiss, J. Aach, J. J. Collins, and G. M. Church. 2015. "Highly efficient Cas9-mediated transcriptional programming." *Nat Methods* 12 (4):326-8. doi: 10.1038/nmeth.3312.
- Chen, D., H. Ma, H. Hong, S. S. Koh, S. M. Huang, B. T. Schurter, D. W. Aswad, and M. R. Stallcup. 1999. "Regulation of transcription by a protein methyltransferase." *Science* 284 (5423):2174-7.

Bibliography

- Cheng, A. W., H. Wang, H. Yang, L. Shi, Y. Katz, T. W. Theunissen, S. Rangarajan, C. S. Shivalila, D. B. Dadon, and R. Jaenisch. 2013. "Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system." *Cell Res* 23 (10):1163-71. doi: 10.1038/cr.2013.122.
- Cho, S. W., S. Kim, Y. Kim, J. Kweon, H. S. Kim, S. Bae, and J. S. Kim. 2014. "Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases." *Genome Res* 24 (1):132-41. doi: 10.1101/gr.162339.113.
- Choudhury, S. R., Y. Cui, K. Lubecka, B. Stefanska, and J. Irudayaraj. 2016. "CRISPR-dCas9 mediated TET1 targeting for selective DNA demethylation at BRCA1 promoter." *Oncotarget* 7 (29):46545-46556. doi: 10.18632/oncotarget.10234.
- Cloos, P. A., J. Christensen, K. Agger, A. Maiolica, J. Rappsilber, T. Antal, K. H. Hansen, and K. Helin. 2006. "The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3." *Nature* 442 (7100):307-11. doi: 10.1038/nature04837.
- Colot, V., and J. L. Rossignol. 1999. "Eukaryotic DNA methylation as an evolutionary device." *Bioessays* 21 (5):402-11. doi: 10.1002/(SICI)1521-1878(199905)21:5<402::AID-BIES7>3.0.CO;2-B.
- Cooper-Kuhn, C. M., M. Vroemen, J. Brown, H. Ye, M. A. Thompson, J. Winkler, and H. G. Kuhn. 2002. "Impaired adult neurogenesis in mice lacking the transcription factor E2F1." *Mol Cell Neurosci* 21 (2):312-23.
- Dame, R. T. 2005. "The role of nucleoid-associated proteins in the organization and compaction of bacterial chromatin." *Mol Microbiol* 56 (4):858-70. doi: 10.1111/j.1365-2958.2005.04598.x.
- Davis, R. L., H. Weintraub, and A. B. Lassar. 1987. "Expression of a single transfected cDNA converts fibroblasts to myoblasts." *Cell* 51 (6):987-1000.
- de Wit, E., and W. de Laat. 2012. "A decade of 3C technologies: insights into nuclear organization." *Genes Dev* 26 (1):11-24. doi: 10.1101/gad.179804.111.
- Doench, J. G., E. Hartenian, D. B. Graham, Z. Tothova, M. Hegde, I. Smith, M. Sullender, B. L. Ebert, R. J. Xavier, and D. E. Root. 2014. "Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation." *Nat Biotechnol* 32 (12):1262-7. doi: 10.1038/nbt.3026.
- Domcke, S., A. F. Bardet, P. Adrian Ginno, D. Hartl, L. Burger, and D. Schübeler. 2015. "Competition between DNA methylation and transcription factors determines binding of NRF1." *Nature* 528 (7583):575-9. doi: 10.1038/nature16462.
- Eckner, R., M. E. Ewen, D. Newsome, M. Gerdes, J. A. DeCaprio, J. B. Lawrence, and D. M. Livingston. 1994. "Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor." *Genes Dev* 8 (8):869-84.
- Feng, S., S. J. Cokus, X. Zhang, P. Y. Chen, M. Bostick, M. G. Goll, J. Hetzel, J. Jain, S. H. Strauss, M. E. Halpern, C. Ukomadu, K. C. Sadler, S. Pradhan, M. Pellegrini, and S. E. Jacobsen. 2010. "Conservation and divergence of methylation patterning in plants and animals." *Proc Natl Acad Sci U S A* 107 (19):8689-94. doi: 10.1073/pnas.1002720107.
- Ferrari, K. J., A. Scelfo, S. Jammula, A. Cuomo, I. Barozzi, A. Stützer, W. Fischle, T. Bonaldi, and D. Pasini. 2014. "Polycomb-dependent H3K27me1 and H3K27me2 regulate active transcription and enhancer fidelity." *Mol Cell* 53 (1):49-62. doi: 10.1016/j.molcel.2013.10.030.

Bibliography

- Fodor, B. D., S. Kubicek, M. Yonezawa, R. J. O'Sullivan, R. Sengupta, L. Perez-Burgos, S. Opravil, K. Mechtler, G. Schotta, and T. Jenuwein. 2006. "Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells." *Genes Dev* 20 (12):1557-62. doi: 10.1101/gad.388206.
- Freudenberg, J. M., S. Ghosh, B. L. Lackford, S. Yellaboina, X. Zheng, R. Li, S. Cuddapah, P. A. Wade, G. Hu, and R. Jothi. 2012. "Acute depletion of Tet1-dependent 5-hydroxymethylcytosine levels impairs LIF/Stat3 signaling and results in loss of embryonic stem cell identity." *Nucleic Acids Res* 40 (8):3364-77. doi: 10.1093/nar/gkr1253.
- Fu, Y., J. A. Foden, C. Khayter, M. L. Maeder, D. Reyon, J. K. Joung, and J. D. Sander. 2013. "High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells." *Nat Biotechnol* 31 (9):822-6. doi: 10.1038/nbt.2623.
- Gallego-Bartolomé, J., J. Gardiner, W. Liu, A. Papikian, B. Ghoshal, H. Y. Kuo, J. M. Zhao, D. J. Segal, and S. E. Jacobsen. 2018. "Targeted DNA demethylation of the." *Proc Natl Acad Sci U S A* 115 (9):E2125-E2134. doi: 10.1073/pnas.1716945115.
- Grosjean, H. 2009. "Nucleic Acids Are Not Boring Long Polymers of Only Four Types of Nucleotides: A Guided Tour."
- Guenzl, P. M., and D. P. Barlow. 2012. "Macro lncRNAs: a new layer of cis-regulatory information in the mammalian genome." *RNA Biol* 9 (6):731-41. doi: 10.4161/rna.19985.
- Guo, C., and S. A. Morris. 2017. "Engineering cell identity: establishing new gene regulatory and chromatin landscapes." *Curr Opin Genet Dev* 46:50-57. doi: 10.1016/j.gde.2017.06.011.
- Gurdon, J. B., T. R. Elsdale, and M. Fischberg. 1958. "Sexually mature individuals of *Xenopus laevis* from the transplantation of single somatic nuclei." *Nature* 182 (4627):64-5.
- Hark, A. T., C. J. Schoenherr, D. J. Katz, R. S. Ingram, J. M. Levorse, and S. M. Tilghman. 2000. "CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus." *Nature* 405 (6785):486-9. doi: 10.1038/35013106.
- Harris, Henry. 1999. *The birth of the cell*. New Haven, Conn. ; London: Yale University Press.
- Hartenian, E., and J. G. Doench. 2015. "Genetic screens and functional genomics using CRISPR/Cas9 technology." *FEBS J* 282 (8):1383-93. doi: 10.1111/febs.13248.
- He, Y., and P. Casaccia-Bonnel. 2008. "The Yin and Yang of YY1 in the nervous system." *J Neurochem* 106 (4):1493-502. doi: 10.1111/j.1471-4159.2008.05486.x.
- He, Y. F., B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He, and G. L. Xu. 2011. "Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA." *Science* 333 (6047):1303-7. doi: 10.1126/science.1210944.
- Helm, M., and Y. Motorin. 2017. "Detecting RNA modifications in the epitranscriptome: predict and validate." *Nat Rev Genet* 18 (5):275-291. doi: 10.1038/nrg.2016.169.
- Hendry, S. H., E. G. Jones, P. C. Emson, D. E. Lawson, C. W. Heizmann, and P. Streit. 1989. "Two classes of cortical GABA neurons defined by differential calcium binding protein immunoreactivities." *Exp Brain Res* 76 (2):467-72.

Bibliography

- Higuchi, M., H. Kiyama, T. Hayakawa, Y. Hamada, and Y. Tsujimoto. 1995. "Differential expression of Notch1 and Notch2 in developing and adult mouse brain." *Brain Res Mol Brain Res* 29 (2):263-72.
- Hilton, I. B., A. M. D'Ippolito, C. M. Vockley, P. I. Thakore, G. E. Crawford, T. E. Reddy, and C. A. Gersbach. 2015. "Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers." *Nat Biotechnol* 33 (5):510-7. doi: 10.1038/nbt.3199.
- Hinz, J. M., M. F. Laughery, and J. J. Wyrick. 2015. "Nucleosomes Inhibit Cas9 Endonuclease Activity in Vitro." *Biochemistry* 54 (48):7063-6. doi: 10.1021/acs.biochem.5b01108.
- Hirase, T., J. M. Staddon, M. Saitou, Y. Ando-Akatsuka, M. Itoh, M. Furuse, K. Fujimoto, S. Tsukita, and L. L. Rubin. 1997. "Occludin as a possible determinant of tight junction permeability in endothelial cells." *J Cell Sci* 110 (Pt 14):1603-13.
- Hobert, O. 2008. "Regulatory logic of neuronal diversity: terminal selector genes and selector motifs." *Proc Natl Acad Sci U S A* 105 (51):20067-71. doi: 10.1073/pnas.0806070105.
- Hotchkiss, R. D. 1948. "The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography." *J Biol Chem* 175 (1):315-32.
- Hu, J., Y. Lei, W. K. Wong, S. Liu, K. C. Lee, X. He, W. You, R. Zhou, J. T. Guo, X. Chen, X. Peng, H. Sun, H. Huang, H. Zhao, and B. Feng. 2014. "Direct activation of human and mouse Oct4 genes using engineered TALE and Cas9 transcription factors." *Nucleic Acids Res* 42 (7):4375-90. doi: 10.1093/nar/gku109.
- Huang, Y., J. Fang, M. T. Bedford, Y. Zhang, and R. M. Xu. 2006. "Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A." *Science* 312 (5774):748-51. doi: 10.1126/science.1125162.
- Illingworth, R. S., and A. P. Bird. 2009. "CpG islands--'a rough guide'." *FEBS Lett* 583 (11):1713-20. doi: 10.1016/j.febslet.2009.04.012.
- Ito, S., L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He, and Y. Zhang. 2011. "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." *Science* 333 (6047):1300-3. doi: 10.1126/science.1210597.
- Iwafuchi-Doi, M., and K. S. Zaret. 2014. "Pioneer transcription factors in cell reprogramming." *Genes Dev* 28 (24):2679-92. doi: 10.1101/gad.253443.114.
- Jenuwein, T., G. Laible, R. Dorn, and G. Reuter. 1998. "SET domain proteins modulate chromatin domains in eu- and heterochromatin." *Cell Mol Life Sci* 54 (1):80-93.
- Jiang, X., and J. Nardelli. 2016. "Cellular and molecular introduction to brain development." *Neurobiol Dis* 92 (Pt A):3-17. doi: 10.1016/j.nbd.2015.07.007.
- Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and E. Charpentier. 2012. "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." *Science* 337 (6096):816-21. doi: 10.1126/science.1225829.
- Julian, L. M., A. C. McDonald, and W. L. Stanford. 2017. "Direct reprogramming with SOX factors: masters of cell fate." *Curr Opin Genet Dev* 46:24-36. doi: 10.1016/j.gde.2017.06.005.
- Kabadi, A. M., D. G. Ousterout, I. B. Hilton, and C. A. Gersbach. 2014. "Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector." *Nucleic Acids Res* 42 (19):e147. doi: 10.1093/nar/gku749.
- Kamachi, Y., and H. Kondoh. 2013. "Sox proteins: regulators of cell fate specification and differentiation." *Development* 140 (20):4129-44. doi: 10.1242/dev.091793.

Bibliography

- Kan, L., N. Israsena, Z. Zhang, M. Hu, L. R. Zhao, A. Jalali, V. Sahni, and J. A. Kessler. 2004. "Sox1 acts through multiple independent pathways to promote neurogenesis." *Dev Biol* 269 (2):580-94. doi: 10.1016/j.ydbio.2004.02.005.
- Kan, L., A. Jalali, L. R. Zhao, X. Zhou, T. McGuire, I. Kazanis, V. Episkopou, A. G. Bassuk, and J. A. Kessler. 2007. "Dual function of Sox1 in telencephalic progenitor cells." *Dev Biol* 310 (1):85-98. doi: 10.1016/j.ydbio.2007.07.026.
- Kearns, N. A., H. Pham, B. Tabak, R. M. Genga, N. J. Silverstein, M. Garber, and R. Maehr. 2015. "Functional annotation of native enhancers with a Cas9-histone demethylase fusion." *Nat Methods* 12 (5):401-403. doi: 10.1038/nmeth.3325.
- Kelsh, R. N. 2006. "Sorting out Sox10 functions in neural crest development." *Bioessays* 28 (8):788-98. doi: 10.1002/bies.20445.
- Khare, S. P., F. Habib, R. Sharma, N. Gadewal, S. Gupta, and S. Galande. 2012. "Histone--a relational knowledgebase of human histone proteins and histone modifying enzymes." *Nucleic Acids Res* 40 (Database issue):D337-42. doi: 10.1093/nar/gkr1125.
- Kho, M. R., D. J. Baker, A. Laayoun, and S. S. Smith. 1998. "Stalling of human DNA (cytosine-5) methyltransferase at single-strand conformers from a site of dynamic mutation." *J Mol Biol* 275 (1):67-79. doi: 10.1006/jmbi.1997.1430.
- Kim, J. B., H. Zaehres, G. Wu, L. Gentile, K. Ko, V. Sebastiano, M. J. Araúzo-Bravo, D. Ruau, D. W. Han, M. Zenke, and H. R. Schöler. 2008. "Pluripotent stem cells induced from adult neural stem cells by reprogramming with two factors." *Nature* 454 (7204):646-50. doi: 10.1038/nature07061.
- Kim, J. S., H. J. Lee, and D. Carroll. 2010. "Genome editing with modularly assembled zinc-finger nucleases." *Nat Methods* 7 (2):91; author reply 91-2. doi: 10.1038/nmeth0210-91a.
- Kuscu, C., S. Arslan, R. Singh, J. Thorpe, and M. Adli. 2014. "Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease." *Nat Biotechnol* 32 (7):677-83. doi: 10.1038/nbt.2916.
- Köferle, A., K. Worf, C. Breunig, V. Baumann, J. Herrero, M. Wiesbeck, L. H. Hutter, M. Götz, C. Fuchs, S. Beck, and S. H. Stricker. 2016. "CORALINA: a universal method for the generation of gRNA libraries for CRISPR-based screening." *BMC Genomics* 17 (1):917. doi: 10.1186/s12864-016-3268-z.
- La Russa, M. F., and L. S. Qi. 2015. "The New State of the Art: Cas9 for Gene Activation and Repression." *Mol Cell Biol* 35 (22):3800-9. doi: 10.1128/MCB.00512-15.
- Laible, G., A. Wolf, R. Dorn, G. Reuter, C. Nislow, A. Lebersorger, D. Popkin, L. Pillus, and T. Jenuwein. 1997. "Mammalian homologues of the Polycomb-group gene Enhancer of zeste mediate gene silencing in Drosophila heterochromatin and at S. cerevisiae telomeres." *EMBO J* 16 (11):3219-32. doi: 10.1093/emboj/16.11.3219.
- Lane, A. B., M. Strzelecka, A. Ettinger, A. W. Grenfell, T. Wittmann, and R. Heald. 2015. "Enzymatically Generated CRISPR Libraries for Genome Labeling and Screening." *Dev Cell* 34 (3):373-8. doi: 10.1016/j.devcel.2015.06.003.
- Lee, H. J., T. A. Hore, and W. Reik. 2014. "Reprogramming the methylome: erasing memory and creating diversity." *Cell Stem Cell* 14 (6):710-9. doi: 10.1016/j.stem.2014.05.008.

Bibliography

- Lee, J. T. 2012. "Epigenetic regulation by long noncoding RNAs." *Science* 338 (6113):1435-9. doi: 10.1126/science.1231776.
- Lendahl, U., L. B. Zimmerman, and R. D. McKay. 1990. "CNS stem cells express a new class of intermediate filament protein." *Cell* 60 (4):585-95.
- Li, S., and C. E. Mason. 2014. "The pivotal regulatory landscape of RNA modifications." *Annu Rev Genomics Hum Genet* 15:127-50. doi: 10.1146/annurev-genom-090413-025405.
- Li, Z., D. Zhang, X. Xiong, B. Yan, W. Xie, J. Sheen, and J. F. Li. 2017. "A potent Cas9-derived gene activator for plant and mammalian cells." *Nat Plants* 3 (12):930-936. doi: 10.1038/s41477-017-0046-0.
- Liu, P., M. Chen, Y. Liu, L. S. Qi, and S. Ding. 2018. "CRISPR-Based Chromatin Remodeling of the Endogenous Oct4 or Sox2 Locus Enables Reprogramming to Pluripotency." *Cell Stem Cell* 22 (2):252-261.e4. doi: 10.1016/j.stem.2017.12.001.
- Liu, X. S., H. Wu, X. Ji, Y. Stelzer, X. Wu, S. Czauderna, J. Shu, D. Dadon, R. A. Young, and R. Jaenisch. 2016. "Editing DNA Methylation in the Mammalian Genome." *Cell* 167 (1):233-247.e17. doi: 10.1016/j.cell.2016.08.056.
- Loh, Y. H., W. Zhang, X. Chen, J. George, and H. H. Ng. 2007. "Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells." *Genes Dev* 21 (20):2545-57. doi: 10.1101/gad.1588207.
- Love, M. I., W. Huber, and S. Anders. 2014. "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2." *Genome Biol* 15 (12):550. doi: 10.1186/s13059-014-0550-8.
- Lowary, P. T., and J. Widom. 1997. "Nucleosome packaging and nucleosome positioning of genomic DNA." *Proc Natl Acad Sci U S A* 94 (4):1183-8.
- Luo, G. Z., M. A. Blanco, E. L. Greer, C. He, and Y. Shi. 2015. "DNA N(6)-methyladenine: a new epigenetic mark in eukaryotes?" *Nat Rev Mol Cell Biol* 16 (12):705-10. doi: 10.1038/nrm4076.
- Maeder, M. L., S. J. Linder, V. M. Cascio, Y. Fu, Q. H. Ho, and J. K. Joung. 2013. "CRISPR RNA-guided activation of endogenous human genes." *Nat Methods* 10 (10):977-9. doi: 10.1038/nmeth.2598.
- Maresch, R., S. Mueller, C. Veltkamp, R. Öllinger, M. Friedrich, I. Heid, K. Steiger, J. Weber, T. Engleitner, M. Barenboim, S. Klein, S. Louzada, R. Banerjee, A. Strong, T. Stauber, N. Gross, U. Geumann, S. Lange, M. Ringelhan, I. Varela, K. Unger, F. Yang, R. M. Schmid, G. S. Vassiliou, R. Braren, G. Schneider, M. Heikenwalder, A. Bradley, D. Saur, and R. Rad. 2016. "Multiplexed pancreatic genome engineering and cancer induction by transfection-based CRISPR/Cas9 delivery in mice." *Nat Commun* 7:10770. doi: 10.1038/ncomms10770.
- Marín, O., and U. Müller. 2014. "Lineage origins of GABAergic versus glutamatergic neurons in the neocortex." *Curr Opin Neurobiol* 26:132-41. doi: 10.1016/j.conb.2014.01.015.
- Masserdotti, G., S. Gillotin, B. Sutor, D. Drechsel, M. Irmeler, H. F. Jørgensen, S. Sass, F. J. Theis, J. Beckers, B. Berninger, F. Guillemot, and M. Götz. 2015. "Transcriptional Mechanisms of Proneural Factors and REST in Regulating Neuronal Reprogramming of Astrocytes." *Cell Stem Cell* 17 (1):74-88. doi: 10.1016/j.stem.2015.05.014.

Bibliography

- Mauer, J., X. Luo, A. Blanjoie, X. Jiao, A. V. Grozhik, D. P. Patil, B. Linder, B. F. Pickering, J. J. Vasseur, Q. Chen, S. S. Gross, O. Elemento, F. Debart, M. Kiledjian, and S. R. Jaffrey. 2017. "Reversible methylation of m." *Nature* 541 (7637):371-375. doi: 10.1038/nature21022.
- Merkenschlager, M., and E. P. Nora. 2016. "CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation." *Annu Rev Genomics Hum Genet* 17:17-43. doi: 10.1146/annurev-genom-083115-022339.
- Mitalipov, S., and D. Wolf. 2009. "Totipotency, pluripotency and nuclear reprogramming." *Adv Biochem Eng Biotechnol* 114:185-99. doi: 10.1007/10_2008_45.
- Mohn, F., M. Weber, M. Rebhan, T. C. Roloff, J. Richter, M. B. Stadler, M. Bibel, and D. Schübeler. 2008. "Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors." *Mol Cell* 30 (6):755-66. doi: 10.1016/j.molcel.2008.05.007.
- Morita, S., H. Noguchi, T. Horii, K. Nakabayashi, M. Kimura, K. Okamura, A. Sakai, H. Nakashima, K. Hata, K. Nakashima, and I. Hatada. 2016. "Targeted DNA demethylation in vivo using dCas9-peptide repeat and scFv-TET1 catalytic domain fusions." *Nat Biotechnol* 34 (10):1060-1065. doi: 10.1038/nbt.3658.
- Nakagawa, M., M. Koyanagi, K. Tanabe, K. Takahashi, T. Ichisaka, T. Aoi, K. Okita, Y. Mochiduki, N. Takizawa, and S. Yamanaka. 2008. "Generation of induced pluripotent stem cells without Myc from mouse and human fibroblasts." *Nat Biotechnol* 26 (1):101-6. doi: 10.1038/nbt1374.
- Nakajima-Koyama, M., J. Lee, S. Ohta, T. Yamamoto, and E. Nishida. 2015. "Induction of Pluripotency in Astrocytes through a Neural Stem Cell-like State." *J Biol Chem* 290 (52):31173-88. doi: 10.1074/jbc.M115.683466.
- Nicetto, D., G. Donahue, T. Jain, T. Peng, S. Sidoli, L. Sheng, T. Montavon, J. S. Becker, J. M. Grindheim, K. Blahnik, B. A. Garcia, K. Tan, R. Bonasio, T. Jenuwein, and K. S. Zaret. 2019. "H3K9me3-heterochromatin loss at protein-coding genes enables developmental lineage specification." *Science* 363 (6424):294-297. doi: 10.1126/science.aau0583.
- O'Geen, H., I. M. Henry, M. S. Bhakta, J. F. Meckler, and D. J. Segal. 2015. "A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture." *Nucleic Acids Res* 43 (6):3389-404. doi: 10.1093/nar/gkv137.
- O'Geen, H., C. Ren, C. M. Nicolet, A. A. Perez, J. Halmai, V. M. Le, J. P. Mackay, P. J. Farnham, and D. J. Segal. 2017. "dCas9-based epigenome editing suggests acquisition of histone methylation is not sufficient for target gene repression." *Nucleic Acids Res* 45 (17):9901-9916. doi: 10.1093/nar/gkx578.
- Perez-Pinera, P., D. D. Kocak, C. M. Vockley, A. F. Adler, A. M. Kabadi, L. R. Polstein, P. I. Thakore, K. A. Glass, D. G. Ousterout, K. W. Leong, F. Guilak, G. E. Crawford, T. E. Reddy, and C. A. Gersbach. 2013. "RNA-guided gene activation by CRISPR-Cas9-based transcription factors." *Nat Methods* 10 (10):973-6. doi: 10.1038/nmeth.2600.
- Perrin, A., J. Rousseau, and J. P. Tremblay. 2017. "Increased Expression of Laminin Subunit Alpha 1 Chain by dCas9-VP160." *Mol Ther Nucleic Acids* 6:68-79. doi: 10.1016/j.omtn.2016.11.004.
- Peterson, B. A., D. C. Haak, M. T. Nishimura, P. J. Teixeira, S. R. James, J. L. Dangl, and Z. L. Nimchuk. 2016. "Genome-Wide Assessment of Efficiency and

Bibliography

- Specificity in CRISPR/Cas9 Mediated Multiple Site Targeting in Arabidopsis." *PLoS One* 11 (9):e0162169. doi: 10.1371/journal.pone.0162169.
- Pevny, L. H., S. Sockanathan, M. Placzek, and R. Lovell-Badge. 1998. "A role for SOX1 in neural determination." *Development* 125 (10):1967-78.
- Pflueger, C., D. Tan, T. Swain, T. Nguyen, J. Pflueger, C. Nefzger, J. M. Polo, E. Ford, and R. Lister. 2018. "A modular dCas9-SunTag DNMT3A epigenome editing system overcomes pervasive off-target activity of direct fusion dCas9-DNMT3A constructs." *Genome Res* 28 (8):1193-1206. doi: 10.1101/gr.233049.117.
- Pollard, S. M., L. Conti, Y. Sun, D. Goffredo, and A. Smith. 2006. "Adherent neural stem (NS) cells from fetal and adult forebrain." *Cereb Cortex* 16 Suppl 1:i112-20. doi: 10.1093/cercor/bhj167.
- Polstein, L. R., P. Perez-Pinera, D. D. Kocak, C. M. Vockley, P. Bledsoe, L. Song, A. Safi, G. E. Crawford, T. E. Reddy, and C. A. Gersbach. 2015. "Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators." *Genome Res* 25 (8):1158-69. doi: 10.1101/gr.179044.114.
- Pombo, A., and N. Dillon. 2015. "Three-dimensional genome architecture: players and mechanisms." *Nat Rev Mol Cell Biol* 16 (4):245-57. doi: 10.1038/nrm3965.
- Qi, L. S., M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman, A. P. Arkin, and W. A. Lim. 2013. "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." *Cell* 152 (5):1173-83. doi: 10.1016/j.cell.2013.02.022.
- Regev, A., S. A. Teichmann, E. S. Lander, I. Amit, C. Benoist, E. Birney, B. Bodenmiller, P. Campbell, P. Carninci, M. Clatworthy, H. Clevers, B. Deplancke, I. Dunham, J. Eberwine, R. Eils, W. Enard, A. Farmer, L. Fugger, B. Göttgens, N. Hacohen, M. Haniffa, M. Hemberg, S. Kim, P. Klenerman, A. Kriegstein, E. Lein, S. Linnarsson, E. Lundberg, J. Lundberg, P. Majumder, J. C. Marioni, M. Merad, M. Mhlanga, M. Nawijn, M. Netea, G. Nolan, D. Pe'er, A. Phillipakis, C. P. Ponting, S. Quake, W. Reik, O. Rozenblatt-Rosen, J. Sanes, R. Satija, T. N. Schumacher, A. Shalek, E. Shapiro, P. Sharma, J. W. Shin, O. Stegle, M. Stratton, M. J. T. Stubbington, F. J. Theis, M. Uhlen, A. van Oudenaarden, A. Wagner, F. Watt, J. Weissman, B. Wold, R. Xavier, N. Yosef, and Human Cell Atlas Meeting Participants. 2017. "The Human Cell Atlas." *Elife* 6. doi: 10.7554/eLife.27041.
- Reyon, D., M. L. Maeder, C. Khayter, S. Q. Tsai, J. E. Foley, J. D. Sander, and J. K. Joung. 2013. "Engineering customized TALE nucleases (TALENs) and TALE transcription factors by fast ligation-based automatable solid-phase high-throughput (FLASH) assembly." *Curr Protoc Mol Biol* Chapter 12:Unit 12.16. doi: 10.1002/0471142727.mb1216s103.
- Roundtree, I. A., M. E. Evans, T. Pan, and C. He. 2017. "Dynamic RNA Modifications in Gene Expression Regulation." *Cell* 169 (7):1187-1200. doi: 10.1016/j.cell.2017.05.045.
- Sanosaka, T., M. Namihira, and K. Nakashima. 2009. "Epigenetic mechanisms in sequential differentiation of neural stem cells." *Epigenetics* 4 (2):89-92.
- Schneider, R., A. J. Bannister, and T. Kouzarides. 2002. "Unsafe SETs: histone lysine methyltransferases and cancer." *Trends Biochem Sci* 27 (8):396-402.

Bibliography

- Schotta, G., M. Lachner, K. Sarma, A. Ebert, R. Sengupta, G. Reuter, D. Reinberg, and T. Jenuwein. 2004. "A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin." *Genes Dev* 18 (11):1251-62. doi: 10.1101/gad.300704.
- Schübeler, D. 2015. "Function and information content of DNA methylation." *Nature* 517 (7534):321-6. doi: 10.1038/nature14192.
- Shang, W., F. Wang, G. Fan, and H. Wang. 2017. "Key elements for designing and performing a CRISPR/Cas9-based genetic screen." *J Genet Genomics* 44 (9):439-449. doi: 10.1016/j.jgg.2017.09.005.
- Shi, Y., F. Lan, C. Matson, P. Mulligan, J. R. Whetstone, P. A. Cole, and R. A. Casero. 2004. "Histone demethylation mediated by the nuclear amine oxidase homolog LSD1." *Cell* 119 (7):941-53. doi: 10.1016/j.cell.2004.12.012.
- Smith, S. S., B. E. Kaplan, L. C. Sowers, and E. M. Newman. 1992. "Mechanism of human methyl-directed DNA methyltransferase and the fidelity of cytosine methylation." *Proc Natl Acad Sci U S A* 89 (10):4744-8.
- Smith, Z. D., C. Sindhu, and A. Meissner. 2016. "Molecular features of cellular reprogramming and development." *Nat Rev Mol Cell Biol* 17 (3):139-54. doi: 10.1038/nrm.2016.6.
- Spitz, F., and E. E. Furlong. 2012. "Transcription factors: from enhancer binding to developmental control." *Nat Rev Genet* 13 (9):613-26. doi: 10.1038/nrg3207.
- Sternberg, S. H., S. Redding, M. Jinek, E. C. Greene, and J. A. Doudna. 2014. "DNA interrogation by the CRISPR RNA-guided endonuclease Cas9." *Nature* 507 (7490):62-7. doi: 10.1038/nature13011.
- Stiles, J., and T. L. Jernigan. 2010. "The basics of brain development." *Neuropsychol Rev* 20 (4):327-48. doi: 10.1007/s11065-010-9148-4.
- Strahl, B. D., P. A. Grant, S. D. Briggs, Z. W. Sun, J. R. Bone, J. A. Caldwell, S. Mollah, R. G. Cook, J. Shabanowitz, D. F. Hunt, and C. D. Allis. 2002. "Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression." *Mol Cell Biol* 22 (5):1298-306.
- Stricker, S. H., and M. Götz. 2018. "DNA-Methylation: Master or Slave of Neural Fate Decisions?" *Front Neurosci* 12:5. doi: 10.3389/fnins.2018.00005.
- Stricker, S. H., A. Köferle, and S. Beck. 2017. "From profiles to function in epigenomics." *Nat Rev Genet* 18 (1):51-66. doi: 10.1038/nrg.2016.138.
- Sussel, L., J. Kalamaras, D. J. Hartigan-O'Connor, J. J. Meneses, R. A. Pedersen, J. L. Rubenstein, and M. S. German. 1998. "Mice lacking the homeodomain transcription factor Nkx2.2 have diabetes due to arrested differentiation of pancreatic beta cells." *Development* 125 (12):2213-21.
- Symmank, J., and G. Zimmer. 2017. "Regulation of neuronal survival by DNA methyltransferases." *Neural Regen Res* 12 (11):1768-1775. doi: 10.4103/1673-5374.219027.
- Tachibana, M., K. Sugimoto, T. Fukushima, and Y. Shinkai. 2001. "Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3." *J Biol Chem* 276 (27):25309-17. doi: 10.1074/jbc.M101914200.
- Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind, and A. Rao. 2009. "Conversion of 5-

Bibliography

- methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." *Science* 324 (5929):930-5. doi: 10.1126/science.1170116.
- Takahashi, K., and S. Yamanaka. 2006. "Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors." *Cell* 126 (4):663-76. doi: 10.1016/j.cell.2006.07.024.
- Takizawa, T., K. Nakashima, M. Namihira, W. Ochiai, A. Uemura, M. Yanagisawa, N. Fujita, M. Nakao, and T. Taga. 2001. "DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain." *Dev Cell* 1 (6):749-58.
- Tan, M., H. Luo, S. Lee, F. Jin, J. S. Yang, E. Montellier, T. Buchou, Z. Cheng, S. Rousseaux, N. Rajagopal, Z. Lu, Z. Ye, Q. Zhu, J. Wysocka, Y. Ye, S. Khochbin, B. Ren, and Y. Zhao. 2011. "Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification." *Cell* 146 (6):1016-28. doi: 10.1016/j.cell.2011.08.008.
- Tanenbaum, M. E., L. A. Gilbert, L. S. Qi, J. S. Weissman, and R. D. Vale. 2014. "A protein-tagging system for signal amplification in gene expression and fluorescence imaging." *Cell* 159 (3):635-46. doi: 10.1016/j.cell.2014.09.039.
- Tang, Y., P. Yu, and L. Cheng. 2017. "Current progress in the derivation and therapeutic application of neural stem cells." *Cell Death Dis* 8 (10):e3108. doi: 10.1038/cddis.2017.504.
- Teif, V. B., D. A. Beshnova, Y. Vainshtein, C. Marth, J. P. Mallm, T. Höfer, and K. Rippe. 2014. "Nucleosome repositioning links DNA (de)methylation and differential CTCF binding during stem cell development." *Genome Res* 24 (8):1285-95. doi: 10.1101/gr.164418.113.
- Tessarz, P., and T. Kouzarides. 2014. "Histone core modifications regulating nucleosome structure and dynamics." *Nat Rev Mol Cell Biol* 15 (11):703-8. doi: 10.1038/nrm3890.
- Tsukada, Y., J. Fang, H. Erdjument-Bromage, M. E. Warren, C. H. Borchers, P. Tempst, and Y. Zhang. 2006. "Histone demethylation by a family of JmjC domain-containing proteins." *Nature* 439 (7078):811-6. doi: 10.1038/nature04433.
- Uchida, N., D. W. Buck, D. He, M. J. Reitsma, M. Masek, T. V. Phan, A. S. Tsukamoto, F. H. Gage, and I. L. Weissman. 2000. "Direct isolation of human central nervous system stem cells." *Proc Natl Acad Sci U S A* 97 (26):14720-5. doi: 10.1073/pnas.97.26.14720.
- Vaquerizas, J. M., S. K. Kummerfeld, S. A. Teichmann, and N. M. Luscombe. 2009. "A census of human transcription factors: function, expression and evolution." *Nat Rev Genet* 10 (4):252-63. doi: 10.1038/nrg2538.
- Vastenhouw, N. L., and A. F. Schier. 2012. "Bivalent histone modifications in early embryogenesis." *Curr Opin Cell Biol* 24 (3):374-86. doi: 10.1016/j.ceb.2012.03.009.
- Vazquez-Vilar, M., J. M. Bernabé-Orts, A. Fernandez-Del-Carmen, P. Ziarsolo, J. Blanca, A. Granell, and D. Orzaez. 2016. "A modular toolbox for gRNA-Cas9 genome engineering in plants based on the GoldenBraid standard." *Plant Methods* 12:10. doi: 10.1186/s13007-016-0101-2.
- Venere, M., Y. G. Han, R. Bell, J. S. Song, A. Alvarez-Buylla, and R. Blelloch. 2012. "Sox1 marks an activated neural stem/progenitor cell in the hippocampus." *Development* 139 (21):3938-49. doi: 10.1242/dev.081133.

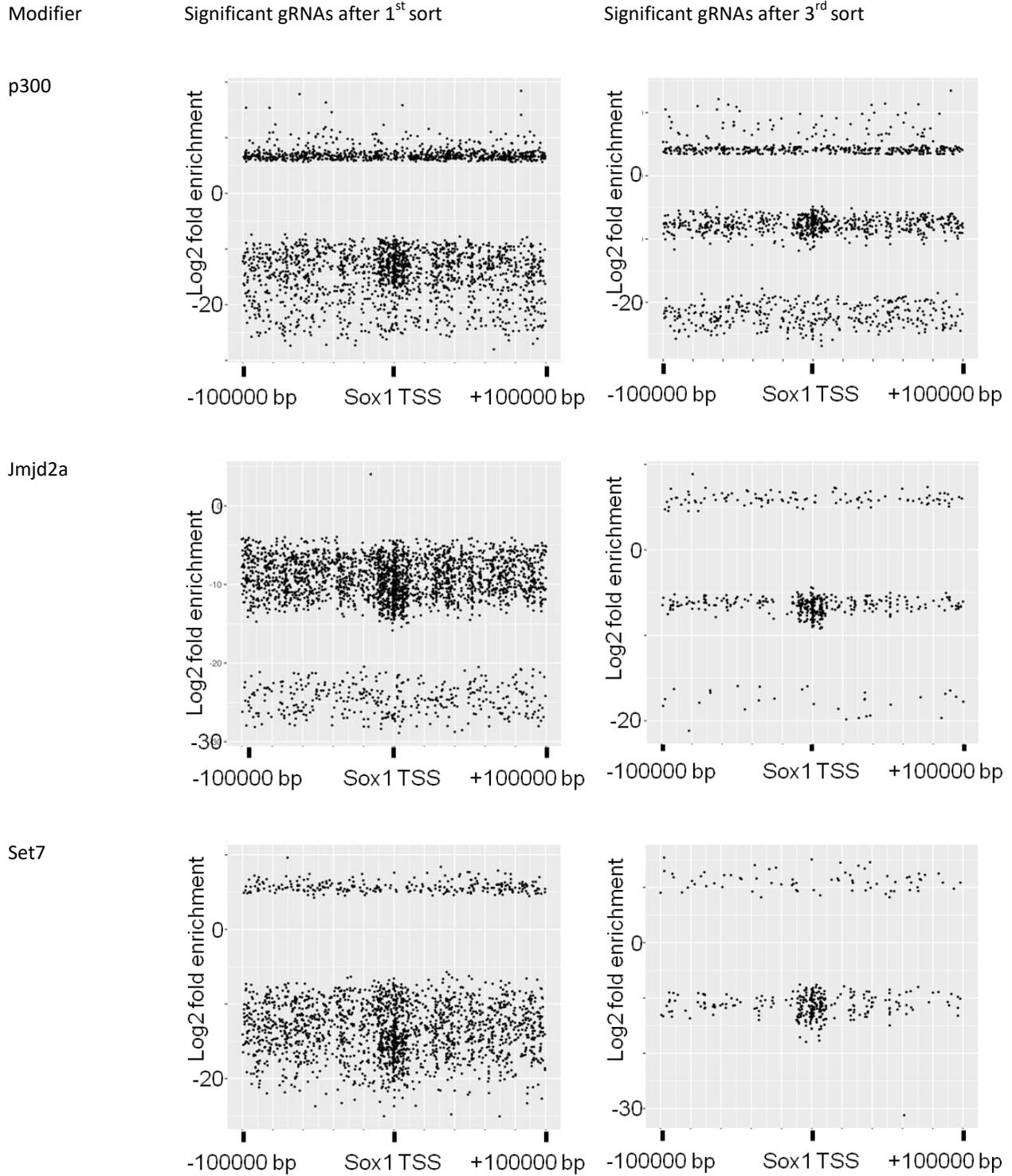
Bibliography

- Verma, N., H. Pan, L. C. Dore, A. Shukla, Q. V. Li, B. Pelham-Webb, V. Teijeiro, F. Gonzalez, A. Krivtsov, C. J. Chang, E. P. Papapetrou, C. He, O. Elemento, and D. Huangfu. 2018. "TET proteins safeguard bivalent promoters from de novo methylation in human embryonic stem cells." *Nat Genet* 50 (1):83-95. doi: 10.1038/s41588-017-0002-y.
- Vickaryous, M. K., and B. K. Hall. 2006. "Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest." *Biol Rev Camb Philos Soc* 81 (3):425-55. doi: 10.1017/S1464793106007068.
- Vierbuchen, T., A. Ostermeier, Z. P. Pang, Y. Kokubu, T. C. Südhof, and M. Wernig. 2010. "Direct conversion of fibroblasts to functional neurons by defined factors." *Nature* 463 (7284):1035-41. doi: 10.1038/nature08797.
- Vojta, A., P. Dobrinić, V. Tadić, L. Bočkor, P. Korać, B. Julg, M. Klasić, and V. Zoldoš. 2016. "Repurposing the CRISPR-Cas9 system for targeted DNA methylation." *Nucleic Acids Res* 44 (12):5615-28. doi: 10.1093/nar/gkw159.
- Waddington, C. H. 2012. "The epigenotype. 1942." *Int J Epidemiol* 41 (1):10-3. doi: 10.1093/ije/dyr184.
- Walther, D. J., J. U. Peter, S. Bashammakh, H. Hörtnagl, M. Voits, H. Fink, and M. Bader. 2003. "Synthesis of serotonin by a second tryptophan hydroxylase isoform." *Science* 299 (5603):76. doi: 10.1126/science.1078197.
- Wang, H., R. Cao, L. Xia, H. Erdjument-Bromage, C. Borchers, P. Tempst, and Y. Zhang. 2001. "Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase." *Mol Cell* 8 (6):1207-17.
- Wang, H., M. La Russa, and L. S. Qi. 2016. "CRISPR/Cas9 in Genome Editing and Beyond." *Annu Rev Biochem* 85:227-64. doi: 10.1146/annurev-biochem-060815-014607.
- Weber, J., R. Öllinger, M. Friedrich, U. Ehmer, M. Barenboim, K. Steiger, I. Heid, S. Mueller, R. Maresch, T. Engleitner, N. Gross, U. Geumann, B. Fu, A. Segler, D. Yuan, S. Lange, A. Strong, J. de la Rosa, I. Esposito, P. Liu, J. Cadiñanos, G. S. Vassiliou, R. M. Schmid, G. Schneider, K. Unger, F. Yang, R. Braren, M. Heikenwälder, I. Varela, D. Saur, A. Bradley, and R. Rad. 2015. "CRISPR/Cas9 somatic multiplex-mutagenesis for high-throughput functional cancer genomics in mice." *Proc Natl Acad Sci U S A* 112 (45):13982-7. doi: 10.1073/pnas.1512392112.
- Weber, M., I. Hellmann, M. B. Stadler, L. Ramos, S. Pääbo, M. Rebhan, and D. Schübeler. 2007. "Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome." *Nat Genet* 39 (4):457-66. doi: 10.1038/ng1990.
- Weltner, J., D. Balboa, S. Katayama, M. Bernal, K. Krjutškov, E. M. Jouhilahti, R. Trokovic, J. Kere, and T. Otonkoski. 2018. "Human pluripotent reprogramming with CRISPR activators." *Nat Commun* 9 (1):2643. doi: 10.1038/s41467-018-05067-x.
- Whyte, W. A., D. A. Orlando, D. Hnisz, B. J. Abraham, C. Y. Lin, M. H. Kagey, P. B. Rahl, T. I. Lee, and R. A. Young. 2013. "Master transcription factors and mediator establish super-enhancers at key cell identity genes." *Cell* 153 (2):307-19. doi: 10.1016/j.cell.2013.03.035.

Bibliography

- Wolfe, S. A., L. Nekludova, and C. O. Pabo. 2000. "DNA recognition by Cys2His2 zinc finger proteins." *Annu Rev Biophys Biomol Struct* 29:183-212. doi: 10.1146/annurev.biophys.29.1.183.
- Wood, H. B., and V. Episkopou. 1999. "Comparative expression of the mouse Sox1, Sox2 and Sox3 genes from pre-gastrulation to early somite stages." *Mech Dev* 86 (1-2):197-201.
- Wysocka, J., and W. Herr. 2003. "The herpes simplex virus VP16-induced complex: the makings of a regulatory switch." *Trends Biochem Sci* 28 (6):294-304. doi: 10.1016/S0968-0004(03)00088-4.
- Ying, Q. L., M. Stavrdis, D. Griffiths, M. Li, and A. Smith. 2003. "Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture." *Nat Biotechnol* 21 (2):183-6. doi: 10.1038/nbt780.
- Zemach, A., I. E. McDaniel, P. Silva, and D. Zilberman. 2010. "Genome-wide evolutionary analysis of eukaryotic DNA methylation." *Science* 328 (5980):916-9. doi: 10.1126/science.1186366.
- Zhang, Y., C. Yin, T. Zhang, F. Li, W. Yang, R. Kaminski, P. R. Fagan, R. Putatunda, W. B. Young, K. Khalili, and W. Hu. 2015. "CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs." *Sci Rep* 5:16277. doi: 10.1038/srep16277.
- Zhou, L., X. Cheng, B. A. Connolly, M. J. Dickman, P. J. Hurd, and D. P. Hornby. 2002. "Zebularine: a novel DNA methylation inhibitor that forms a covalent complex with DNA methyltransferases." *J Mol Biol* 321 (4):591-9.
- Zhou, V. W., A. Goren, and B. E. Bernstein. 2011. "Charting histone modifications and the functional organization of mammalian genomes." *Nat Rev Genet* 12 (1):7-18. doi: 10.1038/nrg2905.
- Zhu, W. G., K. Srinivasan, Z. Dai, W. Duan, L. J. Druhan, H. Ding, L. Yee, M. A. Villalona-Calero, C. Plass, and G. A. Otterson. 2003. "Methylation of adjacent CpG sites affects Sp1/Sp3 binding and activity in the p21(Cip1) promoter." *Mol Cell Biol* 23 (12):4056-65.
- Zhu, X., Z. Ai, X. Hu, and T. Li. 2016. "Efficient Generation of Corticofugal Projection Neurons from Human Embryonic Stem Cells." *Sci Rep* 6:28572. doi: 10.1038/srep28572.
- Ziller, M. J., J. A. Ortega, K. A. Quinlan, D. P. Santos, H. Gu, E. J. Martin, C. Galonska, R. Pop, S. Maidl, A. Di Pardo, M. Huang, H. Y. Meltzer, A. Gnirke, C. J. Heckman, A. Meissner, and E. Kiskinis. 2018. "Dissecting the Functional Consequences of De Novo DNA Methylation Dynamics in Human Motor Neuron Differentiation and Physiology." *Cell Stem Cell* 22 (4):559-574.e9. doi: 10.1016/j.stem.2018.02.012.

Appendix: Chromatin modifier screen gRNA enrichment



List of tables

Table 1: Common Histone modifying enzymes 19

List of Figures

Figure 1: Waddington’s epigenetic landscape.....	11
Figure 2: Expression of Sox1 during neural development.	16
Figure 3: The dCas9/gRNA complex.	23
Figure 4: Clonal lines express high levels of dCas9 mRNA.	56
Figure 5: Targeted induction of Sox1 in NPCs is significant, but minor.	58
Figure 6: Majority of NPCs is unresponsive to Sox1 induction.	60
Figure 7: GFP and Sox1 levels correlate in Sox1 ^{wt/GFP} NPCs.....	61
Figure 8: GFP induction is stable in induced cells	62
Figure 9: Transcriptomes of Sox1 positive and negative cells cluster apart.....	64
Figure 10: Transcriptome analysis reveals phenotypic changes in NPCs after Sox1 induction	65
Figure 11: Sox1 induction leads to expression of neuroepithelial markers in NPCs.	66
Figure 12: Sox1 positive NPCs regain neuronal differentiation potential.....	68
Figure 13: Discrepancy in response to induction not based on technical issues	71
Figure 14: Quantification of histone modifications at the Sox1 promoter	73
Figure 15: DNA methylation strongly reduced in Sox1 expressing NPCs.....	74
Figure 16: DNA Methylation at a control locus does not vary.....	75
Figure 17: Combination of dCas9-VP64 and dCas9-Tet1 leads to an increase in the amount of responsive NPCs....	78
Figure 18: DNA methylation as barrier to transcriptional editing is not specific to Sox1	81
Figure 19: Sox20000 library positions..	82
Figure 20: Enrichment of Sox1 ^{GFP} -positive cells over three subsequent sorts.	84
Figure 21 (next page): gRNAs are equally enriched over the Sox1 locus.....	85
Figure 22: gRNA candidates are functional.....	87
Figure 23: Sox1 transcription is blocked by promoter DNA methylation in NPCs	103

List of abbreviations

(l)ncRNA	(long) non-coding RNA
5caC	5-Carboxymethylcytosine
5fC	5-formylcytosine
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
Actc1	Actin alpha cardiac muscle 1
aRGC	apical radial glia cell
BDNF	Brain-derived neurotrophic factor
BER	base excision repair
bFGF	basic fibroblast growth factor
bp	base pair
BSA	Bovine Serum Ablumin
cAMP	cyclic AMP
Cas	CRISPR associated protein
cDNA	complementary DNA
ChIP	Chromatin immuno precipitation
CNS	central nervous system
CpG	Cytosine-phosphate-Guanine
CRISPR	Clustered regularly interspaced palindromic repeats
Ctcf	CCCTC binding factor
Dapi	4',6-Diamino-2-Phenylindole
dCas9	deactivated Cas9
DMSO	Di-methyl-sulfoxide
DNA	Desoxiribonucleic acid
Dnmt1	DNA methyltransferase 1
Dnmt2	DNA methyltransferase 2
Dnmt3	DNA methyltransferase 3
dTet1	deactivated Tet1
E2F-1	E2F transcription factor 1
ECL	Electrochemiluminescence
EDTA	Ethylendiamitetraacetic Acid
EGF	epithelial growth factor

List of abbreviations

ESC	embryonic stem cell
FACS	Fluorescent Activated Cell Sorting
Gapdh	Glyceraldehyde 3-phosphate dehydrogenase
GDNF	Glial cell line-derived neurotrophic factor
GFAP	glial fibrillary acidic protein
GFP	green fluorescent protein
gRNA	guide RNA
H2A/B	histone 2 A/B
H3	histone 3
H3K4/9/27/36	Histon 3 lysine 4/9/27/36
H3K4/9/27/36 me3	Trimethylation of histon 3 lysine 4/9/27/36
HEK	human embryonic kidney cells
ICC	immunohistochemistry
IgG	Immunoglobulin G
Il1rn	Interleukin 1 receptor antagonist
Jmjd2a	Lysine-specific demethylase 4A
kb	kilobases
Klf4	Krüppel like factor 4
LB medium	lysogeny broth medium
Map2	Microtubuli associated protein 2
miRNA	micro RNA
MOI	multiplicity of infection
mRNA	messenger RNA
MyoD	myogenic differentiation 1
NEC	neuroepithelial cell
Ngn2	Neurogenin 2
NGS	next generation sequencing
NPC	neural progenitor cell
NR	neural rosette
NSC	neural stem cell
NT3	Neurotrophin 3
Oct4	Octamere binding factor 4
ORF	open reading frame
PAGE	Polyacrylamide gel electrophoresis

List of abbreviations

PAM	protospacer adjacent motive
PBS	phosphate buffered saline
PCA	Principal Component Analysis
PSC	pluripotent stem cell
qPCR	quantitative PCR
rcf	relative centrifugal force
RIPA	radioimmunoprecipitation assay
RNA	ribonucleic acid
RNA-seq	RNA Sequencing
S100beta	S100 calcium-binding protein B
SDS	Sodium Dodecyl Sulfate
Set7	Histone-lysine N-methyltransferase SETD7
siRNA	small interfering RNA
Sox1	Sex-determining-region-y-box 1
Sox2	Sex-determining-region-y-box 2
Sox3	Sex-determining-region-y-box 3
Sp1	specificity protein 1
Stagr	String assembly gRNA
TAD	topologically associated domain
TAF	transcription activating factor
TALE	Transcription activator like effector
TEMED	N,N,N',N'-Tetramethyl ethylenediamine
Tet	Ten eleven translocation
TH	Tryptophane hydroxylase
TSS	transcription start site
Tuj1	Neuron-specific class III beta-tubulin
vGlut1	Vesicular glutamate transporter 1
wt	wild type
YY1	Yin Yang 1
ZNF	Zinc finger protein
Zo-1	zona occludens 1

List of Publications

- Baumann, V.**, M. Wiesbeck, C. T. Breunig, J. M. Braun, A. Koflerle, J. Ninkovic, M. Gotz, and S. H. Stricker. 2019. "Targeted removal of epigenetic barriers during transcriptional reprogramming." *Nat Commun* 10 (1):2119. doi: 10.1038/s41467-019-10146-8.
- Breunig, C. T., T. Durovic, A. M. Neuner, **V. Baumann**, M. F. Wiesbeck, A. Köferle, M. Götz, J. Ninkovic, and S. H. Stricker. 2018. "One step generation of customizable gRNA vectors for multiplex CRISPR approaches through string assembly gRNA cloning (STAgR)." *PLoS One* 13 (4):e0196015. doi: 10.1371/journal.pone.0196015.
- Köferle, A., K. Worf, C. Breunig, **V. Baumann**, J. Herrero, M. Wiesbeck, L. H. Hutter, M. Götz, C. Fuchs, S. Beck, and S. H. Stricker. 2016. "CORALINA: a universal method for the generation of gRNA libraries for CRISPR-based screening." *BMC Genomics* 17 (1):917. doi: 10.1186/s12864-016-3268-z.
- Balsevich, G., **V. Baumann**, A. Uribe, A. Chen, and M. V. Schmidt. 2016. "Prenatal Exposure to Maternal Obesity Alters Anxiety and Stress Coping Behaviors in Aged Mice." *Neuroendocrinology* 103 (3-4):354-68. doi: 10.1159/000439087.
- Bellisario, V., P. Panetta, G. Balsevich, **V. Baumann**, J. Noble, C. Raggi, O. Nathan, A. Berry, J. Seckl, M. Schmidt, M. Holmes, and F. Cirulli. 2015. "Maternal high-fat diet acts as a stressor increasing maternal glucocorticoids' signaling to the fetus and disrupting maternal behavior and brain activation in C57BL/6J mice." *Psychoneuroendocrinology* 60:138-50. doi: 10.1016/j.psyneuen.2015.06.012.

Copyright information

Fig. 1: from (Bard 2008)

Title: Waddington's Legacy to
Developmental and Theoretical
Biology
Author: Jonathan B. L. Bard
Publication: Biological Theory
Publisher: Springer Nature
Date: Jan 1, 2009
License # 4544730862244

Fig. 2: from (Perez-Pinera et al. 2013)

Title: RNA-guided gene activation by
CRISPR-Cas9-based
transcription factors
Author: Pablo Perez-Pinera, D Dewran
Kocak, Christopher M Vockley,
Andrew F Adler, Ami M Kabadi et
al.
Publication: Nature Methods
Publisher: Springer Nature
Date: Jul 25, 2013
License # 4544731086880

Fig. 3: from (Tang, Yu, and Cheng 2017)

Title: Current progress in the
derivation and therapeutic
application of neural stem cells
Author: Yuewen Tang, Pei Yu, Lin Cheng
Publication: Cell Death & Disease
Publisher: Springer Nature
Date: Oct 12, 2017

Open Access publication, no permission required according to publisher's policies

Acknowledgements

I want to thank Stefan Stricker for the guidance during this project, but also for giving me the freedom to develop it according to my ideas. Thank you for the great supervision.

I also want to thank the Stricker Lab as a whole for the great atmosphere and the loads of fun I had working with you for the past years. It's time for lunch, muchachos! In particular, I want to thank Julia Braun, Maximilian Wiesbeck, and Christopher Breunig for your help with and contribution to this work.

I would like to thank all members of my thesis committee, Gunnar Schotta, Jovica Ninkovic, and Wolfgang Wurst for advice and ideas.

Chris, thanks for the many hours of fun during nerd sessions, new year's parties, and hikes! I think I still owe you that glass...

Maxi, thank you for teaching me the art of Kicker!

Julia, for teaching me the difference between Baden and Schwaben. Repeatedly.

Freddy, thank you for simply everything. You make it all so much easier!

Lastly, I want to thank my family. You are too many to mention by name, so you'll just have to believe that I mean all of you. Thank you all for your support!

Eidesstattliche Versicherung/Affidavit

Valentin Baumann

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation

“Identification and manipulation of chromatin barriers in transcriptional reprogramming”

selbstständig angefertigt habe, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

I hereby confirm that the dissertation

“Identification and manipulation of chromatin barriers in transcriptional reprogramming”

is the result of my own work and that I have only used sources or materials listed and specified in the dissertation.

12.08.2019

München, den/ Munich, date

Valentin Baumann

Unterschrift/ signature

Declaration of author contributions

Valentin Baumann (V.B.) wrote the thesis text and prepared all except reprinted figures presented in the thesis.

Dr. Stefan Stricker (S.S.) and V.B. designed the study and the experimental approach.

V.B., Julia Braun (J.B.), and Maximilian Wiesbeck (M.W.) conducted the experiments (for a detailed listing, see table below).

V.B. collected and analyzed the data.

Dr. Anna Köferle (A.K.) and V.B. cloned dCas9-fusion vectors

V.B. designed all gRNA sequences.

V.B., C.B., and M.W. cloned gRNA library plasmids and cloned the gRNA libraries.

V.B. and M.W. conducted all virus work, including the generation of clonal NPC lines.

S.S. isolated mouse muscle and heart tissue for RNA preparation.

S.S. generated NSCs and NPCs from ESCs.

Contributing Author	Figure number
Valentin Baumann	4, 5, 6, 7, 8, 9, 10, 11, 12A-C, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22
Maximilian Wiesbeck	5, 6, 11A and B, 12D
Julia Braun	6, 14, 15, 16C, 19, 20B, 21