Molecular characterisation of the transcription factor – chromatin landscape interplay during neuronal cell fate acquisition

Dissertation

submitted to attain the academic degree

“Doctor of Natural Sciences”

at the Department of Biology

of the Johannes Gutenberg University Mainz

by

Johannes Jung

born on 14.05.1985 in Limburg an der Lahn

Mainz, 2017

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung: 27.09.2018

"The man in black fled across the desert and the gunslinger followed"

Stephen King

Table of contents

Abstract

A multicellular organism consists of a huge variety of different cell types that are all derived from the totipotent zygote. In order to achieve the creation of such diversity from a single cell, a fine-tuned regulation of the genomic output is essential. With regard to gene regulation it is increasingly appreciated that cell fate changes are coordinated by multiple regulatory layers on both, the genetic as well as the epigenetic level. The concerted action of both layers integrates cell-intrinsic and cell-extrinsic cues during development to enable for dynamic cellular responses that become manifested in differential gene expression profiles.

This thesis focused on the molecular regulation of cell fate changes during embryonic neuronal differentiation in *Mus musculus*. Although several transcription factors and epigenetic mechanisms have been described in the context of embryonic neurogenesis, a full understanding yet remains elusive. In this thesis, molecular biology and neurobiology techniques were applied in addition to high-throughput sequencing technologies and computational analyses to explore the impact of two neuron-specific transcription factors on the developmental chromatin landscape. The first part of this thesis investigated the transcription factor NeuroD1. This protein was found to occupy *cis*-regulatory elements of neuronal genes in a sequence specific fashion. Interestingly, a significant fraction of such genomic loci were embedded in an inactive chromatin landscape prior to NeuroD1's binding. NeuroD1's occupancy of these sites initiated their conversion into an active euchromatin state and hence led to the induction of the neuronal developmental program. Several identified NeuroD1 target genes encompassed transcriptional regulators, suggesting that this protein initiated a cascade of nuclear factors that further mediated neuronal fate determination. Strikingly, the transient action of NeuroD1 was sufficient to induce a sustained shape of the neuronal chromatin landscape that persisted after its disappearance during later stages of development. This characterisation of NeuroD1's molecular function suggested that it might act as a pioneer transcription factor. In the second part of this thesis Zfp354c was identified as a novel KRAB-zinc-finger protein that interacted with KAP1 to confer H3K9me3-mediated silencing of repetitive elements during neurogenesis. Both, its depletion *in vitro* and *in vivo*, impaired neuronal differentiation, possibly due to the de-repression of endogenous retroviral elements.

In summary, this thesis provides novel insights into coordinated chromatin landscape dynamics during neuronal cell fate acquisition and advocates for the importance of both, the coding as well as the non-coding genome, during this process. These findings open new avenues for the understanding of regulated cellular differentiation and their underlying transcriptome dynamics in the context of neurogenesis. This knowledge can potentially be used for therapeutic reprogramming attempts to ectopically restore degenerated neuronal cell populations.

Zusammenfassung

Ausgehend von einer einzigen totipotenten Zygote wird eine Vielzahl mannigfaltiger Zelltypen gebildet aus denen letztlich ein lebensfähiger, mehrzelliger Organismus hervorgeht. Um dies sicher zu stellen bedarf es einer fein abgestimmten Regulation des genomischen Outputs. Zahlreiche Studien zu dieser Thematik verdeutlichen, dass Veränderungen des zellulären Entwicklungsweges sowohl genetisch als auch epigenetisch über mehrere regulatorische Mechanismen koordiniert werden. Ein engmaschig reguliertes Netzwerk von aufeinanderfolgenden Prozessen integriert hierbei sowohl zellintrinsische als auch zellextrinische Signale während des Entwicklungsprozesses, um dynamische zelluläre Antworten in differentiellen Genexpressionsprofilen zu manifestieren.

In der vorliegenden Dissertation wurden schwerpunktmäßig die Mechanismen der molekularen Regulation von zellulären Entwicklungsschicksalen während der embryonalen neuronalen Differenzierung in *Mus musculus* behandelt. Obwohl im Kontext der embryonalen Neurogenese bereits eine Vielzahl an Transkriptionsfaktoren und die Bedeutung epigenetischer Mechanismen beschrieben wurden, ist ein umfassendes Verständnis aller beteiligten Faktoren sowie deren Zusammenspiel im zeitlichen und räumlichen Kontext nicht umfassend verstanden. Daher wurden neben Hochdurchsatz-Sequenzierungstechnologien und bioinformatischen Analysen molekularbiologische und neurobiologische Techniken eingesetzt, um den Einfluss zweier neuronen-spezifischer Transkriptionsfaktoren auf die entwicklungsimmanente Chromatinlandschaft zu untersuchen. Der erste Teil dieser Arbeit untersuchte den Transkriptionsfaktor NeuroD1. Die durchgeführten Versuche ergaben, dass dieses Protein sequenzspezifisch an *cis*-regulatorische Elemente von neuronalen Genen bindet. Interessanterweise waren eine Vielzahl dieser genomischen Abschnitte vor der NeuroD1-Bindung in einer inaktiven Chromatinlandschaft eingebettet. Die Bindung von NeuroD1 an diese Elemente initiierte deren Umwandlung in einen aktiven Chromatinzustand und führte dabei zur Induktion des neuronalen Entwicklungsprogramms. Mehrere identifizierte NeuroD1-Zielgene umfassten weitere Transkriptionsregulatoren, was nahelegt, dass dieser Faktor eine Kaskade von nuklearen Faktoren aktiviert, die den neuronalen Entwicklungsweg weiter bestimmen. Interessanterweise war die transiente Wirkung von NeuroD1 ausreichend, um eine nachhaltige Manifestation der neuronalen Chromatinlandschaft zu induzieren, die auch in dessen Abwesenheit in der späteren Entwicklung bestehen bleibt. Diese Charakterisierung der molekularen Funktion von NeuroD1 deutete darauf hin, dass es sich um einen Pionier-Transkriptionsfaktor handeln könnte. Der zweite Teil dieser Dissertation identifizierte Zfp354c als ein neues KRAB-Zink-Finger-Protein das mit KAP1 interagiert, um eine H3K9me3-vermittelte Repression repetitiver genomischer Elemente während der Neurogenese zu gewährleisten. Sowohl seine *In-vitro*- als auch *In-vivo*-Depletion beeinträchtigte neuronale Differenzierungsprozesse, möglicherweise aufgrund der Reaktivierung von endogenen retroviralen Elementen.

Zusammengefasst liefert diese Dissertation neue Einblicke in die Mechanismen und Dynamiken der Modulation der Chromatinlandschaft während der Etablierung einer neuronalen Zellidentität und verdeutlicht die Bedeutung sowohl des kodierenden als auch des nicht kodierenden Genoms in diesem Prozess. Diese Erkenntnisse tragen zum umfassenderen Verständnis der Regulation zellulärer

Differenzierung und ihrer zugrundeliegenden Transkriptomdynamiken im Kontext der Neurogenese bei. Dieses Wissen kann möglicherweise für therapeutische Reprogrammierungsversuche verwendet werden, um degenerierte neuronale Zellpopulationen ektopisch wiederherzustellen.

1 Introduction

1.1 The cell as the entity of life

Life exists for several billion years on earth, although its origin is still not fully unravelled [1-4]. The cell is the fundamental entity of life known on earth: First described in 1665 by Robert Hook, Jakob Schleiden and Theodor Schwann founded the cell theory (1838/1839) which explains that all organisms are composed of cells. Rudolph Ludwig Karl Virchow further defined the cell theory in 1855 by stating “omnis cellula e cellula” [5]. The eukaryotic cell consists of two main compartments which are surrounded by a membrane: the nucleus and the cytoplasm. After Friedrich Miescher first isolated “nuclein” (nucleic acid) [6], the term chromatin was coined in 1879 by Walther Flemming for the “highly phosphorus organic acid” found in the nucleus [7]. This organelle contains the majority of the genetic and heritable information of an eukaryotic organism: The genetic information is stored in a linear structure, the deoxyribonucleic acid (DNA) double helix [8], which defines as a chemical entity, the genome, the identity of each cell and hence multicellular organisms. This polymer consists of four nucleotides which comprise the nucleobases adenine (A), guanine (G), cytosine (C) and thymine (T), linked by a sugar-phosphate-backbone. The order of these nucleotides (DNA sequence) in specific genomic regions (genes) encodes for biological active RNAs (ribonucleic acids), which consecutively mediate the cellular phenotype either directly in the nucleus or indirectly after their cytoplasmic translation into proteins (“DNA code” of genes). However, the functional output of the genome is not simply a mirror picture of the DNA sequence: The stored information of the genome is modulated via several cellular and molecular events like nuclear architecture, chromatin structure or epigenetic modifications (see 1.4 and 1.5), which together build up the genomic output while integrating the cellular environment [9-11]. In the course of evolution, from the origin of life around 3,900 million years ago (Mya) and the emergence of the eukaryotic cell (around 2,100 Mya) onwards, a huge diversity has emerged in the tree of life. The first multicellular eukaryote evolved around one billion years ago [12, 13]. However, the fossils of the first animals are dated to more than 600 Mya, marking the rise of multicellular species in the animal kingdom [14-16]. A multicellular organism develops from a single cell during a process termed ontogeny (see 1.2). A primary focus of developmental biology research till date are cell-intrinsic and -extrinsic events which regulate the genome in order to give rise to the distinct cell types of the complex mature organism. These events enable for example the emergence of more than 200 cell types in humans [17], which share with a few exceptions all the same genome of the zygote. In order to generate this variety of cells during ontogeny, two key principles are important: (1) Cell proliferation, a process through which the number of cells increases due to cell division and cell growth [18]. (2) Cell differentiation, during wich the mature, distinct and functional form of a cell is generated from an immature (less specialised) younger precursor [18]. In a recent publication, Morris and colleagues [19] summarise several terminologies important for the further understanding of these two processes: (1) Cell fate: “The developmental destination of a cell if left undisturbed in its environment. The fate of a cell is more restricted than its potential” [19]. (2) Cell states: “The transcriptional output of a gene regulatory network, with a variable degree of stability; development is characterised by sequences of cell states that culminate in specific fates” [19]. (3) Cell

potential: “Biologically, potentials represent the range of fates into which a cell can develop. It is reduced during development and is obscured in, for example, lineage-tracing experiments, which only reveal fates. In physics, potential can be described as the ability to do work and represents an amount of energy stored for that purpose. In both biology and physics, it represents an ability to do something” [19]. The later shows several distinct grades of existence which lead from high to low potential. So called totipotent cells like the zygote exhibit the ability to generate all cell types of an organism with their various phenotypes. In contrast, the potency of embryonic stem cells is more restricted as they are able to differentiate into any cell type of the three germ layers but not extraembryonic tissues and are therefore considered as pluripotent. While cells differentiate into a certain fate they step-wise loose the potential of generating cells of other germ layers and become more restricted and specialised what for example ensures that within a certain tissue only certain cell types are generated.

1.2 Murine embryonic development

The process in which an embryo develops from a fertilized egg, the totipotent zygote, is referred to as embryogenesis. This process encompasses for example that (1) cells have to increase in number due to proliferation and (2) differentiation processes increase the number of different cell types [20]. During early embryonic development, the zygote matures into increasingly complex multicellular structures accompanied by the rise of distinct cell types [21, 22]. After initial cleavages, the multicellular ball-like morula develops into the blastocyst comprising the inner cell mass (ICM) and the trophectoderm (Figure 1), which is derived from the outer cells of the morula and contributes to the development of the placenta.

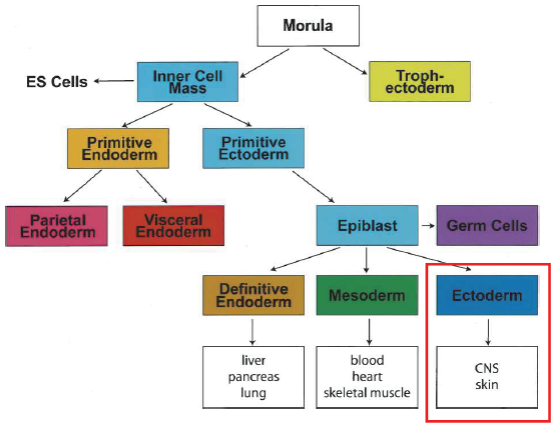


Figure 1: Scheme of early mouse development depicting the relationship of early cell populations to the primary germ layers.
The definitive endoderm, mesoderm and ectoderm are derived from the epiblast. The lineage of interest in this thesis is highlighted with a red box. Figure modified from [23].

Cells of the ICM differentiate into two lineages: (1) the primitive endoderm (PE), which contributes mostly to the extra-embryonic yolk sac and (2) the primitive ectoderm, which gives rise to the epiblast (EPI). During gastrulation, which starts around embryonic day 6.5 (E6.5), the three primary germ layers are derived from

the epiblast: ecto-, meso- and endoderm. These three cell populations give rise to all different cell types of the organisms in a tightly regulated spatio-temporal process. However, germ cells are an exception because they are derived earlier directly from the epiblast. In total, after approximately 19 days of murine embryonic development, the young mouse is born.

1.2.1 Development of the early mouse embryo

The early developmental period of the mouse embryo ranges from the fertilized egg, which originates from the fusion of a sperm and an oocyte, to the gastrulating embryo (Figure 2). During this period, the aforementioned lineage specification events take place under tight regulation where a single cell, the totipotent zygote, advances to a multicellular egg cylinder [20]. Throughout this process, more specialized cell types are generated which is accompanied by a gradual loss of their developmental potential.

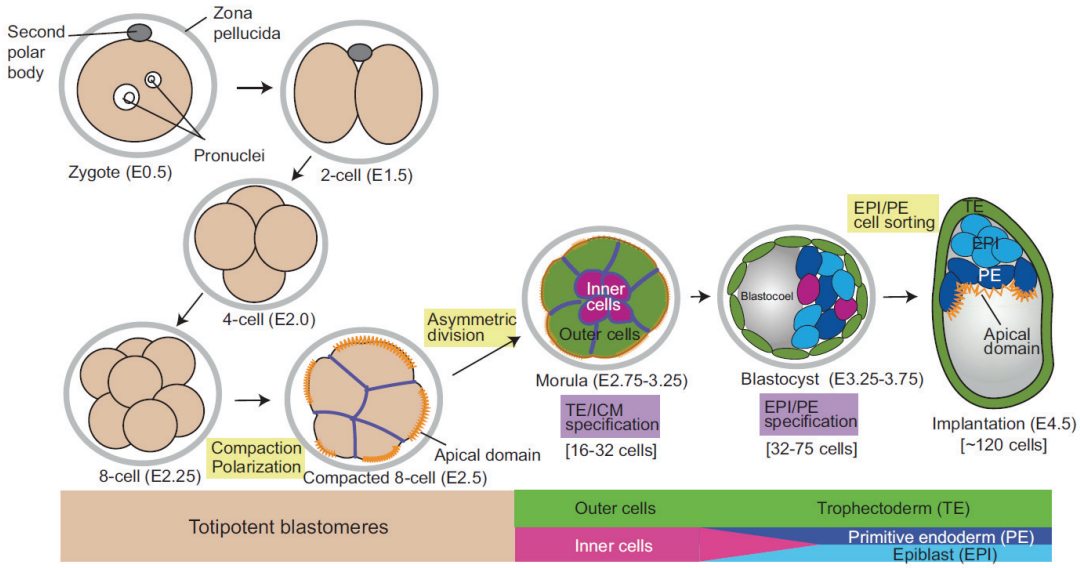


Figure 2: Preimplantation mouse development.
"Schematic of the morphological changes and lineage differentiation steps of mouse preimplantation embryo development. The coloured bars show the sequential lineage progression from totipotent blastomere to the first three lineages: trophectoderm (green), epiblast (light blue) and primitive endoderm (dark blue). Important morphogenetic events are highlighted in yellow. Lineage specification events are highlighted in purple. Orange lines indicate the apical domains of cells." Figure and legend taken from [22].

After the initial three rounds of symmetric cell divisions, the first asymmetric cell divisions take place in the compacted 8-cell stage, which introduces an inside-outside polarity of the embryo (Figure 2): The preimplantation embryo develops into the morula with two distinct cell types, the inner and outer cells, which can be distinguished by the expression of the transcription factors (TFs) SRY-box 2 (Sox2) and inhibitor of DNA binding 2 (Id2), respectively [24, 25]. The blastomeres already exhibit a heterogeneity in their transcriptome as early as from the 2 - 4 cell stage onwards [25-29]. Subsequently, the embryo advances into the blastocyst, which is still surrounded by the zona pellucida. At this stage, the first three cell lineages of the embryo have been established which are characterised by a distinct expression profile: the

trophectoderm (TE), primitive endoderm (PE) and the epiplast (EPI) [21, 25, 30-33]. Interestingly, the specification of ICM cells into EPI/PE cells is accompanied by a distinct 'salt & pepper' pattern of Nanog and GATA binding protein 6 (Gata6) at E 3.75 [25, 34-36]. It has been shown that also Sox2 and Octamer-binding transcription factor 4 (Oct4) are important regulators in the process of EPI/PE formation in addition to Nanog and Gata6 [22, 24, 37, 38]. Apart from transcription factors, the role of diverse signaling pathways like the Notch-, LIF-JAK/STAT-, TGF-β- and FGF-pathways during this process becomes increasingly appreciated [39].

The transition of the morula into the blastocyst, encompassed by a cavitation process, is a key event in cell fate acquisition during ontogeny and has to be tightly controlled [21, 25]. The first critical lineage decisions take place during early embryonic development with the establishment of the three embryonic germinal sheets.

1.2.2 Embryonic stem cells as a model system to study cellular differentiation

Ernst Haeckel coined the term "stem cell" in 1868 for the ancestral unicellular organism ("Stammzelle") from which he believed the multicellular organisms ("Stämme") evolved [40]. Later he also used this term for the zygote as the origin of all cells of an organism [41]. Towards the end of the 19th century the common precursor of the hematopoietic system was referred to as "stem cell" [42], which had been demonstrated experimentally later [43-45]. The ability of a cell to differentiate into cells of all lineages as well as to self-renew is referred to as "pluripotency" as described above. The first research in this direction was conducted on inbreed strains of mice with testicular teratomas, which were composed of many types of embryonic and adult tissues and grew continuously [46]. Further investigations of mouse teratoma (benign) and teratocarcinomas (malignant) and their *in vitro* culture established the link between stem cells and pluripotency and coined the term embryonal carcinoma (EC) cells [47-53]. Taken together, this line of research initiated the current definition of stem cells as cells which exhibit (1) the capacity to give rise to differentiated (more specialized) cells and (2) also have the ability to self-renew [44, 54-56]. Gail Martin and Martin Evans found that clonal EC cells were able to form cellular aggregates ("embryoid bodies") which exhibited a differentiation potential like early embryos during *in vivo* development [57]. Concomitant research on optimised conditions for the continuous culture of EC cells *in vitro*, including the usage of mitotically inactive fibroblast as feeder cells and the presence of serum in the culture medium, led to the first successful isolation of murine embryonic stem cells (mESCs) from the inner cell mass of a mouse blastocyst [58, 59]. From then on, mESCs have been shown to be germline-competent and were used to introduce transgenes into mice by retroviral transfection of mESCs [60] as well as to generate targeted gene modifications *in vivo* by homologous recombination in mESCs *in vitro* [61]. Their property of being pluripotent distinguishes mESCs from other stem cells, as they normally cannot give rise to every cell type in an organism and can exist as tissue-specific populations throughout ontogeny in contrast to mESCs. The culture conditions for mESCs were refined by the discovery of leukemia inhibitory factor (LIF), which inhibited the differentiation of mESCs cultured without feeder cells as this glycoprotein is primarily localized on the cellular surface of feeder cells [62-66]. LIF was found to be a crucial factor for mESC self-renewal by

activating the Janus-associated kinase (JAK) and hence the action of signal transducer and activator of transcription 3 (Stat3), which mediates the transcriptional response to maintain pluripotency [67-70]. The indefinite culture of mESCs in serum supplemented with LIF on gelatin coated plates leads to a pluripotency state which resembles an *in vivo* developmental stage close to the naive epiblast/primitive ectoderm in the late blastocyst [71-76]. However, several spatio-temporal distinct pluripotency states have been described [77].

Intense research identified that the tight regulation and interplay of three TFs build up the core of the pluripotency gene regulatory network (PGRN) by regulating themselves via feed-back and feed-forward loops: Oct4, Sox2 and Nanog [24, 37, 38, 78-88]. Interestingly, all of these TFs have been shown to be important for *in vivo* development as described above. Further in depth characterisation of the PGRN on a molecular basis allowed deeper understanding of embryonic development and established ESCs as a model system to apprehend this process. Their potential to differentiate into cell types of the mesoderm, endoderm and ectoderm lineage led foundation for several differentiation protocols for different cell populations *in vitro* [23, 89, 90] and keeps strong promise for application in regenerative medicine [23, 91-94]. These tools enable researchers to study the earliest stages of lineage induction and specification, which otherwise would be more challenging *in vivo*. Especially the generation of different types of neural cells as a derivate of the ectoderm has been an asset to understand cell-fate determination during neuronal differentiation (Figure 3).

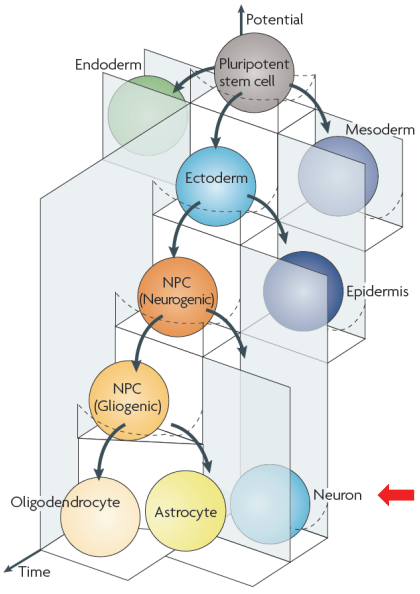


Figure 3: The differentiation potential of cells diminishes during neuronal differentiation
Embryonic Stem cells can give rise to all cell types of the three germ layers. During differentiation along the neural lineage, the potential of neural/neuronal stem cells (NPCs) diminishes over time and their fate is restricted to distinct neural cell types. The cell type of special interest in this thesis is marked with a red arrow. Figure and legend adopted from [95].

A standard method to generally induce differentiation of mESCs *in vitro* is based on the withdrawal of LIF as well as an adhesive substrate, leading to the formation of ball-shaped cellular aggregates (CAs), which were termed embryoid bodies (EBs) (reviewed in [23]). These EBs have been shown to develop further into

cells with neuronal properties after exposure to retinoic acid (RA) as well as serum depletion [96]. Although this protocol underwent several refinements to promote differentiation of EBs along neuroectoderm lineage, the generated neuronal cells were still heterogonous [97-101]. In 2004, Miriam Bibel and colleagues published a neuronal differentiation protocol for mESCs which generates a highly pure population of glutamatergic neurons [102]. These neuronal cells were found to recapitulate embryonic neurogenesis as they were also derived from Pax6-positive neural progenitors (NPs) and exhibited gene expression dynamics comparable to their *in vivo* counterpart [102-106]. Interestingly, the developmental potential of these *in vitro* derived NPs was found to be restricted towards the generation of neurons in the central nervous system (CNS) rather than peripheral sensory neurons [107]. This protocol serves as a source to study embryonic neurogenesis *in vitro* on a highly homogenous cell population, allowing for epigenome profiling of cellular events during differentiation.

1.2.3 Murine embryonic neurogenesis

Animals with nervous systems are found throughout Bilateria as well as in Cnidara and Ctenophora [108, 109]. Interestingly however, neurons as well as complex brains seem to have evolved independently several times [110]. The nervous system originates from the neuroectoderm and is highly diverse throughout the animal kingdom as exemplified by comparing the nerve net of *Caenorhabditis elegans* with its 302 neurons [111] to the human brain containing hundreds of millions to billions of neurons [112].

The nervous system of bilaterally symmetric animals is generally divided into the peripheral nervous system (PNS) and the central nervous system (CNS), which consists of the brain and the spinal cord. Two main cell types are found in the CNS: (1) neurons, which are the basic computational unit of the CNS and (2) glial cells like astrocytes or oligodendrocytes, which play a major role in regulating the homeostasis in the CNS [113]. The majority of neurons in the CNS exhibit a distinct morphology due to their characteristic cellular processes, the dendrites and axons, which are important for the development of a synapse network as a foundation for inter-cell communication. Although the diversity of neuronal cell types was first described by Santiago Ramon y Cajal in 1899 [114], a complete picture of all subtypes and when and how decisions for subtype specifications are taken still remains elusive until today. One aim of modern neurobiology is to understand the molecular mechanisms of neurogenesis and neuronal diversity. This harbors the promise to be utilized for therapeutic intervention towards neurodegenerative diseases like Alzheimer's or Parkinson disease. This is of utmost importance as the vast majority of all neurons is generated prenatally and are afterwards kept in a post-mitotic state throughout ontogeny, whereas the formation of synapses, axon myelination and gliogenesis take place mostly postnatally.

1.2.3.1 Development of the central nervous system and the neocortex

During embryonic development, the ectoderm separates into the epidermal ectoderm and the neuroectoderm. The neuroectoderm gives rise to the neural crest and the neural tube during neurulation [115, 116]. Whereas the neural crest forms components of the PNS, the neural tube gives rise to the CNS [116]: Its posterior region develops into the spinal cord whereas the anterior part of the neural tube

progresses into three primary vesicles, namely the prosencephalon (forebrain), mesencephalon (midbrain) and the rhombencephalon (hindbrain). These regions further subdivide into different compartments: The prosencephalon gives rise to the telencephalon (endbrain) and the diencephalon (interbrain), whereas the rhombencephalon creates the metencephalon and myelencephalon (Figure 4).

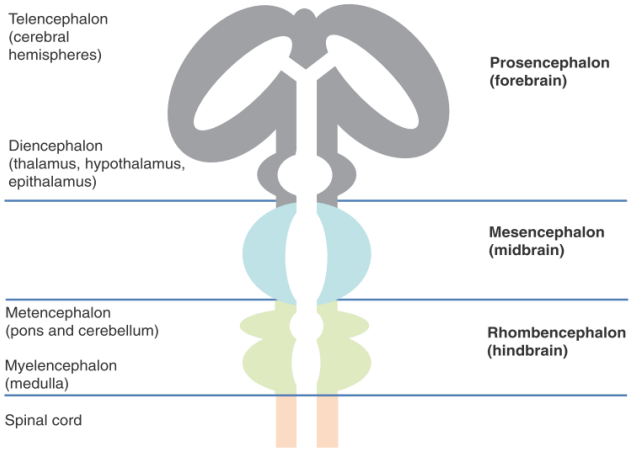


Figure 4: Schematic view of the developing brain subdivisions from the neural tube.
Figure taken from [117].

The embryonic telencephalon can be divided into two main regions, the ventral subpallium and the dorsal pallium (Figure 5). The mammalian neocortex is the newest addition during brain evolution, emerges from the pallium and fulfills associative and analytical functions [118-120]. The mature neocortex consists of approximately 12 million glial cells and 14 million neurons in *Mus musculus* [121]. Interestingly, only around 20% of these cortical neurons are inhibitory, whereas around 80% are excitatory neurons [122]. The inhibitory neurons are generated in the ganglionic eminences of the ventral telencephalon and migrate tangentially into the neocortex, while the glutamatergic excitatory neurons of the neocortex originate from germinal zones directly in the dorsal telencephalon [123-131].

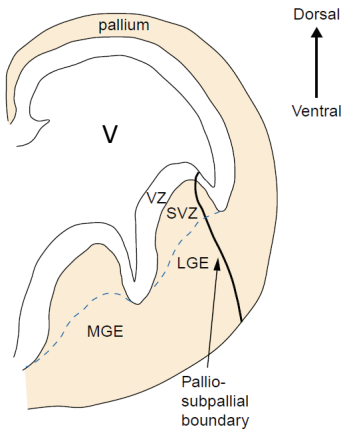


Figure 5: Scheme of a coronal hemisection of the murine telencephalon at E12.5.
"The scheme shows morphologically defined structures and the progenitor subdomains of the embryonic telencephalon. The ventricular zone (VZ) extends along the whole DV axis and contains precursor cells. The subventricular zone (SVZ, indicated by broken blue lines) also contains precursor cells, which is a unique feature of the telencephalon. Note the placement of the DV boundary (i.e. pallio-subpallial boundary) in the dorsal portion of the LGE. LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence." Figure adopted and legend taken from [132]. V, ventricle; DV, dorso-ventral

Introduction

The murine neocortex is generated from E10.5 to E18.5 and comprises a complex layered cytoarchitecture across its radial dimension [127, 133-135]. At the onset of neurogenesis, neuroepithelial cells (NEC) establish the ventricular zone (VZ) by generating radial glial (RG) cells [136-139]. These RG cells exhibit an apical-basal polarity due to their radial processes ranging from the lateral ventricle (apical) to the pial (basal) surface whereas their cell body resides in the VZ. They have the capacity to self-renew and also display astroglial properties [136, 137, 140]. At the beginning of neurogenesis, these neural stem cells switch from symmetric to asymmetric cell division, thereby giving rise to a new RG cell as well as to a differentiating cell such as a newborn neuron or intermediate neuronal progenitor (IP) [119, 134, 141, 142]. These IPs establish the subventricular zone (SVZ) as a second germinal layer dorsal to the VZ and further give rise to neurons directly or after symmetric proliferative divisions [119, 134, 141, 143-145]. The newborn cortical neurons migrate along the radial scaffold build by RG cells to the outer layer of the developing cortex across previously established layers, leading to an “inside-out” manner of corticogenesis (Figure 6): early-born neurons reside in deeper neocortical layers (layer VI and V) and late-born neurons migrate to more superficial layers (layer IV and II/III) [133, 134]. Interestingly however, the Cajal–Retzius neurons which reside in the layer I are born first in distinct focal sites during early neurogenesis [133, 146]. The six-layered structure of neurons in the neocortex is called cortical plate (CP), which is separated from the VZ/SVZ layers by an intermediate zone (IZ) during embryogenesis.

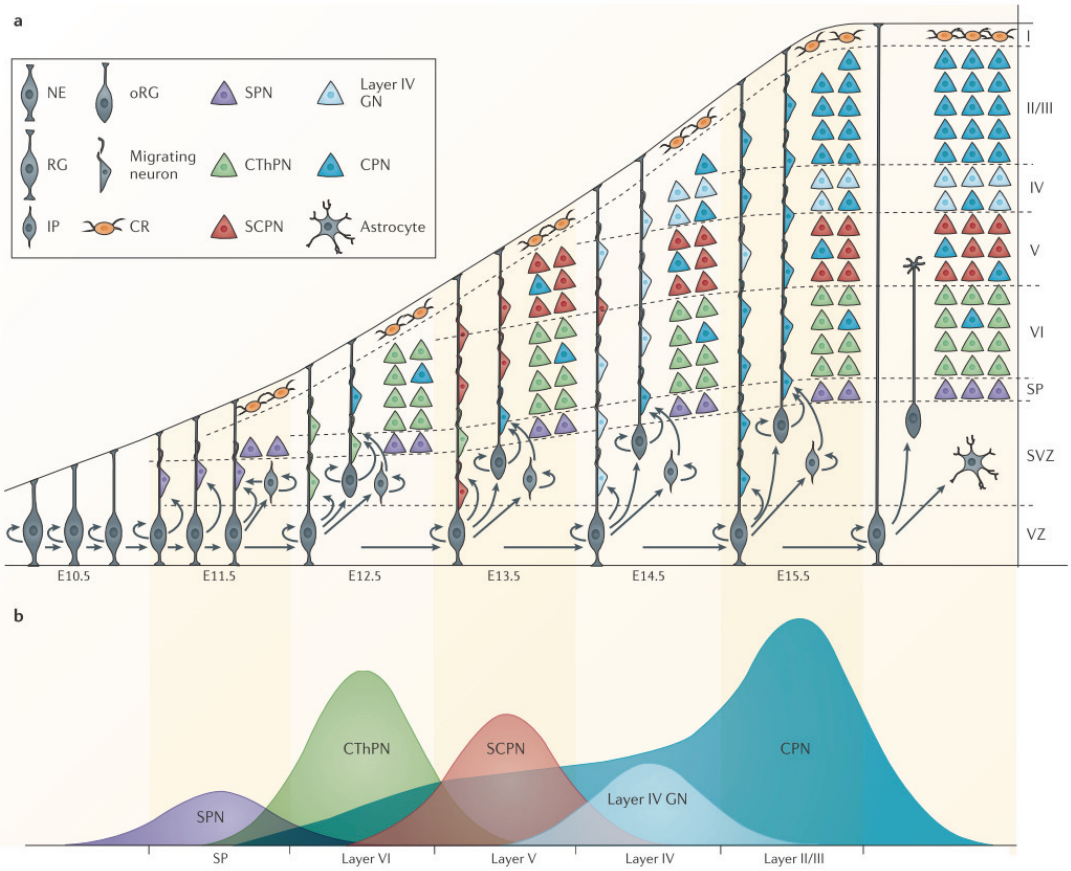


Figure 6: Neocortical projection neurons are generated in an 'inside-out' fashion by diverse progenitor types in the VZ and SVZ.
"This schematic depicts the sequential generation of neocortical projection neuron subtypes and their migration to appropriate layers over the course of mouse embryonic development. (A) Radial glia (RG) in the ventricular zone (VZ) begin to produce projection neurons around embryonic day 11.5 (E11.5). At the same time, RG generate intermediate progenitors (IPs) and outer RG (oRG), which establish the subventricular zone (SVZ) and act as transit-amplifying cells to increase neuronal production. Cajal–Retzius (CR) cells primarily migrate into neocortical layer I from non-cortical locations, whereas other projection neurons are born in the neocortical VZ and/or SVZ and migrate along radial glial processes to reach their final laminar destinations. (B) Distinct projection neuron subtypes are born in sequential waves over the course of neurogenesis. The peak birth of subplate neurons (SPN) occurs around E11.5, with the peak birth of corticothalamic projection neurons (CThPN) and subcerebral projection neurons (SCPN) occurring at E12.5 and E13.5, respectively. Layer IV granular neurons (GN) are born around E14.5. Some callosal projection neurons (CPN) are born starting at E12.5, and those CPN born concurrently with CThPN and SCPN also migrate to deep layers. Most CPN are born between E14.5 and E16.5, and these late-born CPN migrate to superficial cortical layers. Peak sizes are proportional to the approximate number of neurons of each subtype born on each day. NE, neuroepithelial cell." Figure and legend taken from [133].

It is remarkable that at each developmental time point, the neural progenitor population is generating a different subtype of cortical projection neurons [147]. Interestingly, each neuronal layer can be identified by the expression of specific combinations of marker genes [133]. Despite tremendous progress, a full understanding of the molecular dynamics and regulatory processes underlying these cell fate commitment processes still remains elusive.

1.2.3.2 Cell fate determination of neural progenitors

As described above, RG cells in the VZ can give rise to neurons either directly via asymmetric neurogenic cell division or indirectly by generating additional progenitor populations which are committed further towards the neuronal lineage [119, 127]. These RG cells as well as their ancestral NECs belong to the class of apical progenitors (APs) with regard to their location of mitosis and both express the paired box 6 (Pax6) protein as a characteristic marker [138, 148, 149]. As these cells are able to self-renew and give rise to the majority of neocortical cell types including neurons and glia cells, they could be seen as neural stem cells (multipotent). I will refer to this cells as neural progenitors in this thesis. The self-renewing capacity maintains and expands the progenitor pool before these cells give rise to more developmentally-restricted cell types. However, it has to be taken into account that NEC and RG cell populations seem heterogeneous in nature with various levels of cell fate restrictions [127, 150]. Some RG cells retain their undifferentiated state during the course of embryonic development in order to contribute to the population of adult neural stem cells [151, 152]. The additional class of embryonic progenitors, which are found in the SVZ as the second germinal layer dorsal of the VZ, is summarized as basal progenitors (BPs). These BPs exhibit less developmental potential in comparison to apical RG cells. BPs include the aforementioned IPs, which express T-box brain protein 2 (Tbr2) as a characteristic marker [148]. Additionally, these progenitors display a morphology with retracted apical and basal processes in comparison to RG cells. BPs have only a very limited capacity to proliferate before they differentiate into neurons [127, 153, 154]. These cells will be referred to as neuronal progenitors in this thesis as they exclusively contribute to the neuronal population during embryogenesis.
The developmental control of cell fate changes from proliferating neural progenitors towards post-mitotic neurons is a complex process, coordinated by multiple cell-extrinsic and -intrinsic mechanisms in a spatio-temporal manner (Figure 7, see also 1.5.2) [119, 127, 134, 155].

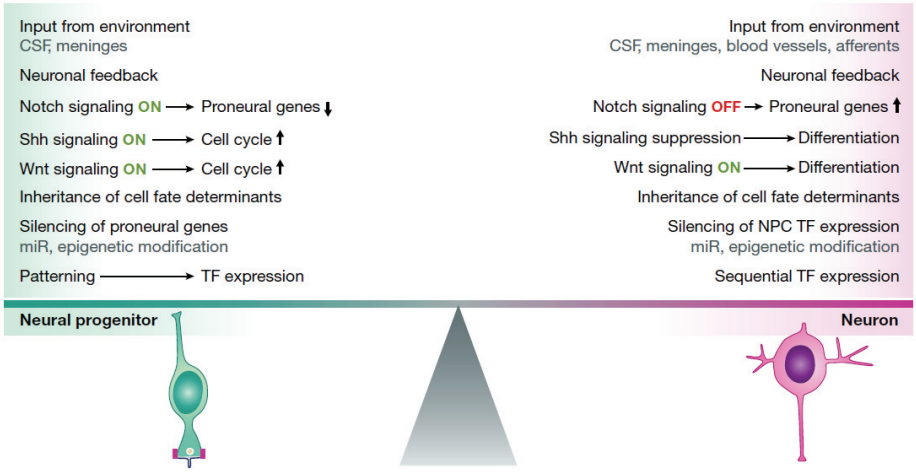


Figure 7: Cell-intrinsic and cell-extrinsic factors affect the balance between neural progenitor proliferation and differentiation.
Figure taken from [134].

Especially the retonic acid (RA) and the Notch-Hes signaling pathways have been described to play an essential role during this processes [156-160]. Notch is a cellular receptor located in the cell membrane of RG cells. After this receptor becomes activated by its Deltalike (Dll) ligands, which are expressed from newborn neurons, a cellular cascade is induced in activated RG cells leading to the inhibition of their neuronal differentiation potential, whereas their self-renewal process is promoted (Figure 8).

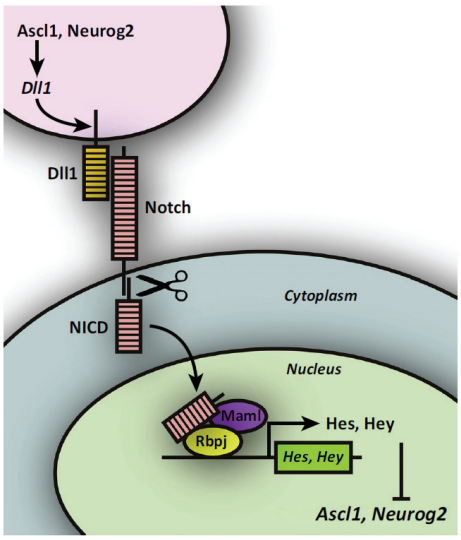


Figure 8: Notch signaling pathway.
"Notch signaling is activated through cell–cell interaction, between ligand (Dll1)-expressing and receptor (Notch)-expressing cells. Proneural factors (Ascl1, Neurog2) induce the expression of Notch ligand, which activate Notch signaling in neighboring cells. Upon activation of Notch signaling, Notch protein is subjected to proteolysis, releasing its intracellular domain (NICD, Notch intracellular domain), which translocates to the nucleus. Nuclear NICD forms a complex with the DNA-binding protein Rbpj and the coactivator Maml and induces the expression of downstream genes such as Hes and Hey. Hes and Hey repress the expression of target genes including proneural genes such as Ascl1 and Neurog2." Figure and legend taken from [161].

The balance between proliferation and differentiation of neural stem and progenitor cells is essential for embryonic brain development in order to fine-tune the overall amount of cells as well as cell type diversity in the neocortex [119, 127, 136]. It is interesting to note, that differences in the size and morphological complexity of the dorsal telencephalon across species can be attributed to changes in the cortical progenitor pool [162, 163]. For example, an additional outer SVZ harboring an expanded pool of BPs and thereby increasing the neuronal output is present in the brains of e.g. primates, which evolved a folded neocortex (gyrencephalic) in contrast to most rodents with their smooth, non-folded (lissencephalic) cortex [162, 163]. The increasing body of research strengthens the understanding that neuronal cell fate determination is driven by changes in chromatin environment and gene expression programs of progenitors during embryogenesis (see 1.5.2). However, the identification of unknown elements is crucial towards completing a comprehensive picture of such fascinating processes.

1.2.4 Cell migration and epithelial-to-mesenchymal transition

Embryonic development comprehends that newborn cells have to change their physical location in order to generate the growing embryo and its organs. In order to migrate to a distinct location in the developing organisms, a cell has to disconnect its cell-to-cell contacts in the local surrounding. The process in which a polarized epithelial cell adopts a mesenchymal, migratory phenotype is called epithelial-to-mesenchymal transition (EMT, Figure 9) [164, 165]. This procedure takes place when for example epithelial cells of the neuroectoderm give rise to neural crest cells during embryonic development [166].

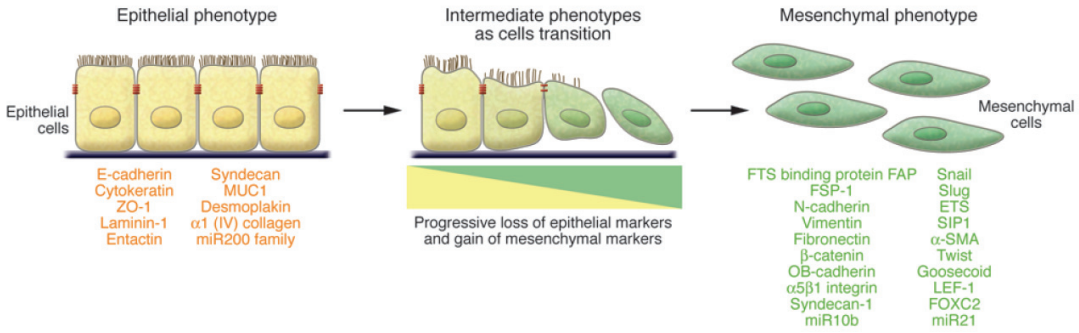


Figure 9: Scheme depicting the principals of epithelial-to-mesenchymal transition (EMT).
"An EMT involves a functional transition of polarized epithelial cells into mobile and ECM component–secreting mesenchymal cells. The epithelial and mesenchymal cell markers commonly used by EMT researchers are listed. Colocalization of these two sets of distinct markers defines an intermediate phenotype of EMT, indicating cells that have passed only partly through an EMT. Detection of cells expressing both sets of markers makes it impossible to identify all mesenchymal cells that originate from the epithelia via EMT, as many mesenchymal cells likely shed all epithelial markers once a transition is completed. For this reason, most studies in mice use irreversible epithelial cell–lineage tagging to address the full range of EMT-induced changes. ZO-1, zona occludens 1; MUC1, mucin 1, cell surface associated; miR200, microRNA 200; SIP1, survival of motor neuron protein interacting protein 1; FOXC2, forkhead box C2." Figure and legend taken from [165].

The phenotypic remodeling underlying EMT encompasses several biochemical and gene regulatory chances, including an increase in the production of extra cellular matrix (ECM) components, resistance to programmed cell death (apoptosis) as well as remodeling of the cytoskeleton [165]. However, the occurrence of EMT is not restricted to embryonic development, it is also linked to cellular processes in tissue repair and highjacked in pathological conditions like cancer metastasis or fibrosis [165, 167-169]. Major regulators of EMT are signaling pathways, especially the transforming growth factor β (TGF-β) pathway. This cytokine induces EMT via two signaling cascades, a MAPK-dependent (non-canonical) and a Smad2/3-dependent (canonical) pathway [165, 170]. Particularly the TGF-β mediated activation of Snail by Smad factors is a key regulatory event because these factors act as transcriptional repressors for epithelial cell markers and inhibit cell proliferation [171, 172]. During the embryonic development of mammalian telencephalon newborn neurons migrate from the germinal zones towards their cortical destiny as described above. This process has also been described as an EMT-like mechanism [173]. In addition to the involvement of Notch signaling in the regulation of developmental EMT via the zinc-finger containing Snail factors [174-176], this signaling pathway has also been attributed to be involved in the regulation of neuronal migration [177-181]. However, a full picture of the short- and long-range guiding cues underlying neuronal migration in the developing neocortex is still elusive [128, 182-185].

1.3 Transcription factors and the regulation of gene expression

Orchestrated cell fate changes are driven by differential gene expression programs. Major mediators of such cellular remodeling are nuclear proteins which bind to specific gene regulatory sites in the genome. These DNA-binding proteins are called transcription factors (TFs) and they influence the expression levels of associated genes as *trans*-acting factors. A substantial fraction of the mammalian proteome is attributed to

TFs as approximately 1,500 TFs are encoded in mammalian genomes, although the lineage-specific repertoire is quite diverse across the tree of life [186-189]. TFs are generally divided into two categories: (1) ubiquitously expressed general TFs that bind to gene promoters and are essential for basal transcription of the associated genes as they are part of the preinitiation complex [190]. In contrast, (2) cell type/sequence-specific TFs which target proximal as well as distal ($\geq$ 1 kb from TSS) regulatory elements like enhancers and are responsible for spatio-temporally controlled transcription profiles at a given time during ontogeny [191-193]. Sequence-specific TFs possess a characteristic domain for the interaction with DNA, the DNA-binding domain (DBD). The sequence of DNA bases which represents the core recognition locus for these TFs is called DNA binding motif and contains typically around 6 - 12 nucleotides in eukaryotes. However, DNA binding motifs are only known for the minority of TFs till date [194]. TFs can in general either promote (activator) or repress transcription (repressor), a function which is often mediated by additional domains like *trans*-activating domains within the same protein. Interestingly, eukaryotic TFs act in a combinatorial manner to ensure the functional transcriptional output of a cell in a particular environment at a given developmental time window and are a major subject of organismal evolution [192, 194-199]. Importantly, not all binding sites in the genome are bound and not all bound motifs are functionally relevant, suggesting that additional mechanisms ensure a fine-tuned functional occupancy of TF-binding sites, including their clustering, to ensure a proper cellular phenotype [192, 196, 200-206]. Additionally, it has been shown in vertebrates that the location of TF binding at regulatory elements exhibits species-specific differences [207, 208]. TFs are not acting in isolation as their function is influenced by the local chromatin environment as well as the presence of co-factors in the nucleus.

1.3.1 Classification of transcription factors based on their DNA-binding domain

Transcription factors interact with genomic DNA via their DBDs and the sequence similarity of these domains can be used to classify them into different families. Although more than 90 different DBDs are known, the majority of metazoan TFs belong to the families of zinc-finger proteins (ZFPs), basic helix-loop-helix (bHLH) and homeodomain (HD) TFs [194]. Of particular interest in this thesis are the two TF families described below.

1.3.1.1 Basic helix-loop-helix transcription factors

The family of bHLH TFs is an ancient class of proteins and plays a major role in lineage specification and cell differentiation [194, 209]. This family contains members like myogenic differentiation factor (MyoD), neurogenic differentiation factor (NeuroD), oligodendrocyte transcription factor (Olig), inhibitor of DNA binding (Id), achaete-scute complex homolog (Ascl), hairy and enhancer of split (Hes) as well as neurogenins (Ngn). They are characterised by an around 60 amino acids long N-terminal bHLH domain which consists of a basic region followed by an α-helix, a variable loop and a second α-helix [210]. Whereas the basic region is responsible for TF-DNA interaction, the HLH section functions as a dimerization domain [194, 211]. Since bHLH factors act as dimers, their functionality can be regulated by differential expression or post-translational modification of their dimerization partners in a given cell [211]. The functional bHLH

homo- or heterodimers typically bind to hexanucleotide DNA motifs like E-box elements (CANNTG) [212]. Interestingly, several additional dimerization domains have been described in the C-terminal region of bHLH factors [194, 211, 213, 214]. The functional diversity of these TFs is also evident from literature describing bHLH factors to function in cellular processes like cell cycle regulation, signal sensing as well as the regulation of developmental processes, including neurogenesis and gliogenesis [161, 209, 215-219]. Important neural bHLH factors like Ngn2 and Ascl1 function as activators in a hetero-dimer with the bHLH factor E47, whereas repressors like Hes1 and Hes5 act as homo-dimers [217]. Interestingly, another neuronal bHLH factor, NeuroD1, has been shown to form a hetero-dimer with E47 while occupying its specific hexanucleotide E-box motif CATCTG (Figure 10) [220, 221].

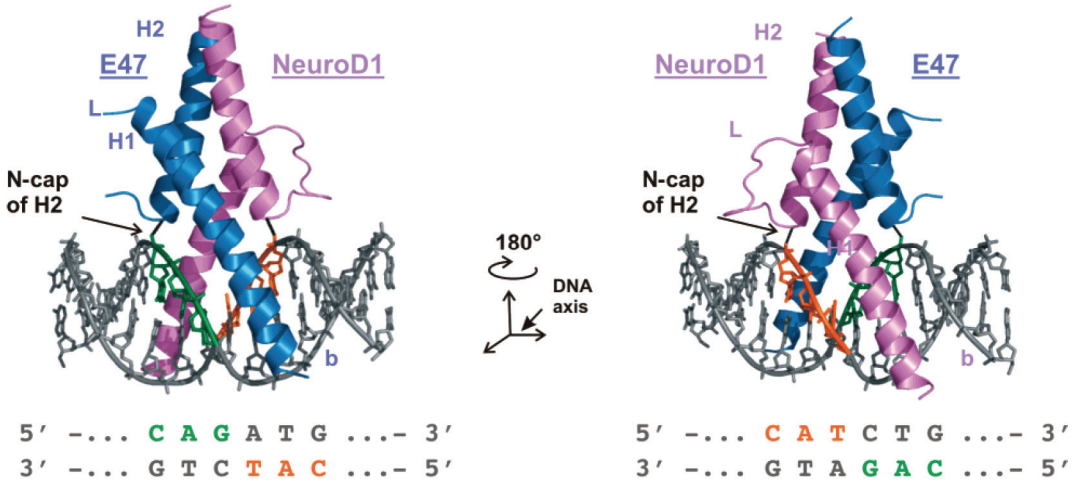


Figure 10: Overview of the E47 - NeuroD1 heterodimer structure.
"Ribbon drawing of the E47 (blue) and NeuroD1 (violet) bHLH domains bound to DNA. On the left the structure is rotated 180 ° around the DNA 2-fold axis with respect to the right structure. The basic (b), helix 1 (H1), loop (L), and helix 2 (H2) regions are labeled for E47 on the left and for NeuroD1 on the right." Figure and legend taken from [220].

1.3.1.2 Zinc-finger proteins

Zinc-finger proteins (ZFPs) comprise the largest family of eukaryotic TFs with several thousand members found in Metazoa. They are assumed to function generally as repressors in a plethora of cellular processes [222-224]. Their DBDs comprise so called zinc-fingers whose structure resembles a finger-like shape which is stabilized by a coordinated zinc ion (Zn^{2+}) between specific cysteine or histidine residues [225, 226]. Although there are sequence variations in this around 30 amino acid long zinc-dependent structure, the most prevalent form comprises a cysteine (C) residue in two β-sheet regions as well as two histidine (H) residues in a single α-helix (C2H2), all of which are highly conserved [194, 222, 223, 226, 227]. The modular arrangement zinc-fingers separated by a linker sequence leads to a diversity that enables C2H2 ZFPs to recognize more DNA motifs than all other TFs combined [228, 229]. Each zinc-finger of this poly-ZFPs can recognize mainly three specific DNA bases with its α-helix, but for the majority of ZFPs no sequence motif is known [228]. Interestingly, zinc-fingers have also been described to bind other ligands like RNA or

proteins, but till date no guidelines are known to precisely predict which ligand or which specific DNA motif will be bound by a particular C2H2 domain in the cellular milieu [229-234].

1.3.1.2.1 Krüppel-associated box containing zinc-finger proteins

Several additional protein domains evolved among polydactyl C2H2 ZFPs. The most abundant N-terminal effector domain is the around 75 amino acids containing Krüppel-associated box (KRAB). The combination of zinc-fingers and KRAB domain exist exclusively in several hundred C2H2 ZFPs in tetrapods, comprising for example up to 25% of all TF genes in *Homo sapiens* [188, 235-237]. The KRAB domain has been shown to interact with the transcriptional co-repressor KRAB-associated protein 1 (KAP1; also known as tripartite motif containing 28 (TRIM28)) which in turn recruits an epigenetic silencing-machinery. This machinery includes the H3K9me3 histone-methyltransferase SET Domain Bifurcated 1 (SETDB1, also known as ESET) in order to facilitate heterochromatin formation at target genomic regions (Figure 11) [235, 238-243].

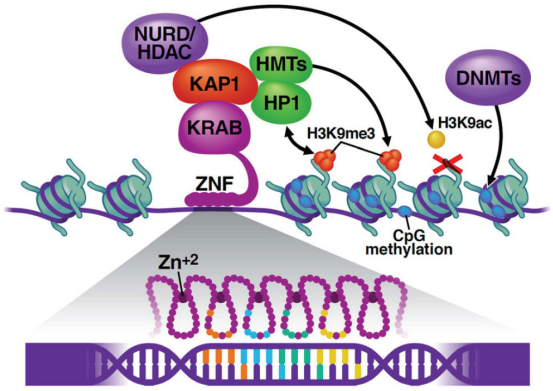


Figure 11: Model of KRAB-ZFP binding to DNA and induction of heterochromatin formation.
"Protein–DNA interaction between ZFPs (here named ZNF) and DNA are mainly mediated by four amino acids at positions −1, 2, 3, and 6 of the α-helix (colored circles). KAP1 is recruited through the KRAB domain and interacts with the NURD/HDAC repressor complex and histone methyltransferases (HMTs) (e.g. SETDB1), which catalyze the removal of H3K9ac and the addition of H3K9me3, respectively. HP1γ interacts with both KAP1 and H3K9me3. DNA methyltransferases (DNMTs) methylate genomic CpG sites, leading to inheritable silencing." Figure and legend taken from [244].

Some studies demonstrated that KRAB-ZFPs target repetitive genomic elements like endogenous retroviral elements (ERVs) to facilitate their transcriptional silencing by the KAP1/SETDB1 machinery, a mechanism which seems to have emerged through an evolutionary arms race [229, 235, 236, 244-249]. This relationship has also been functionally shown as a de-repression of several ERV families was observed in KAP1 or SETDB1 knockout ESCs [250-252]. The rapidly-evolving KRAB-ZFPs are proposed to obtain a specific expression pattern across cell types and tissues in the course of ontogeny as well as speciation [253-258]. However, a deeper functional understanding of the relationship between KRAB-ZFPs and ERV silencing, especially during cell fate determination in several cellular lineages is still vague.

1.3.2 Transcription factor binding sites

The murine genome comprises around 2.7 x 10^9 base pairs (bp) [259-262]. A recent study annotated that around 12% of total DNA might have a potential regulatory function in order to ensure a proper spatio-temporal configuration of gene expression, whereas only around 2% of the genome are annotated as promoters [263]. Cis-regulatory elements such as enhancers and promoters are enriched for TF motifs and are located on the same DNA molecule as the gene they regulate [264, 265]. The activity of promoters as well as enhancers determines the transcriptional profile of any given cell throughout development and orchestrates cell diversity in a spatio-temporal manner. These elements share several properties and function in a hierarchical manner, although they generate different but functional RNA molecules: Whereas the production of spliced, polyadenylated mRNA is regulated by most promoters, the RNAs transcribed from enhancers (eRNAs) are less stable and shorter than mRNAs [266-271]. Interestingly, the repertoire of proteins and TF motifs at *cis*-regulatory elements of housekeeping genes differs from the configuration found at developmental genes [272, 273]. Additionally, a difference in the transcriptional output of orthologous genes in the murine and human genome might be attributed to location changes of TF motifs in their regulatory elements (Figure 12) [263, 265, 274-280].

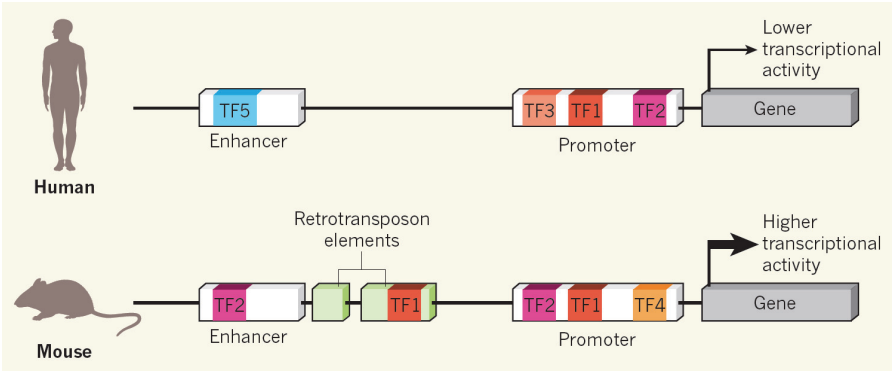


Figure 12: TF binding and gene regulation in mice and man.
Gene expression rates are regulated by TFs, which bind to promoter regions close to the specific gene or to distal enhancers, which might harbor different TF motifs at orthologous genes across species. An additional layer constitutes the presence of retrotransposon elements which contain TF motifs and thereby might influence gene expression. Figure and legend adapted from [277].

Active promoters and enhancers share several chromatin marks and they are mostly depleted of nucleosomes, making them accessible for TF binding [280-282]. Strikingly, novel lines of research propose that these *cis*-regulatory elements might be functionally interchangeable [267, 271, 283-285]. In addition, the traditional view of gene regulation where only (distal) enhancers increase the transcriptional output of gene promoters is progressively challenged by a shift in the understanding of transposable elements: As these frequent genomic elements have been shown to exhibit regulatory characteristics, they emerge as a functional component in the nucleus [286, 287]. The genomic fraction previously seen as “junk” DNA becomes increasingly noticed as a regulatory part of the genome [288, 289].

1.3.2.1 Promoters as *cis*-regulatory elements

A promoter refers to a DNA sequence spanning around 1000 bp from the transcription start site (TSS) of a gene which is responsible for its transcriptional initiation and the adjustment of gene expression levels [290, 291]. This regulatory element is often CpG rich and generally comprises a small core and a larger proximal promoter region, although the sequence composition in Metazoa is quite diverse [281, 291, 292]. The core promoter includes the TSS +/- 40 bp and facilitates the correct positioning of the preinitiation complex (PIC) including general TFs and a DNA-dependent RNA-polymerase [190, 290]. The classical consensus RNA-Polymerase II (Pol II) core promoter contains a TATA-box (consensus sequence TATAAA) around 30 bp upstream of the TSS which is recognized by the TATA-box binding protein (TBP) of the general TF TFIID [290, 293]. However, less than 15% of mammalian promoters contain a TATA-box whereas other core promoter elements like the initiator sequence (Inr) are more common [294-296]. In addition, several different types of Pol II promoters have been described, increasing the complexity of gene regulatory mechanisms [297, 298]. Overall, the binding of TFs to proximal promoter elements as well as to distal regulatory sites regulates the transcriptional rate initiated by the core promoter [272, 290, 299]. The variation of sequence motifs in the core promoter is directly responsible for the specificity of its interaction with certain enhancers to integrate the complete regulatory input from TFs [272, 273, 300].

1.3.2.2 Enhancers as *cis*-regulatory elements

The ubiquitous expression of housekeeping genes can be adequately controlled by their promoter elements alone [291, 293] (Lenhard et al., 2012; Haberle & Lenhard, 2016). In contrast, the highly specific, spatio-temporally restricted expression pattern of several developmental genes throughout ontogeny is mediated by enhancers [293, 301-303]. Enhancers are mostly CpG poor *cis*-regulatory elements, were first discovered in the early 1980s, comprise around 200 - 1500 bp and are more prevalent in the genome than promoters [263, 271, 304-306]. Kim and Shiekhattar summarise several key characteristics of enhancers as follows: “(1) Enhancers increase transcription of a linked gene from its correct initiation site specified by the core promoter, (2) enhancer activity is independent of orientation relative to its target gene, (3) enhancers can function independent of their position relative to the target genes, and also over long distances, (4) enhancers can function with a heterologous promoter, (5) enhancers exhibit DNase I hypersensitivity (HS), which reflects a less compacted chromatin state as a result of the binding of various transcription factors” [267]. Especially their position-independent and long-range mode of operation is in contrast to promoters and emphasises that over the course of ontogeny a given gene can be regulated by several specifically active enhancers which might be located in intergenic regions as well as in introns or exons. Interestingly, several studies strengthened these observations by positioning enhancers in ectopic genomic locations while still preserving their physiological regulatory function [307-309]. As enhancers are activated by TF binding and harbor clusters of TF motifs including motifs of signaling pathway effectors, they play an important role in the coordination of multiple cell-extrinsic and -intrinsic regulatory mechanism to integrate the context-specific regulatory input at the core promoter [293, 310-312]. The mechanism trough which

developmental enhancers function, especially in lineage commitment and cell type specification, is complex and not fully unraveled. However, it has been shown that they can mediate the formation of chromatin loops between intrachromosomal regulatory regions, facilitate an increase of PIC as well as histone modifying enzymes at target promoters and regulate transcriptional elongation [271, 313-322]. The formation of loop structures is not only facilitated by TFs alone but several assisting structural players like CCCTC-binding factor (CTCF), Mediator or Cohesin have been identified to bring an enhancer in close physical proximity to its associated promoter [316, 323-327]. However, the dynamics of enhancer-promoter looping are still under debate [328, 329]. A common experimental procedure to associate an enhancer to its potential target promoter uses a "nearest gene" approach, although the functionality of this prediction has to be experimentally validated, as the complexity of the enhancer regulatory networks does usually not reflect this 1:1 relationship [330-333].

1.3.2.3 The non-coding genome apart from promoters and enhancers

The non-coding genome classically refers to the DNA portion which does not contain protein-coding sequences and is largely comprised of intergenic DNA [334, 335]. Interestingly, the variation in total (haploid) genome size (C value) across metazoan genomes is attributed to a variation of the non-coding, potentially regulatory genomic fraction, whereas the amount of genes coding for proteins is mostly comparable [288, 336-340]. Furthermore, the emergence of the metazoan non-coding regulatory genome was an essential step for the evolution of multicellularity [289]. Increasing research efforts show that these genome size differences in Metazoa can partially be explained by the expansion of repetitive DNA elements [341-344]. The classical constitution of the murine non-coding genome contains several classes of repetitive DNA elements including simple repeats, satellite DNA as well as transposable elements (transposons) such as long interspersed nuclear elements (LINE), short interspersed nuclear elements (SINE) or long terminal repeat (LTR) retrotransposons (Figure 13).

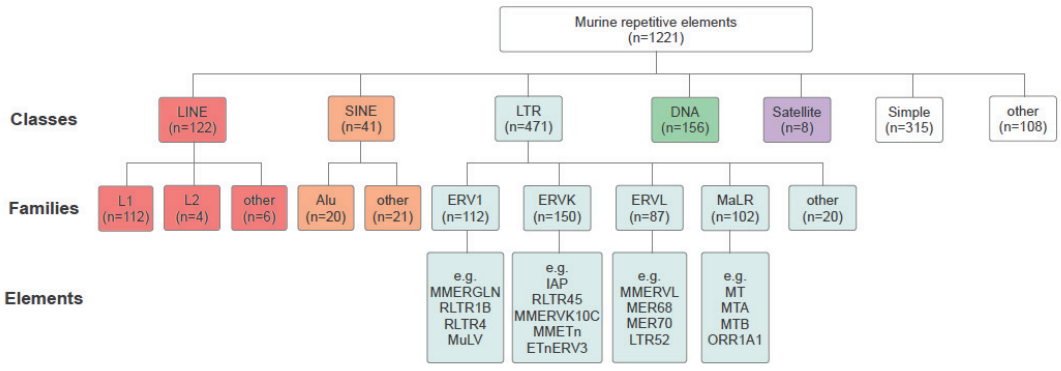


Figure 13: Selected overview of murine repetitive elements.
The murine genome contains 1221 different consensus sequences for repetitive elements which are categorized by Repeatmasker into 16 classes comprising 45 families. A subset of these element classes and families is highlighted with different colors. Figure and legend adapted from [345].

Transposons are genetic elements with an ability to “jump” (transpose) into a genomic location of the host and cover around 40% of the murine genome [259, 346]. Whereas DNA transposons use a “*cut-and-paste*” mechanism, retrotransposons like LINE, SINE and LTR elements transpose via RNA intermediates (“*copy-and-paste*”) by employing reverse transcription [347, 348]. Transposable elements have co-evolved intrinsic *cis*-regulatory sequences for their open reading frames important for viral replication in order to highjack the cellular machinery of their host for gene transcription, although not all elements in the genome are active anymore [347, 349-352]. The 5´-untranslated region (5´-UTR) of LINE retrotransposon encodes for example an internal Pol II promoter and a polyadenylation signal is also present in these elements [349, 353]. Furthermore, the 200 - 600 nt containing long terminal repeat of LTR retrotransposons, which cover around 10% of the murine genome, reflects the characteristics of a *cis*-regulatory element [247, 354-356]. The middle regulatory region (R) of an LTR resembles a Pol II core promoter with the TSS and a TATA-box whereas the flanking unique (U) regions harbor TF motifs and can also contain a splice donor (SD) site at the 5’ region [356]. It has been estimated that these retroviral TF motifs have contributed on average to approximately 20% of the functional TF binding sites in higher mammals, displaying a tendency of species-specific variation [356-358]. Prominent human examples are found in several TF motifs of core pluripotency genes SOX2, OCT4 and NANOG [359, 360]. Furthermore, several retrotransposons have been co-opted by their host genome as *cis*-regulatory elements like enhancers or promoters: Among various examples, a LTR derived from Intracisternal A-Particle (IAP) elements has been shown to regulate the murine agouti allele as a promoter whereas other LTRs function as species-specific enhancers during murine placenta development [287, 349, 355, 361-363].

LTR elements generally originate from exogenous retroviruses which infected germ cells. LTR transposons that lost their ability of extracellular mobility in the course of evolution are referred to as endogenous retroviral elements (ERVs) [355, 364-366]. Murine ERVs are diverse in their sequence composition and display structural differences (Figure 14).

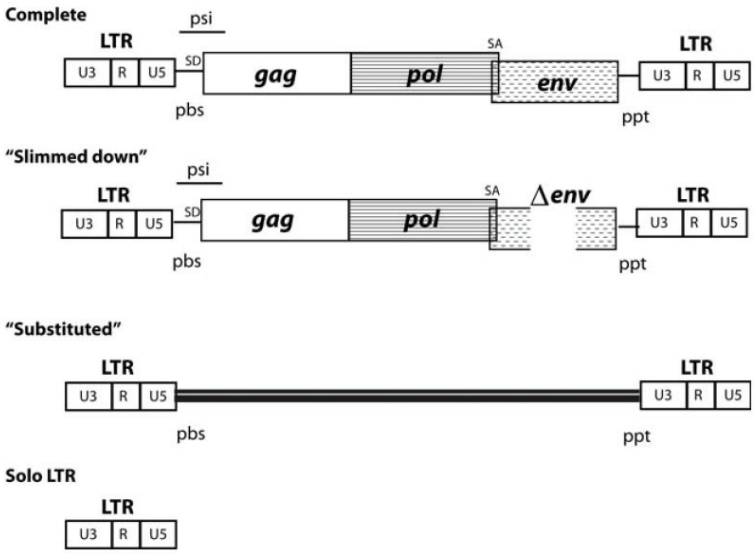


Figure 14: Overview of LTR and basic ERV structures.
"Complete ERVs are essentially identical to the integrated proviruses of simple exogenous retroviruses; they contain two LTRs made up of unique 3′ (U3), repeat (R), and unique 5′ (U5) regions, a primer binding site (pbs) and polypurine tract (ppt), as well as a full complement of coding sequences (group specific antigen, gag; reverse transcriptase (polymerase), pol; envelope, env), splice donor (SD) and acceptor (SA) sites, and an RNA packaging signal (psi). "Slimmed down" ERVs are elements lacking coding sequences compared to a complete ERV-here illustrated with a deletion in env. "Substituted" ERVs are elements in which the ERV coding sequences have been replaced with nonviral sequences. "Solo LTRs" are single LTRs generated by homologous recombination between the two LTRs of a complete element." Figure and legend taken from [354].

Several ERV elements like IAP (~ 1000 copies per cell) or *Mus musculus* virus-like 30 (MMVL30, ~ 150 copies per cell) are transcribed during ontogeny in a tissue-specific matter [355, 356]. On the other hand, highly active ERVs like IAP have been attributed to the occurrence of spontaneous mutations in the murine genome [352, 356, 367]. Therefore, as retroviral elements represent a potential threat to the host by perturbing their genome through various mechanisms, several silencing strategies like DNA methylation and the above-mentioned KRAB-ZFP/KAP1/SETDB1 system have co-evolved to control their activity [247-249, 368, 369]. The contribution of transposable elements to nuclear architecture and gene regulation, especially in the context of health and disease, is increasingly appreciated and changes the perception of these genomic parasites in the non-coding genome [247, 349, 356].

1.3.3 A hierarchical model of gene expression by RNA-Polymerase II

Around the middle of the last century, Francis Crick postulated the central dogma of molecular biology by describing that the genetic information flows between three biological polymers: from DNA to RNA into proteins and from RNA into DNA by reverse transcription, but not from proteins into nucleic acids [370, 371]. The genetic information is encoded in genes, but it is evident that only a subset of genes encodes for proteins and genes should therefore rather be seen as a blueprint for biologically active (regulatory) RNAs [372, 373]. During the process of eukaryotic transcription, three different DNA-depended RNA polymerases (Pol) produce diverse RNA types from the DNA template by employing altered peripheral factors: Whereas

Pol I and Pol III produce mostly ribosomal RNAs (rRNA, I + III), small RNAs (I + III) and transfer RNAs (tRNA, III), Pol II transcribes genes into messenger RNA (mRNA) coding for proteins/peptides in addition to regulatory non-coding RNAs of various lengths [373-375]. A typical eukaryotic protein-coding gene consist of coding sequences ("expressed regions", exons), internal non-coding sequences ("intragenic regions", introns) and untranslated regulatory sequences (UTRs) flanking the 5' and 3' ends. Whereas UTRs are involved in the regulation of mRNA translation, localization and stability, introns are typically not contained in the mature mRNA because they are removed ("spliced") during the mRNA maturation process in contrast to exons [376-379]. Eukaryotic gene expression is a complex and highly dynamic process driven by a fine-tuned interplay of *trans*-acting factors and *cis*-regulatory elements in the local chromatin environment of an autonomous cell in order to facilitate its functional phenotype during ontogeny [9, 380-383]. The initiation of transcription at an accessible RNA Pol II promoter correlates with the step-wise assembly of the multi-subunit pre-initiation complex (PIC, > 85 polypeptides) on the core promoter where co-factors in concert with general TFs recruit Pol II to the correct position relative to the TSS and further prepare the DNA template for transcriptional elongation [190, 384]. As mentioned above, the cell- and context-specific activity of TFs, *cis*-regulatory elements and co-factors incorporates the environmental regulatory input at the promoter in a hierarchical fashion in order to regulate the transcriptional activity of the promoter by promoting Pol II activity (Figure 15).

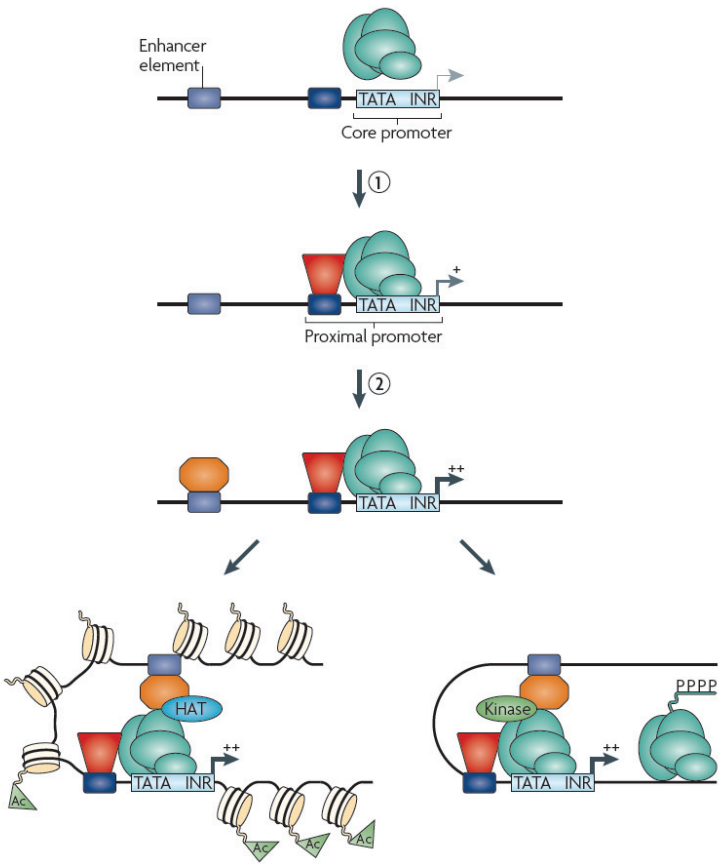


Figure 15: Model for transcriptional regulation at Pol II promoters.
General transcription factors and Pol II (PIC, green ovals) bind to an accessible core promoter containing TATA boxes and initiator (INR) sequences and facilitate a basal level of transcription. (1) The recruitment of sequence-specific TFs (red trapezoid) to the proximal promoter (dark blue box) stabilizes the PIC by direct protein-protein interaction between these TFs, which enhances the transcriptional activity of the promoter (+). The transcriptional output of a promoter can be further increased (++) by the cooperativity of sequence- and cell type-specific factors (orang and pink octagons, respectively) at accessibly distal enhancers: A site-specific TF (2) facilitates the recruitment of co-activators and consecutively cell type-specific TFs (3) as well as PIC. The cell type-specific enhansosomal complex can further recruit chromatin modifying enzymes (bottom left) or kinases that regulate the phosphorylation (P) status of Pol II at its C-terminal domain (CTD) to stimulate transcriptional elongation. Ac, acetylated histone; HAT, histone acetyltransferase. PIC, pre-initiation complex. Figure and legend adapted from [385].

The recruitment of Pol II to active enhancers [386-388] displays another regulatory layer in the light of several reports that showed an altered recruitment of general TFs and Pol II to promoters upon enhancer removal but not vice versa [316, 318, 321]. The notion that enhancers might contribute to the regulation of transcriptional output of a promoter by delivering (active) Pol II is further supported by the ability of Pol II to elongate and proof-read RNA transcripts on its own once its initiation is successfully completed [389]. Several variations have been proposed for this model of enhancer-promoter communication in the context of chromatin, which requires further functional investigation for a deeper understanding. The proposed scenario of gene regulation is furthermore revised by the increasing body of research on pioneer transcription factors (see 1.5.1), which are believed to initiate the formation of accessible chromatin at regulatory elements in order to pave the ground for their activation by the aforementioned mechanisms [390, 391]. It has also been proposed that transcription appears for many genes at once in close spatial vicinity inside the nucleus, a concept referred to as “transcription factories”, further increasing the regulatory complexity [392]. Overall, the concerted action of several regulatory layers ensures that a cell can respond adequately in a spatio-temporal manner to internal- and external cues during ontogeny [383, 393, 394].

1.4 Chromatin and Epigenetics

The term “epigenetics” was coined in 1942 by Conrad Hal Waddington for the (unknown) developmental mechanisms that cause alterations in phenotypes without modifications in the genotype [395, 396]. He later defined his “epigenetic landscape” [397] as a metaphor to illustrate the canalising epigenetic influences that determine the fate of a cell during its developmental journey towards a terminal differentiated state as an analogy to a ball rolling downhill across branching points by gravitational force (Figure 16A). However, the underlying molecular basis was not adequately unrevealed at that time, but continuing research has advanced the profile of Waddington’s landscape to a modern, more dynamic version (Figure 16B).

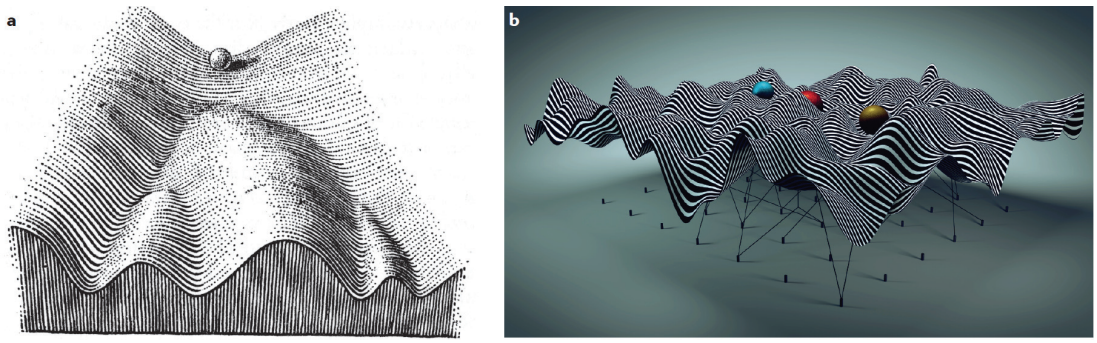


Figure 16: The evolution of Waddington’s epigenetic landscape.
(A) The classical epigenetic landscape as depicted by Waddington in 1957. (B) A modern version of Waddington’s epigenetic landscape model. The landscape and therefore in analogy cellular phenotypes (coloured balls) are dynamically changing due to the influence of epigenetic manipulations as represented by the ropes dragging the landscape’s silhouette. Figure and legend adapted from [398].

These increasing insights advanced Waddington's view of epigenetics into a modern adaption: "Almost three-quarters of a century later, we know that epigenetic mechanisms transduce the inheritance of gene expression patterns without altering the underlying DNA sequence but by adapting chromatin, which is the physiological form of our genetic information" [9]. However, some ambiguities are still discussed in the rapidly growing field of epigenetic research [399, 400]. Nonetheless, it is established that chromatin plays a prominent role in these processes and several mechanisms have been unraveled showing its regulatory impact as an epigenetic barrier for the flow of genomic information during cell fate determination [9, 192, 398, 401-404].

1.4.1 Chromatin organisation and modification

Chromatin is a macromolecular complex that consists of DNA and RNA in addition to (histone) proteins and is embedded within the eukaryotic nucleus. After their first description by Emil Heitz in 1928, two categories of chromatin were initially distinguished based on their degree of compaction and therefore visual appearance upon staining: Loose euchromatin and highly compacted heterochromatin [405]. Continuing research has improved this model with functional annotations, where euchromatin reflects the "light-staining, decondensed and transcriptionally accessible regions of the genome" [9], whereas heterochromatin describes the "dark-staining, condensed and gene-poor regions of the genome" [9].

The building blocks of chromatin are nucleosomes: a histone (H) octamer is formed by two copies of each histone protein (H2A, H2B, H3 and H4) and 147 bp of DNA that is wrapped around this canonical core particle [406-409]. Another histone protein, the linker histone H1, stabilises this "bead" structure by binding to the flanking DNA regions of the nucleosome [410]. Nucleosomes are interconnected by variable stretches of DNA ("linker DNA"), which on average comprise around 200 bp and lead to a "beads on a string" appearance of the 10 nm chromatin fibre [407, 411, 412]. This chromatin fibre can be compacted into higher order chromatin structures by condensation of nucleosomal arrays and the action of architectural proteins like linker histones [412]. Chromatin organisation ensures that the genome is compacted several thousand-fold in order to fit into the eukaryotic nucleus. Furthermore, the packaging of DNA into nucleosomes and higher order chromatin structures reduces the accessibility of DNA sequences for interaction with nuclear factors and therefore influences every DNA-based cellular process like transcription, DNA replication or DNA repair [408, 412-416]. The local chromatin structure of a gene with its underlying modifications at a developmental time point needs to be tightly regulated in order to control the correct transcriptional output. Nonetheless, reasearch dedicated to long-range chromatin interactions and the description of chromatin territories or topologically-associated domains (TADs) shows that attention should also be drawn to chromatin dynamics in the whole nuclear environment in order to fully understand the functional relevance of the different organisational levels (Figure 17) [9, 417-420]. Several regulatory chromatin layers relevant to this thesis are described in the following sections.

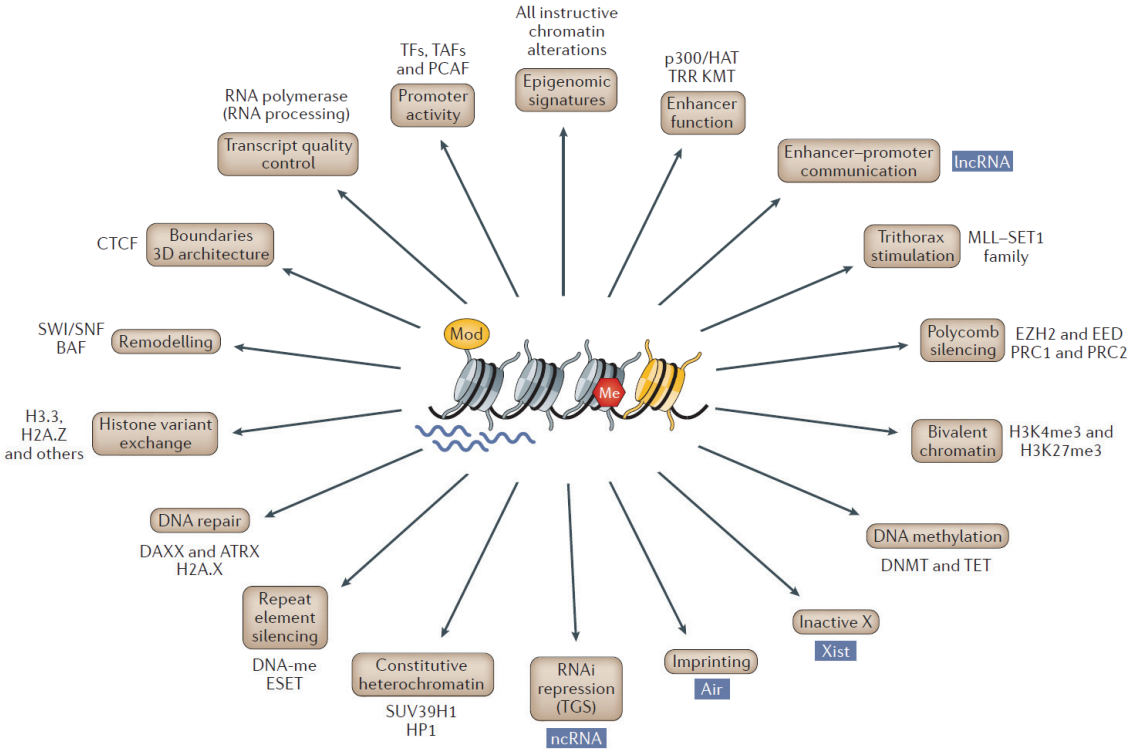


Figure 17: Schematic overview of several chromatin functions as an integrative part of the epigenome.
"A chromatin template with four nucleosomes is depicted in the middle of the figure, together with chief mechanisms, such as histone modifications (Mod), DNA methylation (Me), histone variants and remodelling (yellow nucleosome) and non-coding RNA (ncRNA; wavy blue lines), that alter chromatin structure and function in an inter-dependent fashion. Distinct adaptations of this chromatin template have been associated with various functions of the epigenome (boxed examples). Also shown are some of the major chromatin factors that regulate these chromatin transitions. Air, antisense insulin-like growth factor 2 receptor RNA; ATRX, α-thalassemia/mental retardation syndrome X-linked; BAF, BRG1-associated factor; DAXX, death-domain-associated protein; CTCF, CCCTC-binding factor; DNA-me, DNA methylation; DNMT, DNA (cytosine-5)-methyltransferase; EED, embryonic ectoderm development; ESET, ERG-associated protein with SET domain; EZH2, Enhancer of zeste homologue 2; H2A.X, histone H2 variant; H3K4me3, histone H3 lysine 4 trimethylation; HAT, histone acetyltransferase; HP1, heterochromatin protein 1; KMT, lysine methyltransferase; lncRNA, long non-coding RNA; MLL, mixed-lineage leukaemia; PCAF, p300/CBP-associated factor; PRC, Polycomb repressive complex; RNAi, RNA-mediated interference; SUV39H1, Su(var)3–9 homologue 1; TAFs, TATA-box binding protein associated factors; TET, ten-eleven translocation; TFs, transcription factors; TGS, transcriptional gene silencing; TRR, Trithorax related; Xist, X-inactive specific transcript." Figure and legend taken from [9].

1.4.1.1 DNA cytosine methylation

The first chemical modifications of DNA nucleotides were described in 1948 [421], but a functional description was lacking at that time. Research over the last decades showed that especially the methylation of cytosine at the fifth carbon atom in the pyrimidine ring (5-methylcytosine, 5mC) plays a role in gene regulation, especially gene repression [422-427]. Mammals express three DNA methyltransferase enzymes (DNMTs) which catalyse the formation of 5mC: DNMT3A and DNMT3B act as *de novo* DNA methyltransferases and are very important especially during early development in the preimplantation embryo and primordial germ cells, whereas DNMT1 plays a major role for the maintenance of DNA methylation patterns on the hemi-methylated DNA double helix generated during DNA replication [428-432]. In mammals, DNA methylation frequently appears at palindromic CpG dinucleotides, but promoters with

stretches of CpG-rich regions (CpG islands) tend to exhibit regulatory dynamics of their 5mC content in order to retain CpG dinucleotides throughout evolution [427, 431, 433, 434]. The methylation of a TSS is generally associated with long−term silencing of the corresponding gene whereas gene body methylation has been linked to transcriptional elongation [427, 435-438]. Interestingly, recent studies have identified a new class of differentially expressed enzymes, ten-eleven translocation 1-3 (TET1-3), which catalyse the oxidation of 5mC into 5hmC (5-hydroxymethylcytosine), 5fC (5-formylcytosine) and further 5caC (5 carboxylcytosine) [439-442]. Although the function and stability of these intermediates in the DNA demethylation/DNA methylation cycle (Figure 18) is not fully unravelled, these studies show that the epigenetic layer of DNA methylation is more complex than previously expected [443].

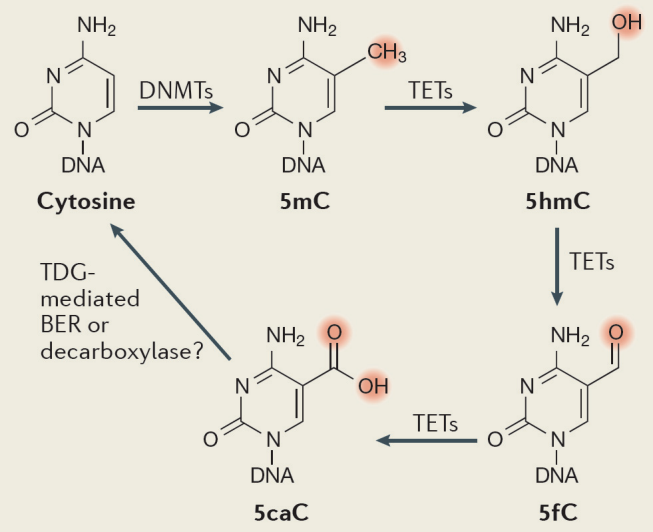


Figure 18: Schematic overview of the DNA-methylation and DNA-demethylation cycle. The concerted action of TET enzymes oxidises 5mC to 5caC via several intermediates. The action of Thymine-DNA glycosylase (TDG)-mediated base excision repair (BER) can potentially regenerate cytosines lacking methylation. 5mC, 5-methylcytosine; 5hmC, 5-hydroxymethylcytosine; 5fC, 5-formylcytosine; 5caC, 5 carboxylcytosine. Figure and legend adapted from [444].

Similar as for histone modifications, several nuclear factors have been identified that bind to 5mC and its oxidised intermediates, like the methyl-CpG-binding protein 2 (MeCP2; [445]) or methyl-CpG binding domain proteins (Mbds, [446]), and therefore function as “readers” of DNA methylation status. Furthermore, it has been proposed that DNA methylation patterns might influence the binding of certain TFs to regulatory regions by generating novel methylation-dependent binding sites [444, 447]. It is therefore not surprising, that DNA methylation has been shown to be involved in a plethora of biological processes including X-chromosome inactivation, genomic imprinting, silencing of genes as well as transposable elements and has been shown to be aberrant in diseases like cancer [432, 448-455].

1.4.1.2 Post-translational modifications of histones

The histone-DNA interactions in a nucleosome occur prevalently on the globular histone domains whereas the N-terminal histone domains ("tails") protrude from the nucleosomal surface [408, 456, 457]. These flexible tails are therefore accessible for nuclear proteins, can be post-translationally modified and interact with nucleosomes in close vicinity [412, 456]. Even though the covalent addition of acetyl-, methyl- and phosphoryl-groups on histone tails are the most studied post-translational modifications (PTMs) (Figure 19), a huge variety of PTMs have been described with an increasing focus on the globular histone domains as modification substrate [402, 458-460].

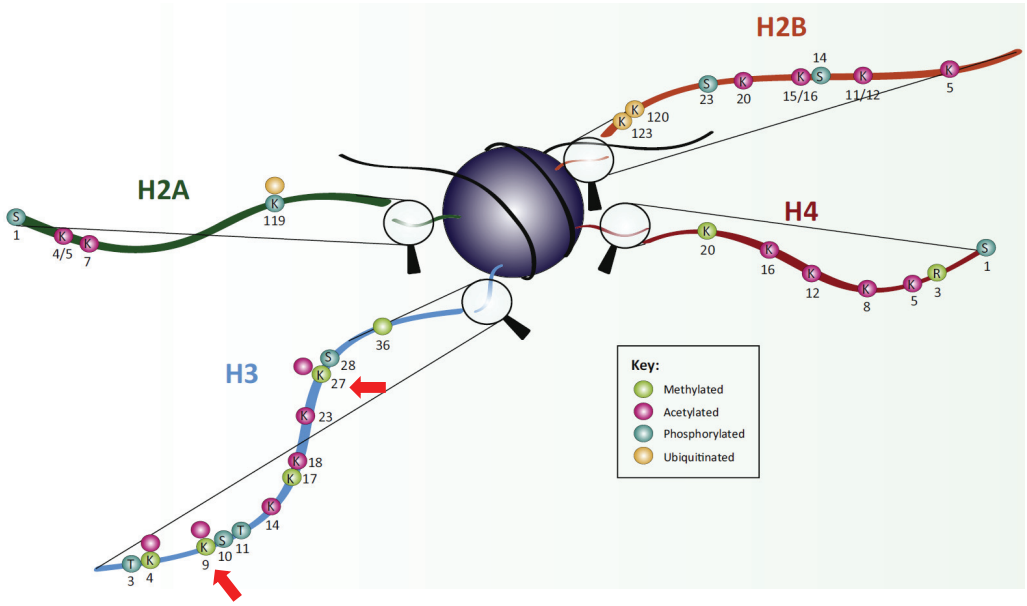


Figure 19: Schematic overview of histone tail post-translational modifications.
The location of the following amino acids at the N-terminal tails of each core histone are shown in the zoomed cutouts: K = lysine, S = serine, R = arginine and T = threonine. Coloured balls represent observed PTMs at a specific amino acid position as indicated by the numbers: green = methylated, pink = acetylated, turquoise = phosphorylated and beige = ubiquitinated. The nomenclature of histone PTMs follows a simple rule: the acetylation of the lysine residue at position 27 of histone H3 would be referred to as H3K27ac for example. Red arrows point to positions of specific interest in this thesis (H3K9, H3K27). For simplicity, only one of each histone type in the nucleosomal octamer is shown. PTM, post-translational modification. Figure and legend adapted from [402].

The figure above illustrates the sheer number of possible PTM combinations and it is well established that each modified nucleosomes carries a certain amount of epigenetic information stored as the so called "histone code" [461-463]. The same modification of the N-terminal tail could arise on both copies of each core histone, however, it has been proposed that histones can be asymmetrically modified [464, 465].
Certain modifications of these highly basic histone tails such as acetylation of positively-charged lysine residues can reduce their electrostatic interaction with the negatively-charged nucleosomal DNA backbone or chromatin-associated proteins by diminishing their affinity and thereby influencing chromatin structure directly [414, 466-468]. A plethora of nuclear factors is involved in writing, interpreting and deleting the

histone code: The so called “writers”, “readers” and “erasers”, which are categorized into several subgroups according to their catalysed PTM reaction or histone recognition domains (Figure 20).

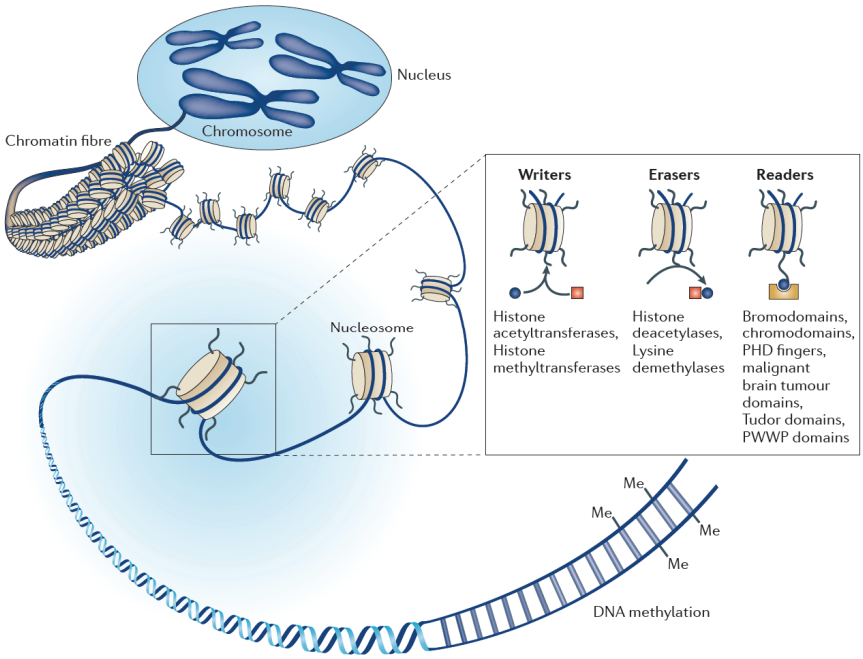


Figure 20: Protein families linked to covalent modifications of histones and their readout.
The concerted action of writers, readers and erasers of histone PTMs in combination with DNA methylation influences the functional status of chromatin and transcriptional landscape at a certain developmental time point. Figure and legend adapted from [469].

Enzymes that catalyse the covalent addition of histone modifications on chromatin are called “writers”, whereas the reverse reaction is catalysed by “erasers” (Figure 20), a finding that further illustrates the reversible nature of the epigenetic code. These histone-modifying enzymes can be recruited to defined DNA sites by other nuclear factors like TFs, non-coding RNAs or co-factors, promoting a locus-specific action of the epigenetic machinery. These recruiters could also be proteins that exhibit recognition domains for histone PTMs like Bromo-, Chromo- or PHD domains that enable them to bind specifically modified residues, which is why they are termed “readers”. The landscape of histone PTMs impacts several chromatin properties that facilitate tremendous functional assets such as the regulation of DNA access to chromatin-remodeling complexes, recruitment of nuclear factors like TFs or the storage of cellular signals [9, 470, 471]. The action of writers, readers and erasers is influenced by the local chromatin landscape (“crosstalk”), including the PTM status of the targeted nucleosome [471-475].

Interestingly, writers like histone acetyltransferases (HATs) and methyltransferases (HMTs) were found in active Pol II complexes [476, 477], providing a direct link between transcriptional regulation and histone PTMs. Histone methylation occurs mostly on two amino acid residues, lysine and arginine, but the extend and function differs between these two: lysine methyltransferases (KMTs) transfer one, two or three methylgroups, leading two a mono-, di- or try-methylated lysine (Kme1/Kme2/Kme3). A prominent example is enhancer of zeste (Ezh2 or Ezh1) as a subunit of the polycomb-repressive complex 2 (PCR2), which

catalyses the methylation of lysine 27 of histone H3 (H3K27me1/me2/me3), leading for example to the formation of facultative heterochromatin at repressed developmental genes with the characteristic marker H3K27me3 [478-481]. The family of H3K9 specific methyltransferases with members like SETDB1 and Su(var)3–9 homologue 1/2 (SUV39H1, SUV39H2) facilitates the concerted formation of constitutive heterochromatin in several cellular processes, often in combination with other heterochromatin proteins like HP1 [242, 482-490]. In contrast, the dimethylation of arginine residues by arginine methyltransferases occurs in either a symmetrical or asymmetrical fashion in addition to monomethylation [491-493]. Several histone demethylases have been identified, for example lysine-specific histone demethylase 1 (LSD1) or the diverse family of jumonji domain-containing proteins where several members like Jmjd3 erase H3K27me3 [494-497]. The acetylation of histone tails by HATs correlates in general with gene activation in euchromatic regions [9, 498-503]. A prominent example is the acetylation of lysine 27 on histone H3 (H3K27ac) by CBP/p300 (CREB-binding protein and p300) [504-507]. Interestingly, it has been shown that a subunit of the general transcription factor TFIID exhibits HAT activity [508]. In the same year when the first nuclear HAT was described [499], also the first histone deacetylase (HDAC) was discovered [509] and further research established the role of histone acetylation/deacetylation in gene activation and repression, respectively [9, 502, 507, 510].

Apart from the described post-translational modifications on canonical histones a further layer of regulatory complexity is added by the incorporation of so called histone variants in nucleosomal octamers that in turn can also harbour PTMs [511, 512]. Histone variants are highly conserved proteins, differ at several amino acid positions compared to their canonical counterparts, but in contrast to them these variants are incorporated into nucleosomes independent of replication [513-517]. Histone variants exhibit often small but yet significant structural difference to their canonical counterparts and facilitate diverse roles in the epigenetic regulation of chromatin and gene expression [511, 512, 518-520]. One of their functional properties is the regulation of DNA accessibility to nuclear factors such as TFs by altering nucleosomal histone-DNA interactions. The histone variant H2A.Z has for example been shown to keep genomic regions like promoters accessibly for both, repressors as well as activators, and plays an important role in cell lineage commitment [521-525]. Another histone variant, the non-histone globular domain containing macroH2A, has been shown to be involved in transcriptional repression [520, 526-528]. Histone variants have also been described to impact on several biological processes apart from transcriptional regulation: DNA replication and DNA repair, heterochromatin maintenance, chromosome packaging or transposon silencing [512, 529, 530].

Taken together, the enormous possibilities of covalent histone modifications and their crosstalk in concert with the evolved cellular machinery to read, write and delete the histone code demonstrates the immense impact of the chromatin template in orchestrating cell fate decisions during ontogeny. The regulation and combinatorial readout of the chromatin (histone) landscape can be influenced itself by the differential expression of chromatin machinery components in a tissue-specific manner. Although a general correlation between nucleosomal modification states and gene activity has been observed, functional interpretations should be carried out carefully.

1.4.1.3 Chromatin regulation beyond DNA methylation and histones modifications

1.4.1.3.1 Chromatin remodeling complexes and chromatin accessibility

As mentioned above, DNA-dependent processes are affected by compaction of nucleosomes into higher order chromatin structures [412, 531]. The position and distribution of nucleosomes in the chromatin fiber determines which DNA regions are accessible to TFs and other nuclear factors, which is of utmost importance for example at *cis*-regulatory regions. Interestingly, ATP-dependent chromatin remodeling complexes (CRCs) have been shown to alter histone-DNA interactions and therefore influence nucleosome positioning by facilitating DNA translocations as well as nucleosome composition/assembly in the chromatin template [532-540]. The CRCs can be classified into four families based on the sequence conservation of their catalytic ATPase subunit: switch/sucrose non-fermentable (SWI/SNF), chromodomain helicase DNA-binding (CHD), imitation switch (ISWI) and INO80 [532, 534, 541, 542]. Apart from the ATPase–translocase domain, these multi-protein complexes comprise several other domains/proteins for protein-protein interactions or histone recognition [532, 534]. These families have been associated with different functional mechanisms (Figure 21). However, some complexes are not solely facilitating only one [534].

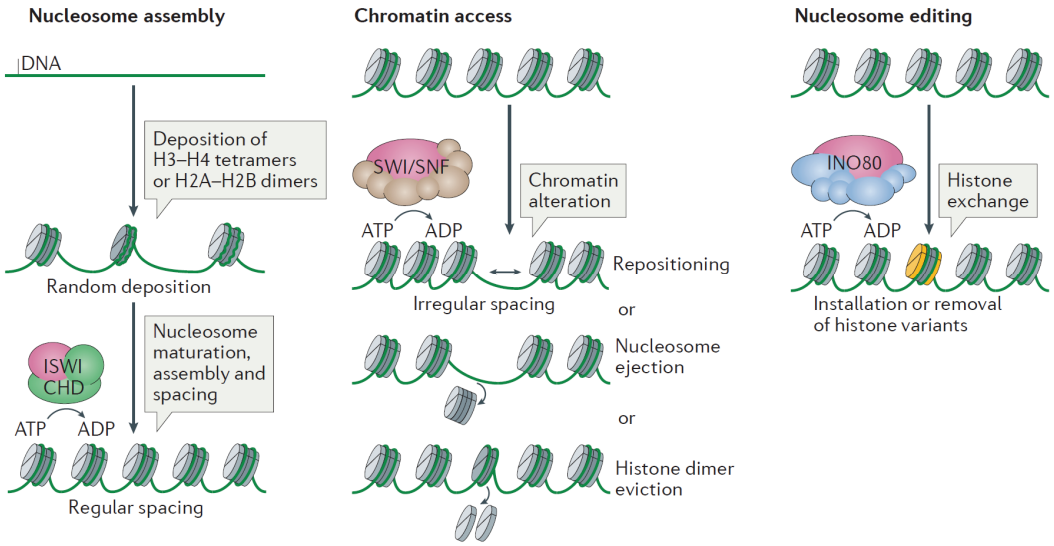


Figure 21: Schematic overview of the functional mechanisms of chromatin remodellers.
"The ATPase–translocase subunit of all remodellers is depicted in pink; additional subunits are depicted in green (imitation switch (ISWI) and chromodomain helicase DNA-binding (CHD)), brown (switch/sucrose non-fermentable (SWI/SNF)) and blue (INO80). Nucleosome assembly: particular ISWI and CHD subfamily remodellers participate in the random deposition of histones, the maturation of nucleosomes and their spacing. Chromatin access: primarily, SWI/SNF subfamily remodellers alter chromatin by repositioning nucleosomes, ejecting octamers or evicting histone dimers. Nucleosome editing: remodellers of the INO80 subfamily (INO80C or Swr1 complex (SWR1C)) change nucleosome composition by exchanging canonical and variant histones, for example, and installing H2A.Z variants (yellow). We note that this functional classification is a simplification, as INO80C, the ISWI remodeller nucleosome remodelling factor (NURF) and certain CHD remodellers can promote chromatin access." Figure and legend taken from [534].

Interestingly, several components of these multi-protein complexes are differentially expressed during cell fate decision processes. This adds another regulatory layer to the action of CRCs in a cell type and

developmental specific dimension by providing other differentially expressed nuclear factors as potential interaction partners which can recruit them to the chromatin template [534, 543, 544]. This is of further interest as CRCs have been shown to be regulated and targeted to specific genomic loci by the interplay of histone PTMs, histone variants and sequence-specific TFs [527, 545-551].

1.4.1.3.2 Non-coding RNAs

All RNA species which are not translated into peptides/proteins are summarised as non-coding RNAs (ncRNAs). Apart from classical examples like transfer RNAs (tRNAs) or ribosomal RNAs (rRNA), both important for protein translation, more and more RNA species transcribed from the non-coding genome were discovered to regulate gene expression and genome organisation through various mechanisms [289, 372, 552-555]. Although ncRNAs can either be expressed in sense or antisense from introns of protein coding genes or intergenic regions, several distinct genes for ncRNAs are additionally present in mammalian genomes [305, 372, 553, 556-559].

Non-coding RNAs with less than 200 nt in length are summarised as short ncRNAs. This group contains for example microRNAs (miRNAs, ~ 22 nt) and small interfering RNAs (siRNAs, ~22 nt). These two classes have gained special interest upon the discovery of a phenomenon termed RNA interference (RNAi; [560, 561]), where gene expression is regulated on post-transcriptional level by sense-antisense RNA pairs [562-567]. On the one hand, miRNAs are derived from a double stranded hairpin-containing precursor RNA structure and are thought to function via repression of protein translation and mRNA degradation due to an imperfect complementarity to their target mRNA [372, 568]. On the other hand, siRNAs are derived from (exogenous) double-stranded RNAs and mostly initiate the degradation of RNA targets due to their perfect complementarity [372, 568]. Non-coding RNAs with more than 200 nt are summarised as long ncRNAs (lncRNAs) and have been implicated in various cellular processes during ontogeny such as X-chromosome inactivation, regulation of transcription or genomic imprinting [569-578]. Current research functionally interconnects the action of non-coding RNAs with TFs and chromatin modifiers (Figure 22) and these RNA species are increasingly appreciated as prominent players in lineage commitment and cell fate determination [574].

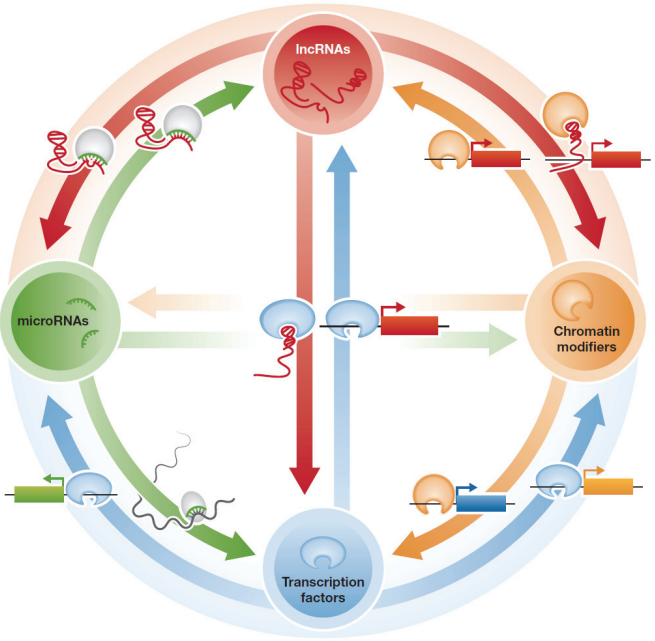


Figure 22: Examples of the interconnection between non-coding RNAs, chromatin modifiers and TFs. "Long noncoding-RNAs (lncRNAs, red RNAs and red arrows) can regulate microRNAs as target site decoys, directly bind to transcription factors as target mimics or as allosteric regulators, and participate in assembly of chromatin-modifying complexes as structural components and recruiters to genomic targets. microRNAs (green RNAs and arrows) post-transcriptionally regulate RNAs from transcription factor, chromatin modifier or lncRNA loci by directly base-pairing to short stretches of RNA. Transcription factors (blue proteins and arrows) can regulate transcription of all the other regulators by directly binding to their promoters. Similarly, chromatin modifiers (orange proteins and arrows) also regulate transcription of the other network components through chromatin modification. Regulatory relationships between microRNAs and chromatin modifier are not depicted." Figure and legend taken from [574].

1.4.2 Chromatin landscape of *cis*-regulatory elements

A new era in epigenomics emerged with the advent of next-generation sequencing (NGS) technologies and their combination with several assays such as DNase footprinting [579], chromatin immunoprecipitation (ChIP; [580, 581]), FAIRE [582, 583] or ATAC [584]. These methods enabled genomewide profiling of histone PTMs, DNA methylation, chromatin accessibility and nuclear factor occupancy at genomic regions of several Metazoa in combination with the increasing availability of their whole genome reference sequences [263, 265, 305, 585-590]. Increasing research characterises the chromatin landscape of *cis*-regulatory elements across several cell types and developmental stages, which advances the understanding of its combinatorial nature inside the nucleus. Distinct histone PTMs, DNA methylation or chromatin accessibility patterns can now be consistently correlated with active or inactive genes (Figure 23).

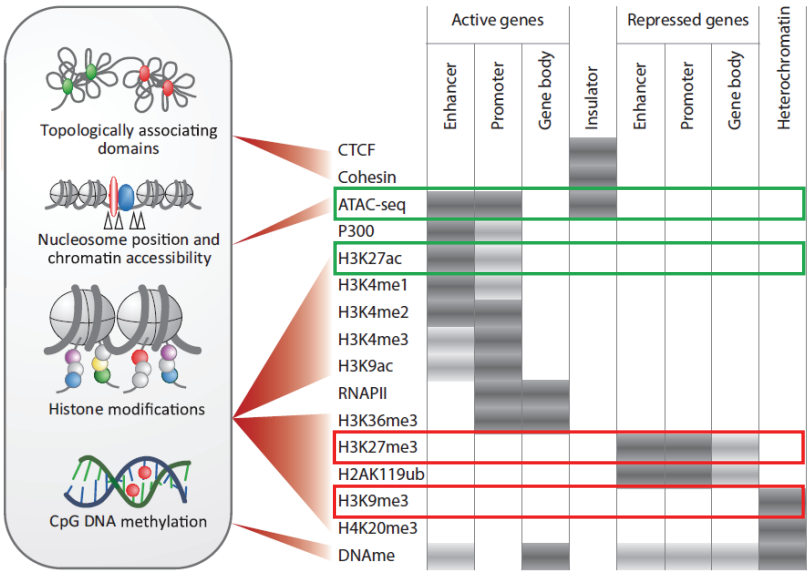


Figure 23: Overview of the chromatin landscape of *cis*-regulatory elements.
"CTCF and Cohesin are involved in the organization of chromosomal loops. ATAC sequencing (indicated) and DNase I sequencing (not shown) can be used to assess chromatin accessibility. The p300 coactivator protein acetylates H3K27 and is mainly found at enhancers. RNAPII causes deposition of H3K36me3 in actively transcribed genes. Lighter gray shades indicate a lower or variable degree of factor binding or histone-tail modification. For DNA methylation (DNAme), light gray indicates that the sequence could be either methylated or unmethylated depending on the example considered. H3K27me3-repressed genes tend to be unmethylated. Active promoters tend to be unmethylated unless they have a low CpG density." Figure and legend taken from [591]. The highlighted euchromatin (green boxes) and heterochromatin markers (facultative and constitutive, red boxes) are of special interest for this thesis.

However, as indicated above, the nature of the chromatin landscape is more dynamic and complex than illustrated in this figure and several other factors are involved, especially the action of TFs.

H3K4me3 is typically associated with active promoters which usually comprise a nucleosome-free region upstream of the TSS. Interestingly it was observed that the polycomb-repressive mark H3K27me3 can coexist with H3K4me3 at certain promoters of key developmental genes in mESCs which are transcriptionally silent [592-594]. These simultaneous occurrence of active and repressive marks is termed "bivalent" chromatin and these promoters of lineage regulators are considered to be in a "poised" state in order to get activated or repressed during proceeding development and differentiation [592, 593, 595-597]. Strikingly, these two antagonistic histone PTMs were proposed to not exist on the same N-terminal tail of the two H3 core histones and therefore cause the appearance of asymmetrically modified nucleosomes [465]. The concept of a "poised" chromatin state has also been extended to enhancers, which are caracterised by H3K4me1 and often H3K27me3 but are less accessible to nuclear factors [279, 598, 599]. In contrast, active enhancers comprise accessible chromatin and are enriched for H3K4me1 and especially H3K27ac, which replaces the repressive H3K27me3 mark of the poised state during differentiation [279, 598, 599]. Interestingly, the enrichment of p300 or H3K27ac has been used to identify tissue-specific enhancers and the implementation of whole transcriptome profiling in the respective tissues with RNA-seq is a major asset for functional interpretations [265, 302, 587, 589, 600, 601]. As active enhancers and promoters are mostly depleted from nucleosomes, the measurement of chromatin accessibility by

FAIRE-seq and ATAC-seq has also been implemented to advance these predictions of active *cis*-regulatory elements [584, 585, 589]. Furthermore, histone variants such as H3.3 or H2A.Z have also been found within enhancer elements [511, 602, 603]. The functional understanding of two characteristic heterochromatin marks at *cis*-regulatory elements, H3K27me3 and H3K9me3, is furthermore advancing by increasing research on their mechanistic differences and genomic distribution. H3K9me3 enrichment was predominantly shown in constitutive heterochromatin which is mostly found in genomic regions with an increased content of repetitive DNA elements and which is silenced across cell types [604, 605]. Additionally, H3K9me3 regions tend to be inaccessible for TFs whereas H3K27me3 repressed promoters can be engaged by Pol II and general TFs [606-609]. The H3K27me3 marked facultative heterochromatin is prevalently found in a cell type-specific manner at *cis*-regulatory regions important for lineage development [610-612]. However, H3K9me3 might also play a role in the cell type-specific regulation of heterochromatin formation, for example at gene clusters encoding for ZFP families [613-616]. Interestingly, the distribution of 5mC and its oxidised intermediates does not occur arbitrarily at *cis*-regulatory elements and their methylation in the context of chromatin landscape has also been linked to transcriptional activation [617-621]. Furthermore, enhancers exhibit 5hmC enrichment and 5fC was found enriched at poised enhancers or exons [265, 622-625]. Nevertheless, these epigenomic assays are mostly based on cell populations. The advent of single-cell assays promises to improve the understanding of cell fate changes within heterogeneous *in vivo* settings, although several technical challenges have to be tackled [626-628].

1.5 Gene regulation during cellular differentiation

As described above, the functional genomic output during ontogeny is affected by the internal and external cellular environment. Therefore, not only the genetic information alone is important for understanding the dynamics of cell fate acquisition, also the multiple players orchestrating the genomic readout are an essential part to gather a comprehensive picture of these processes.

1.5.1 Pioneer transcription factors and cellular reprogramming

Throughout development, cell fate changes are paralleled by the acquisition of distinct cell type-specific transcriptional programs and chromatin landscapes. The competence for these conversions of cellular identity is established by the action of a specific group of TFs called pioneer transcription factors (PTFs) [390, 391, 629]. They have been proposed to exhibit specific functional characteristics which are summarized as follows (adapted from [390]): (1) PTFs sequence-specifically target regulatory elements in a closed, silent chromatin environment prior to their activation, (2) they lead to an increase of chromatin accessibility at target sites upon binding event, also allowing other nuclear proteins to bind and (3) are able to (re)program cell fate (Figure 24).

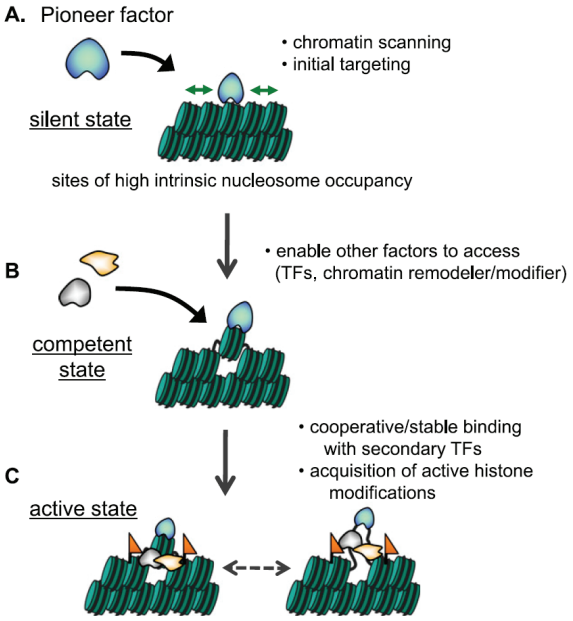


Figure 24: Model for the function of pioneer transcription factors.
Pioneer transcription factors scan and bind target sequences at *cis*-regulatory elements which are embedded in a silent chromatin landscape and recruit other nuclear factors such as CRCs, TFs or histone modifying complexes. Such competent chromatin state is converted into an active state by the collective action of chromatin regulators, leading to the transcriptional induction of associated genes. Figure and legend adapted from [390].

PTFs typically target silent regulatory elements of developmental genes and "set the stage" for other nuclear factors, including Pol II [630], to proceed with the activation of gene regulatory networks determining cell fate decisions and lineage commitment [390, 391, 629, 631, 632]. Characteristic for their mode of action is the ability to access their target sequences directly on nucleosomes, which has been a major focus of PTF research. Interestingly, the DBD of forkhead box A (FoxA) resembles the nucleosomal binding domain of linker histones, which enables this PTF to target its complete binding motif on nucleosomal DNA, replacing linker histones and promoting chromatin accessibility [633-637]. Several other PTFs like Oct4, Klf4 or Ascl1 exhibit different DBD structures and are believed to target partial DNA binding motifs on the nucleosomal surface as their DBD adapts to packaged DNA structures [638].

It is increasingly appreciated that PTFs might play a role for cell fate decision in many lineages during development. A new field of molecular biology (cellular reprogramming) was initiated in the 1980s by studies that used a single TF essential for muscle differentiation, the bHLH protein MyoD (myogenic differentiation), to ectopically convert fibroblasts into myoblasts [639, 640]. Several cocktails of TF combinations (including PTFs) have since been used to initiate cell fate changes across differentiated cell types (transdifferentiation) as well as into induced pluripotent states and the underlying dynamics of the chromatin landscape are increasingly unraveled [641-647]. Especially the field of TF-mediated direct lineage reprogramming into different murine and human glial and neuronal cell types is expanding tremendously [648-651]. However, it has been observed that reprogramming efficiencies were quite low. This has been attributed in part to H3K9me2/H3K9me3-marked heterochromatin regions which seem to constitute a barrier for cell fate changes as for example PTFs are not able to access nucleosomal DNA in this repressed genomic environment [390, 391, 608, 609, 652-654].

1.5.2 Gene regulation during embryonic neurogenesis by transcription factors and chromatin dynamics

The development of the enormous diversity and complexity of neuronal cell types and networks in the brain and specifically in the cortex is a fascinating process and has intrigued researchers to investigate the underlying gene regulatory circuits since decades. Numerous transcription factors and epigenetic mechanisms have been implicated to control the proliferation, cell cycle exit and neuronal differentiation of progenitor cells in the developing embryonic cortex [95, 147, 655-660].

1.5.2.1 Transcriptional regulation of embryonic neurogenesis

Pax6 as major orchestrator of cortical development plays a dual role in regulating progenitor proliferation as well as neurogenesis [661-664]. This homeodomain and paired domain containing TF is expressed in RG cells of the VZ along a cortical gradient during embryonic development and has been shown to regulate several key aspects of forebrain patterning, progenitor proliferation and differentiation as well as neuronal migration [149, 664-677]. Its impact on brain development is apparent as numerous brain abnormalities have been shown to be caused by mutations in the *Pax6* gene [664]. Studies with these mutants have revealed several Pax6 downstream factors including important neuronal TFs such as T-box brain protein 1 (Tbr1), Tbr2, special AT-rich sequence-binding protein 2 (Satb2), Sox2, Neurogenin 1 (Ngn1), Ngn2 or NeuroD1 [148, 666-668, 678-680]. Interestingly, the function of Pax6 in promoting the contrary processes of either progenitor self-renewal or neuronal differentiation was shown to depend on its expression level and affinity to binding-sites, domain activity in protein isoforms as well as its cooperation with other nuclear factors [151, 664, 681-685]. A major focus of recent and ongoing research is the regulation of the *Pax6* gene and protein itself and further the exploration of its role during adult neurogenesis [151, 661, 664]. Interestingly, the pluripotency factor Sox2 also promotes the proliferation of neural progenitors [686, 687] and the sequential occupancy of certain enhancers by Sox2, Sox3 and Sox11 has been described to be crucial for cellular differentiation along neuronal lineage [688, 689]. Apart from *Pax6*, several other fate-determining TFs exhibit expression gradients in VZ progenitors along the rostro-caudal and lateral-medial body axes such as *specificity protein 8* (*Sp8*), *chicken ovalbumin upstream promoter transcription factor 1* (*Coup-TF1*) or *empty spiracles homeobox 2* (*Emx2*), which generates distinct combinations of their expression levels in NSCs facilitating coordinated brain development by providing positional information [133, 690-694]. *Pax6* is among a group of TFs including *Emx1/2*, *LIM homeobox protein 2* (*Lhx2*), *forkhead box protein G1* (*FoxG1*) or the Notch signaling effectors Hes1/5 that are expressed in neural progenitors throughout embryonic neurogenesis in order to preserve the progenitor pool [658, 695-700]. In contrast, several important neuronal TFs such as *Id4*, *Fez family zinc finger 2* (*Fezf2*), *cut-like homeobox 1* (*Cux1*), *Cux2*, *brain-specific homeobox/POU domain protein 1* (*Brn1*) or *Brn2* are differentially expressed in VZ progenitors during embryonic development [658, 701-705]. Furthermore, TFs such as *Insulinoma-associated 1* (*Insm1*) and the aforementioned Pax6 targets *Ngn2* and *Tbr2* are important factors for indirect neurogenesis as they promote the generation of BPs populating the SVZ [144, 658, 706, 707].

The functional interplay of Pax6 with TFs of the bHLH family is crucial for the regulation of the NSC pool [683]. Several members of the bHLH TFs exhibit a prominent role in NSC self-renewal as well as the commitment of neural progenitors towards neurons, oligodendrocytes or astrocytes (Figure 25).

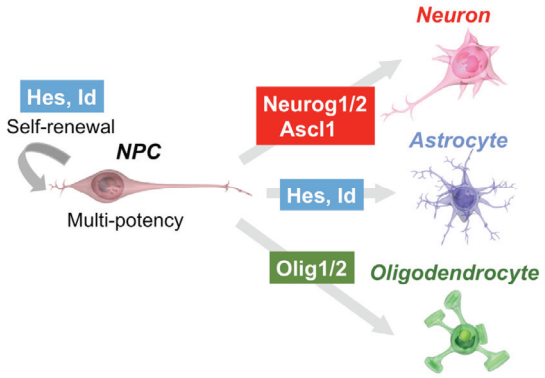


Figure 25: Distinct bHLH TFs are important for NSC self-renewal and cell fate determination.
Figure taken from [217].

Multipotent progenitors are kept in a self-renewing state with the support of Notch-Hes signaling pathway [217, 708-711]. The bHLH factors Hes1 and Hes5 inhibit proneural genes such as *Ngn2* in order to maintain the NSC pool and their depletion along with *Hes3* leads to premature differentiation of progenitors [217, 712, 713]. A similar phenotype was observed in mice with a double knock out for the bHLH factors *Id1* and *Id3* [714]. Interestingly, the expression of bHLH factors such as *Hes1* oscillates in some NSCs via a negative feedback loop, leading to cyclical transcription waves of repressed targets like *Ngn2* [161, 715-717]. This oscillatory control of Notch signaling and bHLH proteins is a key aspect for the maintenance of the progenitor pool in the developing embryonic cortex [161, 716, 718]. Proneural bHLH factors like Ascl1 (Mash1) and Ngn1/2 are promoting neuronal cell fate acquisition of neural progenitors in a spatio-temporal fashion together with other TFs [719-721]. Their importance for the commitment towards neuronal lineage is further supported by the notion that progenitors depleted for *Ngn2* and *Ascl1* fail to generate neurons and rather switch to glial fate [722, 723]. Additionally, several studies have successfully used proneural bHLH factors in combination with other factors like Brn2 or myelin transcription factor 1l (Myt1l) to induce neuronal fate in various cell types [648, 724-726]. Proneural genes like Neurogenins activate the expression of several downstream targets that promote terminal differentiation of the specified neuronal lineage. These targets include bHLH differentiation factors such as *NeuroDs* or *nescient helix-loop-helix 1* (*Nhlh1*) and *Nhlh2* that cooperatively regulate neuronal maturation processes including migration and axon/dendrite- or synapse formation [217, 680, 727-730]. The interdependency of bHLH activators and repressors is therefore a key regulatory layer to control coordinated differentiation of neural progenitors into neurons and glial cells and at the same time maintains the progenitor pool in concert with other nuclear factors and signaling pathways (Figure 26).

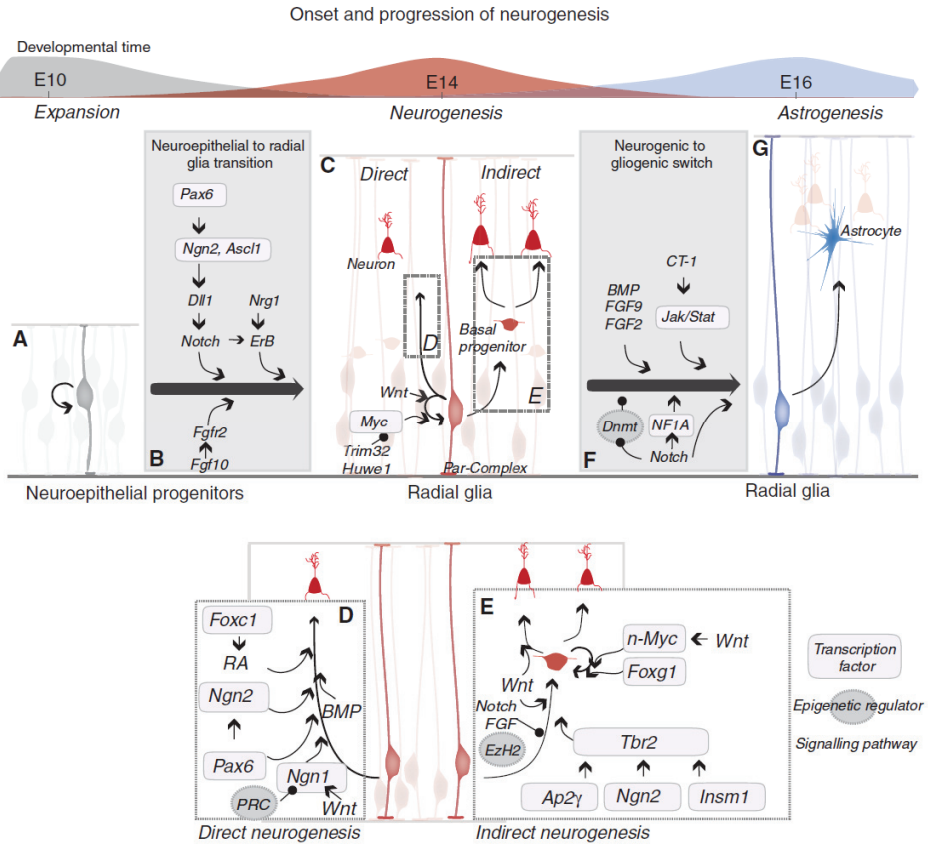


Figure 26: Schematic overview of several molecular layers regulating neuronal differentiation of progenitors and their switch towards gliogenesis.
NE cells (A) switch into RG cells via the action of Pax6 (B) which marks the onset of neurogenesis in the dorsal telencephalon. These RG cells can either directly (D) or indirectly (E) give rise to neurons. During later embryonic development, RG cells produce glial cells (F, G). Each of these steps is regulated by multiple TFs, signaling pathways as well as epigenetic regulators which integrate the external and internal developmental cues into the coordinated formation of a functional brain. Figure and legend adapted from [731].

Many of the above described TFs like neurogenins, Satb2, Tbr1, Tbr2, Emx1 and Emx2 are important for the generation of glutamatergic projection neurons in the dorsal telencephalon [219, 656, 732]. In contrast, the progenitors for inhibitory interneurons in the ventral telencephalon are exposed to a different microenvironment and express another repertoire of TFs including *distal-less homeobox* (*Dlx*) factors or *genomic screened homeobox 1* (*Gsx1*) and *Gsx2* [656, 733]. A plethora of TFs have been identified to build up complex transcriptional networks regulating neuronal differentiation of progenitors and neurons in the developing cortex [656, 658, 731, 734-736]. However, its has to be considered that they all function in the context of chromatin and a full picture of TF-induced chromatin landscape changes during embryonic neurogenesis remains elusive.

1.5.2.2 Chromatin dynamics during embryonic neurogenesis

The field of neuroepigenetics has contributed to the identification of several regulatory elements during neurogenesis and their impact on neuronal development and identity [660, 737-744]. Several studies

employed ChIP-seq technology to predict (fore)brain-specific enhancers by e.g. tissue-specific H3K27ac enrichment or p300 binding [302, 308, 745, 746] and serve as an excellent resource for further functional research. Time course analyses of DNA-methylation products *in vivo* supported the observation that several enhancers are dynamically active during neurogenesis [747]. This is further promoted by differential promoter-enhancer associations during *in vitro* neurogenesis [327]. Interestingly, forebrain specific enhancers that act during early cortical development exhibit a higher sequence conservation in vertebrates in comparison to specific enhancers of other tissues or later (cortical) development [586, 656, 746, 748-751].

Several key regulators of neuronal fate specification like *Pax6*, *Neurogenins* and *Sox* family members are embedded in a bivalent chromatin state in mESCs and gain euchromatin landscape upon differentiation into neural/neuronal lineage [95, 593, 596]. However, fate determinants of the neurogenic lineage become silenced in terminally differentiated neurons to prevent cell death as demonstrated for *Pax6* [752]. In line with these observations, pluripotency genes loose active histone PTMs during neural lineage commitment whereas genes important for neuronal differentiation and neuronal function tend to remain initially poised in NSCs [95, 596, 655]. Not surprisingly, genes of non-neuronal lineages are permanently silenced in NSCs and their progeny [753]. These dynamics of the chromatin landscape have also been observed during reprogramming experiments into neuronal cells (Figure 27). A recent publication provided genome wide H3K4me3 and H3K27me3 ChIP-seq data for several murine cortical progenitor cell types and neurons *in vivo* [754]. These data extended the concept of bivalent chromatin also to neurons as neuronal maturation genes which were expressed in postnatal neurons such as *Grin1*, *Npy* or several calcium channel subunits were found embedded in a poised chromatin landscape in embryonic neurons [754].

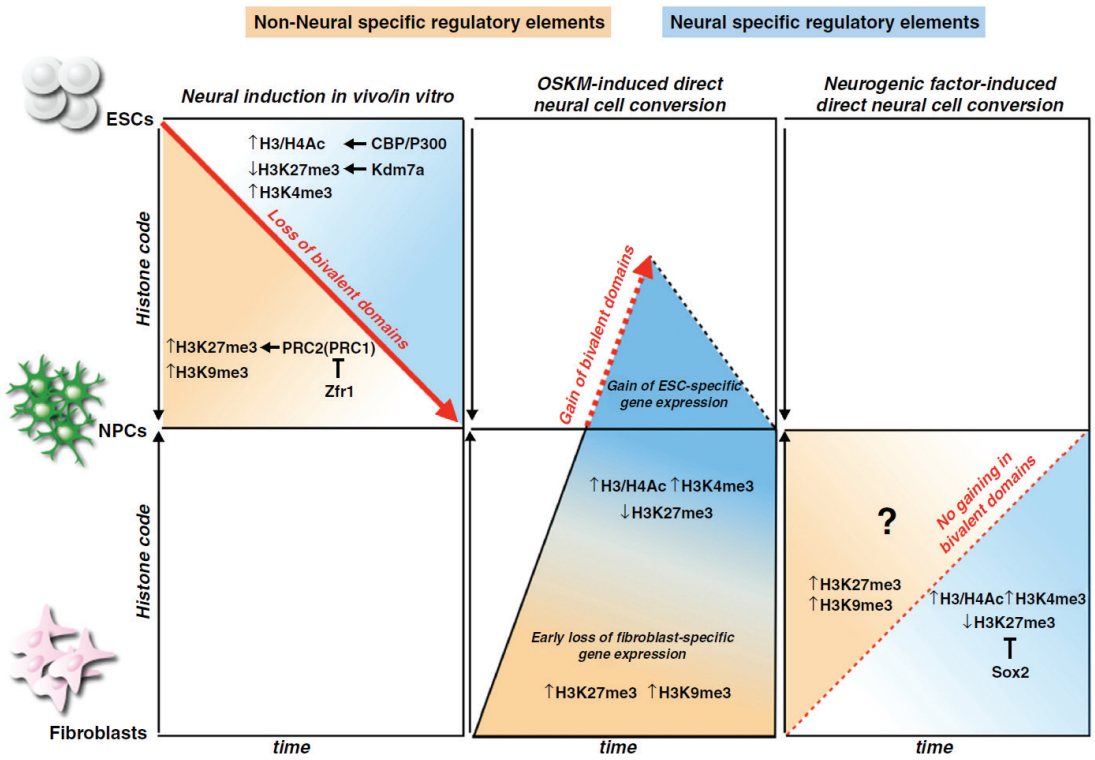


Figure 27: Chromatin landscape dynamics during neuronal differentiation of ESCs and direct neuronal conversion of fibroblasts.
"Schematic representation of histone modifications occurring at both non-neural (orange) and neural specific (blue) regulatory elements during *in vitro* and *in vivo* differentiation of ESCs into NPCs (left panel) and from direct conversion of somatic fibroblasts (middle and right panel). In ESC differentiation into NPCs, regulative elements progressively lose their bivalent state. In particular, in neural specific regulatory regions repressive H3K27me3 and H3K9me3 marks are removed by specific demethylases (e.g. Kdm7a). Moreover, HATs, such as CBP and P300, cooperate in transcriptional activation by acetylation of H3/H4. Repression of non-neural genes is achieved by deposition of H3K27me3 and H3K9me3 marks. In the direct conversion of fibroblasts to NSCs by OSKM transient expression, a progressive gain in bivalent state is achieved at least for some gene regulatory elements. Initially, fibroblast specific genes are silenced thanks to H3K27me3 and H3K9me3. In the past days, when neural medium is added, probably new epigenetic rearrangements stabilize neural gene expression while repress other fates. In direct conversion of fibroblasts to NPCs by neurogenic TF expression, probably no ESC-like bivalent state is achieved. We hypothesize that thanks to the action of the pioneer factors (e.g. Sox2 and others) repressed chromatin is unfolded and associated H3K27/H3K9 methylation is erased in neural-specific genetic loci selectively." Figure and legend taken from [755].

The dynamics of the chromatin landscape suggests that writers and erasers of these histone PTMs are eminent players in the process of cell fate commitment and differentiation into neurons. Indeed, the depletion of the PRC2 catalytic subunit *Ezh2* in NSCs during early neurogenesis results in a loss of H3K27me3 and concomitantly premature neurogenesis, both directly and indirectly [756]. Jmjd3, an erasers of H3K27me2/3 and therefore antagonist of Ezh2, has been implicated as essential for neural cell fate commitment as it regulates key effectors of neurogenesis like *Pax6*, *Nestin* and *Sox1* [757, 758]. The H3K4me2/3 demethylase Jarid1b was shown to contribute to the silencing of pluripotency genes during neuronal differentiation of mESCs [759]. Interestingly, the ability of neural progenitors to self-renew was severely impaired upon depletion of the histone demethylase *Lsd1* [760]. Additionally, *Setdb1* is highly expressed during early neurogenesis and its mediated methylation of H3K9 is essential to prevent premature differentiation of neural progenitors into astrocytes and severe brain defects leading to early lethality [761].

Apart from histone methylation turnover, also histone acetylation plays an essential role during neurogenesis. The lysine-specific HAT Kat6b was identified as a central player for embryonic cortical development by activating several genes for neuronal cell fate commitment [762]. Another HAT, CBP, cooperates with Ngn1 in order to promote neuronal fate in NSCs and inhibiting their glial differentiation, a principle which is also employed by proneural bHLH factors and p300 to activate gene transcription [730, 763-765]. The availability of several epigenetic regulators and TFs which target these regulators to *cis*-regulatory elements at distinct developmental time points in specific cell types plays an essential role in the regulation of neural development. This is further shown by cell type-specific expression patterns of certain histone deacetylases: although *HDAC1* and *HDAC2* are expressed in NSCs, only *HDAC2* is present in neurons whereas *HDAC1* expression is restricted to glial cells [766]. HDAC2 is important for the maintenance of progenitor proliferation, a function which is also associated to HDAC5 or HDAC3 [767]. Furthermore, DNA methyltransferases are differentially expressed during embryonic neurogenesis, suggesting a distinct functionality: Whereas *Dnmt1* is expressed in NSCs and neurons and *Dnmt3a* in SVZ progenitors and neurons, *Dnmt3b* expression is not detectable in the brain after E15.5 [768, 769]. Interestingly, the depletion of *Dnmt1* in neural progenitors induces a precocious glial differentiation by activating JAK-STAT signaling genes [770]. The silencing of non-neuronal lineage genes in NSCs by DNA methylation has also been shown to function in concert with H3K27me3 Polycomb repression [104], suggesting a crosstalk between different epigenetic regulatory layers. Also the role of the TET family enzymes in neurogenesis is increasingly appreciated as 5hmC has been found to be particularly enriched in neurons, for example in introns of neuron-specific genes [771-774].

Chromatin remodeling complexes are associated with specific functions during neurogenesis [775]. The *Chromodomain Helicase DNA-binding protein 5* (*Chd5*) is transcriptionally induced during early neurogenesis end remains highly expressed in postnatal neurons [776]. Several neuronal genes are activated by Chd5 [777] and its depletion causes neurogenesis defects, partly due to its second role in maintaining H3K27me3 marked facultative heterochromatin at certain developmental genes [778]. A well-studied chromatin remodeling complex important for neurogenesis is the BRG1- or hBRM-associated factor (BAF) complex of the SWI/SNF family which exhibits a cell type-specific subunit exchange during differentiation of mESCs (esBAF) to neural progenitors (npBAF) and further into neurons (nBAF) and hence fulfills specific functions [779-782] (Figure 28).

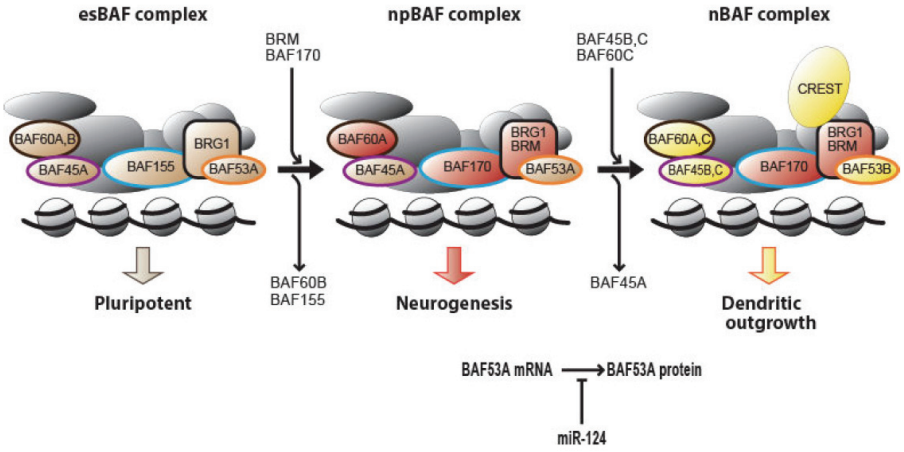


Figure 28: Dynamics of BAF complex components during neuronal differentiation of mESCs. "The exchange of the components is essential for the transition from ESCs to post-mitotic neurons. The exchangeable subunits are colored as follows: esBAF complex-specific subunit, ocher; npBAF complex-specific subunit, red; and nBAF complex-specific subunit, yellow. Cell type-specific BAF complexes have distinct functions that are indispensable for their properties. The BAF complex in ESCs, NPCs, and neurons are defined as esBAF, npBAF, and nBAF, respectively. The microRNA miR-124, binds to BAF53A mRNA transcripts to suppress its expression, thereby facilitating the replacement of BAF53 in the npBAF complexes." Figure and legend taken from [753].

The npBAF complex was demonstrated as an essential regulator of neocortical size and thickness by contributing to NSC proliferation, possibly by interacting with Pax6 as well as promoting Notch signaling by activating *Hes1/5* [783-787]. The npBAF and nBAF complexes are believed to be essential regulators of the neural specific chromatin landscapes, also during reprogramming [788]. Some BAF subunits like Brg1 and BAF170 were found in a complex with the repressor element 1-silencing transcription factor (REST) and its co-repressors [789], suggesting that REST is involved in the BAF targeting and recruitment of other histone modifiers [95, 790-792]. REST is considered to be a master regulator of neuronal differentiation by repressing neuronal genes in non-neuronal cells [775, 790, 793]. Interestingly, *REST* depletion in mESCs does not lead to an upregulation of neural genes like *Pax6*, *Ngn1* or *Ngn2* whereas neuronal target genes become de-repressed [794]. The physiological downregulation of *REST* in neurons contributes to the induction of the neuronal transcriptome [775, 790, 793, 795]. A recent study has additionally revealed that REST binding at *cis*-regulatory elements is directly required to prevent them from DNA-methylation in NSCs, especially enhancers [106]. Furthermore it has been shown that the cell fate commitment of mESCs towards neuronal lineage and the progression from NSCs to neurons is dependent on the regulation of *REST* levels itself [95, 788, 796, 797]. One additional player in this scenario is another CRC, CHD2, which directly regulates the expression of *REST* in VZ progenitors [798]. Interestingly, also rearrangements of the global chromatin architecture in the nucleus have been implicated to regulate neurogenesis [799-801].

Unraveling the role of lncRNAs during embryonic neurogenesis is a relatively recent field of research and only few examples are described so far [574, 802]. Some screening studies revealed several lncRNAs with a cell type and spatio-temporal specific expression profile in the murine brain [803, 804], but a functional understanding is still vague. Interestingly, the *Tcl1 Upstream Neuron-Associated* (*TUNA*) lncRNA plays an important role in the maintenance of pluripotency in mESCs and its depletion impairs their neural

differentiation potential [805]. Furthermore, the depletion of the neural-specific lncRNA *Pinky* (*Pnky*) in the embryonic cortex resulted in a depletion of the NSC pool and premature neuronal differentiation [806]. In contrast, the ectopic overexpression of *Nkx2.2as*, an antisense lncRNA to *Nkx2.2*, in neural progenitors enhanced their differentiation into glial lineage, namely oligodendrocytes [807]. Strikingly, it has been shown that *Pax6* is directly regulated by *Paupar*, a lncRNA which is expressed upstream of the *Pax6* locus, in a *cis*-acting mechanism [808]. In contrast, several miRNAs have been observed to regulate brain development [809-813]. Numerous miRNAs like miR-9*, miR-124a, miR-124b and miR-135 are expressed brain-specifically in certain cell types and regulate important nuclear factors involved in neurogenesis like REST or BAF53a [809, 812].

The regulation of epigenetic layers in NSCs plays an essential role during ontogeny. Although several mechanisms have been described for the switch from neurogenesis to gliogenesis during later embryonic development [104, 655, 814, 815], the underlying molecular mechanisms of this temporal NPC fate switch during embryonic corticogenesis are not fully understood. This is also increasingly complex in the context of the observed derivation of adult NSCs as a progeny of embryonic progenitors [816, 817]. However, the evidence of mutations in epigenetic regulators causing neurodevelopmental or neurodegenerative diseases [818-820] warrants for further investigations on these regulatory layers and how their crosstalk facilitates brain development.

1.5.3 Regulation of transposable elements during development

The impact of transposable elements on the host genome is a growing field of biomedical research [247, 249, 347, 352, 356, 361, 821, 822]. However, a significant part of these studies have been conducted in cell types of the early embryo or germ cells. These studies show for example that ERV-driven transcription serves a regulatory function in the PGRN and is important for oocyte function and fertility [367, 821, 823-826]. This phenomenon seems not to be restricted to protein-coding genes as several examples show that ERVs can also function as promoter elements for ncRNA species in a tissue-specific manner [356, 361, 827, 828]. Several ERV classes seem to accumulate at *cis*-regulatory regions and it has also been observed that ERVs have been co-opted as (species-specific) enhancer elements [349, 356, 362]. However, apart from examples of a co-evolved beneficial impact on host genomes, the majority of transposable elements has to be repressed in order to preserve genome integrity throughout evolution and ontogeny. It was proposed that in mESCs the histone variant H3.3 is involved in silencing ERVs by regulating H3K9me3 levels via the KAP1/SETDB1 system, notably for IAP and MusD elements [829]. Strikingly, the *de novo* methylation of ERVs in mESCs is directed by the same heterochromatin machinery in concert with KRAB-ZFPs, recapitulating DNA methylation based ERV silencing during early *in vivo* development [249, 368, 830]. These mechanism are furthermore important, as it has been shown for (only) a small number of ERV elements to prevent them from potentially functioning as enhancers by gaining active chromatin states and perturbing transcriptome dynamics (Figure 29).

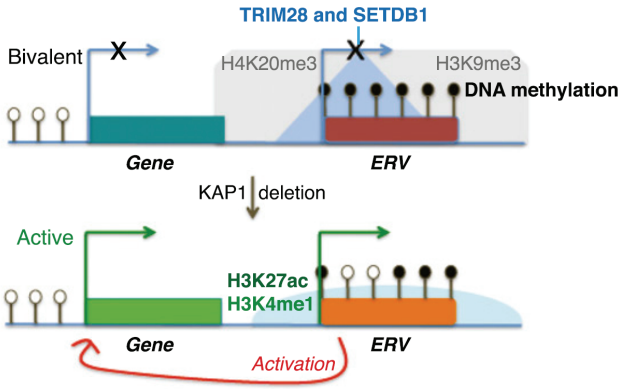


Figure 29: Model for ERV elements gaining enhancer function upon impaired heterochromatin mediated silencing.
ERV elements that escape KRAB-ZFP/KAP1/SETDB1-mediated silencing in mESCs can gain an active chromatin landscape and act as enhancers at specific genomic loci. Figure taken from [831].

Nonetheless, the KAP1/SETDB1 system is believed to be the crucial asset for regulation of ERV silencing throughout development. Increasing attention is drawn to the tetrapod-specific family of KRAB-ZFP proteins as these are believed to be the sequence-specific linkers between retrotransposon DNA sequence and heterochromatin machinery [235, 356]. However, only a few factors such as ZFP809 or ZFP57 were mechanistically studied in ESCs to date, but recent findings suggest that KRAB-ZFPs and their co-repressor are also involved in silencing transposable elements in adult tissues [244, 245, 256, 832, 833]. Interestingly, pioneering work by Fasching and colleagues observed that several normally H3K9me3-marked ERV elements like IAP1 and MMERVK10C were reactivated and influenced the transcription of nearby genes, including lncRNAs, after a conditional knock out of KAP1 in neural progenitors [834]. The mechanistic regulation of ERV silencing in brain development is not well explored but of increasing interest, especially in the context that several neuronal enhancers of the mammalian neocortex originate from ERV elements [835-837] and that KAP1 deletion in the adult murine forebrain alters behavioral response to stress [838]. Remarkably, the misregulation of transposon silencing in the human brain has also been associated with several neurological disorders including Rett syndrome, schizophrenia or neurodegeneration [839-844]. Additionally, LINE1 transposons are implicated to contribute to the complexity of regulatory neuronal networks by actively transposing during the transition from neuronal progenitors to neurons and thereby exposing their *cis*-regulatory sequences to new chromatin landscapes [821, 845-851]. This somatic mosaicism is belief to contribute to the interindividual diversity of behavior, for example [247, 852-854]. A comprehensive mechanistic understanding of the regulation of transposon silencing during ontogeny in order to prevent detrimental effects to the host (Figure 30) is still not fully unraveled. Future research is warranted especially in the light of the tetrapod- and cell type-specific KRAB-ZFPs and mammalian neocortex development.

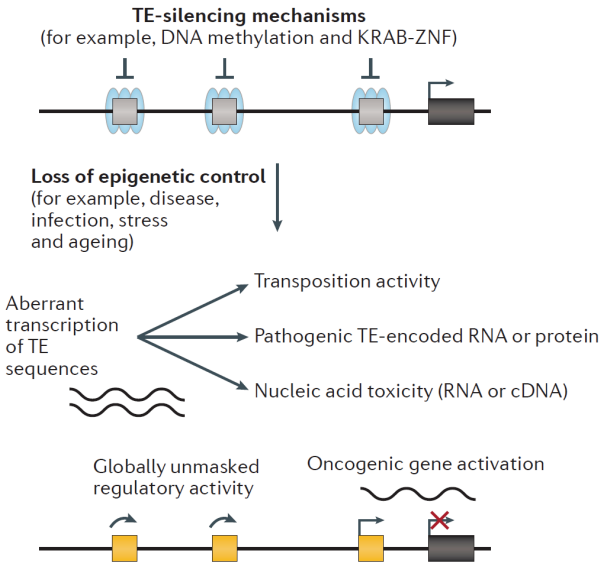


Figure 30: Potential detrimental cellular consequence of transposon reactivation upon loss of epigenetic silencing. Figure taken from [349].

1.6 References Introduction

1. Bell, E.A., et al., *Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon.* Proc Natl Acad Sci U S A, 2015. **112**(47): p. 14518-21.
2. Gollihar, J., M. Levy, and A.D. Ellington, *Biochemistry. Many paths to the origin of life.* Science, 2014. **343**(6168): p. 259-60.
3. Jheeta, S., *The Landscape of the Emergence of Life.* Life (Basel), 2017. **7**(2).
4. Walker, S.I., *Origins of life: a problem for physics, a key issues review.* Rep Prog Phys, 2017. **80**(9): p. 092601.
5. Virchow, R., *Cellular-Pathologie. In: Virchows Archiv.* 1855(Band 8): p. 3–39.
6. Miescher, F., *Über die chemische Zusammensetzung der Eiterzellen.* Hoppe. Seyler's Med. Chem. Unters., 1871. **4**: p. 441–460.
7. Flemming, W., *Neue Beiträge zur Kenntniss der Zelle (1st part).* Arch. Mikr. Anat., 1879. **29**: p. 389–463.
8. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid.* Nature, 1953. **171**(4356): p. 737-8.
9. Allis, C.D. and T. Jenuwein, *The molecular hallmarks of epigenetic control.* Nat Rev Genet, 2016. **17**(8): p. 487-500.
10. Misteli, T., *The cell biology of genomes: bringing the double helix to life.* Cell, 2013. **152**(6): p. 1209-12.
11. Raj, A. and A. van Oudenaarden, *Nature, nurture, or chance: stochastic gene expression and its consequences.* Cell, 2008. **135**(2): p. 216-26.
12. Rokas, A., *The origins of multicellularity and the early history of the genetic toolkit for animal development.* Annu Rev Genet, 2008. **42**: p. 235-51.
13. Grosberg, R.K. and R.R. Strathmann, *The evolution of multicellularity: A minor major transition?* Annual Review of Ecology Evolution and Systematics, 2007. **38**: p. 621-654.
14. Knoll, A.H., *The multiple origins of complex multicellularity.* Annu. Rev. Earth Planet. Sci. , 2011. **39**(1): p. 217–239.
15. Narbonne, G.M., *The Ediacara biota: neoproterozoic origin of animals and their ecosystems.* Annu. Rev. Earth Planet. Sci. , 2005. **33**(1): p. 421–442.
16. Richter, D.J. and N. King, *The genomic and cellular foundations of animal origins.* Annu Rev Genet, 2013. **47**: p. 509-37.
17. Medvedev, S.P., A.I. Shevchenko, and S.M. Zakian, *Molecular basis of Mammalian embryonic stem cell pluripotency and self-renewal.* Acta Naturae, 2010. **2**(3): p. 30-46.
18. Dye, F.J., *Developmental Cell Biology.* Reviews in Cell Biology and Molecular Medicine., 2016. **2**: p. 118–143.
19. Moris, N., C. Pina, and A.M. Arias, *Transition states and cell fate decisions in epigenetic landscapes.* Nat Rev Genet, 2016. **17**(11): p. 693-703.
20. Takaoka, K. and H. Hamada, *Cell fate decisions and axis determination in the early mouse embryo.* Development, 2012. **139**(1): p. 3-14.
21. Rossant, J. and P.P. Tam, *Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse.* Development, 2009. **136**(5): p. 701-13.
22. Chazaud, C. and Y. Yamanaka, *Lineage specification in the mouse preimplantation embryo.* Development, 2016. **143**(7): p. 1063-74.
23. Keller, G., *Embryonic stem cell differentiation: emergence of a new era in biology and medicine.* Genes Dev, 2005. **19**(10): p. 1129-55.
24. Avilion, A.A., et al., *Multipotent cell lineages in early mouse development depend on SOX2 function.* Genes Dev, 2003. **17**(1): p. 126-40.
25. Guo, G., et al., *Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst.* Dev Cell, 2010. **18**(4): p. 675-85.
26. Biase, F.H., X. Cao, and S. Zhong, *Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing.* Genome Res, 2014. **24**(11): p. 1787-96.
27. Deng, Q., et al., *Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells.* Science, 2014. **343**(6167): p. 193-6.
28. Piras, V., M. Tomita, and K. Selvarajoo, *Transcriptome-wide variability in single embryonic development cells.* Sci Rep, 2014. **4**: p. 7137.
29. Shi, J., et al., *Dynamic transcriptional symmetry-breaking in pre-implantation mammalian embryo development revealed by single-cell RNA-seq.* Development, 2015. **142**(20): p. 3468-77.

30. Rossant, J. and W.T. Lis, *Potential of isolated mouse inner cell masses to form trophectoderm derivatives in vivo.* Dev Biol, 1979. **70**(1): p. 255-61.

31. Rossant, J. and K.M. Vijh, *Ability of outside cells from preimplantation mouse embryos to form inner cell mass derivatives.* Dev Biol, 1980. **76**(2): p. 475-82.

32. Suwinska, A., et al., *Blastomeres of the mouse embryo lose totipotency after the fifth cleavage division: expression of Cdx2 and Oct4 and developmental potential of inner and outer blastomeres of 16- and 32-cell embryos.* Dev Biol, 2008. **322**(1): p. 133-44.

33. Ziomek, C.A., M.H. Johnson, and A.H. Handyside, *The developmental potential of mouse 16-cell blastomeres.* J Exp Zool, 1982. **221**(3): p. 345-55.

34. Chazaud, C., et al., *Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway.* Dev Cell, 2006. **10**(5): p. 615-24.

35. Kurimoto, K., et al., *An improved single-cell cDNA amplification method for efficient high-density oligonucleotide microarray analysis.* Nucleic Acids Res, 2006. **34**(5): p. e42.

36. Plusa, B., et al., *Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst.* Development, 2008. **135**(18): p. 3081-91.

37. Mitsui, K., et al., *The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells.* Cell, 2003. **113**(5): p. 631-42.

38. Nichols, J., et al., *Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4.* Cell, 1998. **95**(3): p. 379-91.

39. Menchero, S., et al., *Signaling pathways in mammalian preimplantation development: Linking cellular phenotypes to lineage decisions.* Dev Dyn, 2017. **246**(4): p. 245-261.

40. Haeckel, E., *Natürliche Schöpfungsgeschichte.* Berlin: Georg Reimer, 1868.

41. Haeckel, E., *Anthropogenie, 3rd edn.* Leipzig: Wilhelm Engelmann, 1877.

42. Pappenheim, A., *Über Entwicklung und Ausbildung der Erythroblasten.* Virchows Arch., 1896. **145**: p. 587-643.

43. Becker, A.J., C.E. Mc, and J.E. Till, *Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells.* Nature, 1963. **197**: p. 452-4.

44. Till, J.E. and C.E. Mc, *A direct measurement of the radiation sensitivity of normal mouse bone marrow cells.* Radiat Res, 1961. **14**: p. 213-22.

45. Till, J.E., E.A. McCulloch, and L. Siminovitch, *A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells.* Proc Natl Acad Sci U S A, 1964. **51**: p. 29-36.

46. Stevens, L.C. and C.C. Little, *Spontaneous Testicular Teratomas in an Inbred Strain of Mice.* Proc Natl Acad Sci U S A, 1954. **40**(11): p. 1080-7.

47. Kleinsmith, L.J. and G.B. Pierce, Jr., *Multipotentiality of Single Embryonal Carcinoma Cells.* Cancer Res, 1964. **24**: p. 1544-51.

48. Stevens, L.C., *The biology of teratomas.* Adv Morphog, 1967. **6**: p. 1-31.

49. Pierce, G.B., *Teratocarcinoma: model for a developmental concept of cancer.* Curr Top Dev Biol, 1967. **2**: p. 223-46.

50. Evans, M.J., *The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells.* J Embryol Exp Morphol, 1972. **28**(1): p. 163-76.

51. Kahan, B.W. and B. Ephrussi, *Developmental potentialities of clonal in vitro cultures of mouse testicular teratoma.* J Natl Cancer Inst, 1970. **44**(5): p. 1015-36.

52. Rosenthal, M.D., R.M. Wishnow, and G.H. Sato, *In vitro growth and differetiation of clonal populations of multipotential mouse clls derived from a transplantable testicular teratocarcinoma.* J Natl Cancer Inst, 1970. **44**(5): p. 1001-14.

53. Finch, B.W. and B. Ephrussi, *RETENTION OF MULTIPLE DEVELOPMENTAL POTENTIALITIES BY CELLS OF A MOUSE TESTICULAR TERATOCARCINOMA DURING PROLONGED CULTURE in vitro AND THEIR EXTINCTION UPON HYBRIDIZATION WITH CELLS OF PERMANENT LINES.* Proc Natl Acad Sci U S A, 1967. **57**(3): p. 615-21.

54. Morrison, S.J., N.M. Shah, and D.J. Anderson, *Regulatory mechanisms in stem cell biology.* Cell, 1997. **88**(3): p. 287-98.

55. Till, J.E. and E.A. McCulloch, *Hemopoietic stem cell differentiation.* Biochim Biophys Acta, 1980. **605**(4): p. 431-59.

56. Weissman, I.L., *Stem cells: units of development, units of regeneration, and units in evolution.* Cell, 2000. **100**(1): p. 157-68.

57. Martin, G.R. and M.J. Evans, *Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro.* Proc Natl Acad Sci U S A, 1975. **72**(4): p. 1441-5.

58. Evans, M.J. and M.H. Kaufman, *Establishment in culture of pluripotential cells from mouse embryos.* Nature, 1981. **292**(5819): p. 154-6.

59. Martin, G.R., *Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells.* Proc Natl Acad Sci U S A, 1981. **78**(12): p. 7634-8.

60. Robertson, E., et al., *Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector.* Nature, 1986. **323**(6087): p. 445-8.

61. Thompson, S., et al., *Germ line transmission and expression of a corrected HPRT gene produced by gene targeting in embryonic stem cells.* Cell, 1989. **56**(2): p. 313-21.

62. Smith, A.G. and M.L. Hooper, *Buffalo rat liver cells produce a diffusible activity which inhibits the differentiation of murine embryonal carcinoma and embryonic stem cells.* Dev Biol, 1987. **121**(1): p. 1-9.

63. Smith, A.G., et al., *Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides.* Nature, 1988. **336**(6200): p. 688-90.

64. Williams, R.L., et al., *Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells.* Nature, 1988. **336**(6200): p. 684-7.

65. Rathjen, P.D., et al., *Developmentally programmed induction of differentiation inhibiting activity and the control of stem cell populations.* Genes Dev, 1990. **4**(12B): p. 2308-18.

66. Stewart, C.L., et al., *Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor.* Nature, 1992. **359**(6390): p. 76-9.

67. Boeuf, H., et al., *Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells.* J Cell Biol, 1997. **138**(6): p. 1207-17.

68. Martello, G., P. Bertone, and A. Smith, *Identification of the missing pluripotency mediator downstream of leukaemia inhibitory factor.* EMBO J, 2013. **32**(19): p. 2561-74.

69. Matsuda, T., et al., *STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells.* EMBO J, 1999. **18**(15): p. 4261-9.

70. Niwa, H., et al., *Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3.* Genes Dev, 1998. **12**(13): p. 2048-60.

71. Gardner, R.L. and F.A. Brook, *Reflections on the biology of embryonic stem (ES) cells.* Int J Dev Biol, 1997. **41**(2): p. 235-43.

72. Nichols, J. and A. Smith, *The origin and identity of embryonic stem cells.* Development, 2011. **138**(1): p. 3-8.

73. Nichols, J. and A. Smith, *Pluripotency in the embryo and in culture.* Cold Spring Harb Perspect Biol, 2012. **4**(8): p. a008128.

74. Martello, G. and A. Smith, *The nature of embryonic stem cells.* Annu Rev Cell Dev Biol, 2014. **30**: p. 647-75.

75. Nagy, A., et al., *Derivation of completely cell culture-derived mice from early-passage embryonic stem cells.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8424-8.

76. Li, M. and J.C. Belmonte, *Ground rules of the pluripotency gene regulatory network.* Nat Rev Genet, 2017. **18**(3): p. 180-191.

77. Wu, J. and J.C. Izpisua Belmonte, *Dynamic Pluripotent Stem Cell States and Their Applications.* Cell Stem Cell, 2015. **17**(5): p. 509-25.

78. Ambrosetti, D.C., C. Basilico, and L. Dailey, *Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites.* Mol Cell Biol, 1997. **17**(11): p. 6321-9.

79. Chambers, I., et al., *Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells.* Cell, 2003. **113**(5): p. 643-55.

80. Chambers, I., et al., *Nanog safeguards pluripotency and mediates germline development.* Nature, 2007. **450**(7173): p. 1230-4.

81. Ivanova, N., et al., *Dissecting self-renewal in stem cells with RNA interference.* Nature, 2006. **442**(7102): p. 533-8.

82. Masui, S., et al., *Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells.* Nat Cell Biol, 2007. **9**(6): p. 625-35.

83. Niwa, H., J. Miyazaki, and A.G. Smith, *Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells.* Nat Genet, 2000. **24**(4): p. 372-6.

84. Scholer, H.R., et al., *A family of octamer-specific proteins present during mouse embryogenesis: evidence for germline-specific expression of an Oct factor.* EMBO J, 1989. **8**(9): p. 2543-50.

85. Silva, J., et al., *Nanog is the gateway to the pluripotent ground state.* Cell, 2009. **138**(4): p. 722-37.

86. Suzuki, A., et al., *Nanog binds to Smad1 and blocks bone morphogenetic protein-induced differentiation of embryonic stem cells.* Proc Natl Acad Sci U S A, 2006. **103**(27): p. 10294-9.

87. Thomson, M., et al., *Pluripotency factors in embryonic stem cells regulate differentiation into germ layers.* Cell, 2011. **145**(6): p. 875-89.

88. Okamoto, K., et al., *A novel octamer binding transcription factor is differentially expressed in mouse embryonic cells.* Cell, 1990. **60**(3): p. 461-72.

89. Smith, A.G., *Embryo-derived stem cells: of mice and men.* Annu Rev Cell Dev Biol, 2001. **17**: p. 435-62.

90. Turksen, K., *Preface. Human Embryonic Stem Cell Protocols.* Methods Mol Biol, 2016. **1307**: p. v.

91. Daley, G.Q. and D.T. Scadden, *Prospects for stem cell-based therapy.* Cell, 2008. **132**(4): p. 544-8.

92. Mahla, R.S., *Stem Cells Applications in Regenerative Medicine and Disease Therapeutics.* Int J Cell Biol, 2016. **2016**: p. 6940283.

93. Tabar, V. and L. Studer, *Pluripotent stem cells in regenerative medicine: challenges and recent progress.* Nat Rev Genet, 2014. **15**(2): p. 82-92.

94. Daley, G.Q., M.A. Goodell, and E.Y. Snyder, *Realistic prospects for stem cell therapeutics.* Hematology Am Soc Hematol Educ Program, 2003: p. 398-418.

95. Hirabayashi, Y. and Y. Gotoh, *Epigenetic control of neural precursor cell fate during development.* Nat Rev Neurosci, 2010. **11**(6): p. 377-88.

96. Bain, G., et al., *Embryonic stem cells express neuronal properties in vitro.* Dev Biol, 1995. **168**(2): p. 342-57.

97. Barberi, T., et al., *Neural subtype specification of fertilization and nuclear transfer embryonic stem cells and application in parkinsonian mice.* Nat Biotechnol, 2003. **21**(10): p. 1200-7.

98. Gaspard, N., et al., *An intrinsic mechanism of corticogenesis from embryonic stem cells.* Nature, 2008. **455**(7211): p. 351-7.

99. Kawasaki, H., et al., *Induction of midbrain dopaminergic neurons from ES cells by stromal cell-derived inducing activity.* Neuron, 2000. **28**(1): p. 31-40.

100. Okabe, S., et al., *Development of neuronal precursor cells and functional postmitotic neurons from embryonic stem cells in vitro.* Mech Dev, 1996. **59**(1): p. 89-102.

101. Tropepe, V., et al., *Direct neural fate specification from embryonic stem cells: a primitive mammalian neural stem cell stage acquired through a default mechanism.* Neuron, 2001. **30**(1): p. 65-78.

102. Bibel, M., et al., *Differentiation of mouse embryonic stem cells into a defined neuronal lineage.* Nat Neurosci, 2004. **7**(9): p. 1003-9.

103. Bibel, M., et al., *Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells.* Nat Protoc, 2007. **2**(5): p. 1034-43.

104. Mohn, F., et al., *Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors.* Mol Cell, 2008. **30**(6): p. 755-66.

105. Lienert, F., et al., *Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells.* PLoS Genet, 2011. **7**(6): p. e1002090.

106. Stadler, M.B., et al., *DNA-binding factors shape the mouse methylome at distal regulatory regions.* Nature, 2011. **480**(7378): p. 490-5.

107. Plachta, N., et al., *Developmental potential of defined neural progenitors derived from mouse embryonic stem cells.* Development, 2004. **131**(21): p. 5449-56.

108. Hartenstein, V. and A. Stollewerk, *The evolution of early neurogenesis.* Dev Cell, 2015. **32**(4): p. 390-407.

109. Moroz, L.L., et al., *The ctenophore genome and the evolutionary origins of neural systems.* Nature, 2014. **510**(7503): p. 109-14.

110. Moroz, L.L., *On the independent origins of complex brains and neurons.* Brain Behav Evol, 2009. **74**(3): p. 177-90.

111. Schafer, W., *Nematode nervous systems.* Curr Biol, 2016. **26**(20): p. R955-R959.

112. Enard, W., *The Molecular Basis of Human Brain Evolution.* Curr Biol, 2016. **26**(20): p. R1109-R1117.

113. Verkhratsky, A. and A. Butt, *Neuroglia: Definition, Classification, Evolution, Numbers, Development. In: Glial Physiology and Pathophysiology.* John Wiley & Sons, Ltd, Chichester, UK., 2013.

114. Cajal, S.R.Y., *Textura del Sistema Nervioso del Hombre y de los Vertebrados.* Imprenta N. Moya, 1899.

115. Colas, J.F. and G.C. Schoenwolf, *Towards a cellular and molecular understanding of neurulation.* Dev Dyn, 2001. **221**(2): p. 117-45.

116. Martinez, S., *Molecular Regionalization of the Developing Neural Tube. In: The Mouse Nervous System. Editors: Watson CR, Paxinos G, Puelles L.* Academic Press, 2012: p. 2-18.

117. Azzarelli, R., L.J. Hardwick, and A. Philpott, *Emergence of neuronal diversity from patterning of telencephalic progenitors.* Wiley Interdiscip Rev Dev Biol, 2015. **4**(3): p. 197-214.

118. Borrell, V. and I. Reillo, *Emerging roles of neural stem cells in cerebral cortex development and evolution.* Dev Neurobiol, 2012. **72**(7): p. 955-71.

119. Florio, M. and W.B. Huttner, *Neural progenitors, neurogenesis and the evolution of the neocortex.* Development, 2014. **141**(11): p. 2182-94.

120. Kirkcaldie, M.T.K., *Neocortex. In: The Mouse Nervous System. Editors: Watson CR, Paxinos G, Puelles L.* Academic Press, 2012: p. 52-111.

121. Herculano-Houzel, S., et al., *The basic nonuniformity of the cerebral cortex.* Proc Natl Acad Sci U S A, 2008. **105**(34): p. 12593-8.

122. Nieuwenhuys, R., *The neocortex. An overview of its evolutionary development, structural organization and synaptology.* Anat Embryol (Berl), 1994. **190**(4): p. 307-37.

123. Anderson, S.A., et al., *Distinct origins of neocortical projection neurons and interneurons in vivo.* Cereb Cortex, 2002. **12**(7): p. 702-9.

124. Wichterle, H., et al., *In utero fate mapping reveals distinct migratory pathways and fates of neurons born in the mammalian basal forebrain.* Development, 2001. **128**(19): p. 3759-71.

125. Tan, S.S., et al., *Separate progenitors for radial and tangential cell dispersion during development of the cerebral neocortex.* Neuron, 1998. **21**(2): p. 295-304.

126. Parnavelas, J.G., *The origin and migration of cortical neurones: new vistas.* Trends Neurosci, 2000. **23**(3): p. 126-31.

127. Taverna, E., M. Gotz, and W.B. Huttner, *The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex.* Annu Rev Cell Dev Biol, 2014. **30**: p. 465-502.

128. Evsyukova, I., C. Plestant, and E.S. Anton, *Integrative mechanisms of oriented neuronal migration in the developing brain.* Annu Rev Cell Dev Biol, 2013. **29**: p. 299-353.

129. Marin, O. and J.L. Rubenstein, *A long, remarkable journey: tangential migration in the telencephalon.* Nat Rev Neurosci, 2001. **2**(11): p. 780-90.

130. Marin, O. and J.L. Rubenstein, *Cell migration in the forebrain.* Annu Rev Neurosci, 2003. **26**: p. 441-83.

131. Marin, O., et al., *Guiding neuronal cell migrations.* Cold Spring Harb Perspect Biol, 2010. **2**(2): p. a001834.

132. Campbell, K., *Dorsal-ventral patterning in the mammalian telencephalon.* Curr Opin Neurobiol, 2003. **13**(1): p. 50-6.

133. Custo Greig, L.F., et al., *Molecular logic of neocortical projection neuron specification, development and diversity.* Nat Rev Neurosci, 2013. **14**(11): p. 755-69.

134. Paridaen, J.T. and W.B. Huttner, *Neurogenesis during development of the vertebrate central nervous system.* EMBO Rep, 2014. **15**(4): p. 351-64.

135. Mountcastle, V.B., *The columnar organization of the neocortex.* Brain, 1997. **120 (Pt 4)**: p. 701-22.

136. Gotz, M. and W.B. Huttner, *The cell biology of neurogenesis.* Nat Rev Mol Cell Biol, 2005. **6**(10): p. 777-88.

137. Kriegstein, A.R. and M. Gotz, *Radial glia diversity: a matter of cell fate.* Glia, 2003. **43**(1): p. 37-43.

138. Malatesta, P., et al., *Neuronal or glial progeny: regional differences in radial glia fate.* Neuron, 2003. **37**(5): p. 751-64.

139. Malatesta, P., E. Hartfuss, and M. Gotz, *Isolation of radial glial cells by fluorescent-activated cell sorting reveals a neuronal lineage.* Development, 2000. **127**(24): p. 5253-63.

140. Kriegstein, A. and A. Alvarez-Buylla, *The glial nature of embryonic and adult neural stem cells.* Annu Rev Neurosci, 2009. **32**: p. 149-84.

141. Noctor, S.C., et al., *Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases.* Nat Neurosci, 2004. **7**(2): p. 136-44.

142. Haubensak, W., et al., *Neurons arise in the basal neuroepithelium of the early mammalian telencephalon: a major site of neurogenesis.* Proc Natl Acad Sci U S A, 2004. **101**(9): p. 3196-201.

143. Angevine, J.B., Jr. and R.L. Sidman, *Autoradiographic study of cell migration during histogenesis of cerebral cortex in the mouse.* Nature, 1961. **192**: p. 766-8.

144. Sessa, A., et al., *Tbr2 directs conversion of radial glia into basal precursors and guides neuronal amplification by indirect neurogenesis in the developing neocortex.* Neuron, 2008. **60**(1): p. 56-69.

145. Shitamukai, A., D. Konno, and F. Matsuzaki, *Oblique radial glial divisions in the developing mouse neocortex induce self-renewing progenitors outside the germinal zone that resemble primate outer subventricular zone progenitors.* J Neurosci, 2011. **31**(10): p. 3683-95.

146. Meyer, G., et al., *Different origins and developmental histories of transient neurons in the marginal zone of the fetal and neonatal rat cortex.* J Comp Neurol, 1998. **397**: p. 493–518.

147. Molyneaux, B.J., et al., *Neuronal subtype specification in the cerebral cortex.* Nat Rev Neurosci, 2007. **8**(6): p. 427-37.

148. Englund, C., et al., *Pax6, Tbr2, and Tbr1 are expressed sequentially by radial glia, intermediate progenitor cells, and postmitotic neurons in developing neocortex.* J Neurosci, 2005. **25**(1): p. 247-51.

149. Heins, N., et al., *Glial cells generate neurons: the role of the transcription factor Pax6.* Nat Neurosci, 2002. **5**(4): p. 308-15.

150. Pinto, L. and M. Gotz, *Radial glial cell heterogeneity--the source of diverse progeny in the CNS.* Prog Neurobiol, 2007. **83**(1): p. 2-23.

151. Gotz, M., M. Nakafuku, and D. Petrik, *Neurogenesis in the Developing and Adult Brain-Similarities and Key Differences.* Cold Spring Harb Perspect Biol, 2016. **8**(7).

152. Tramontin, A.D., et al., *Postnatal development of radial glia and the ventricular zone (VZ): a continuum of the neural stem cell compartment.* Cereb Cortex, 2003. **13**(6): p. 580-7.

153. Miyata, T., et al., *Asymmetric production of surface-dividing and non-surface-dividing cortical progenitor cells.* Development, 2004. **131**(13): p. 3133-45.

154. Kowalczyk, T., et al., *Intermediate neuronal progenitors (basal progenitors) produce pyramidal-projection neurons for all layers of cerebral cortex.* Cereb Cortex, 2009. **19**(10): p. 2439-50.

155. Guillemot, F., *Spatial and temporal specification of neural fates by transcription factor codes.* Development, 2007. **134**(21): p. 3771-80.

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

157. Janesick, A., S.C. Wu, and B. Blumberg, *Retinoic acid signaling and neuronal differentiation.* Cell Mol Life Sci, 2015. **72**(8): p. 1559-76.

158. Pierfelice, T., L. Alberi, and N. Gaiano, *Notch in the vertebrate nervous system: an old dog with new tricks.* Neuron, 2011. **69**(5): p. 840-55.

159. Siegenthaler, J.A., et al., *Retinoic acid from the meninges regulates cortical neuron generation.* Cell, 2009. **139**(3): p. 597-609.

160. Zhou, Z.D., et al., *Notch as a molecular switch in neural stem cells.* IUBMB Life, 2010. **62**(8): p. 618-23.

161. Imayoshi, I. and R. Kageyama, *Oscillatory control of bHLH factors in neural progenitors.* Trends Neurosci, 2014. **37**(10): p. 531-8.

162. Kaas, J.H., *The Evolution of Brains from Early Mammals to Humans.* Wiley Interdiscip Rev Cogn Sci, 2013. **4**(1): p. 33-45.

163. Fietz, S.A. and W.B. Huttner, *Cortical progenitor expansion, self-renewal and neurogenesis-a polarized perspective.* Curr Opin Neurobiol, 2011. **21**(1): p. 23-35.

164. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.

165. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* J Clin Invest, 2009. **119**(6): p. 1420-8.

166. Duband, J.L. and J.P. Thiery, *Appearance and distribution of fibronectin during chick embryo gastrulation and neurulation.* Dev Biol, 1982. **94**(2): p. 337-50.

167. Potenta, S., E. Zeisberg, and R. Kalluri, *The role of endothelial-to-mesenchymal transition in cancer progression.* Br J Cancer, 2008. **99**(9): p. 1375-9.

168. Kalluri, R. and E.G. Neilson, *Epithelial-mesenchymal transition and its implications for fibrosis.* J Clin Invest, 2003. **112**(12): p. 1776-84.

169. Nieto, M.A., *Epithelial-Mesenchymal Transitions in development and disease: old views and new perspectives.* Int J Dev Biol, 2009. **53**(8-10): p. 1541-7.

170. Lamouille, S., J. Xu, and R. Derynck, *Molecular mechanisms of epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol, 2014. **15**(3): p. 178-96.

171. Barrallo-Gimeno, A. and M.A. Nieto, *The Snail genes as inducers of cell movement and survival: implications in development and cancer.* Development, 2005. **132**(14): p. 3151-61.

172. Miyazono, K., P. ten Dijke, and C.H. Heldin, *TGF-beta signaling by Smad proteins.* Adv Immunol, 2000. **75**: p. 115-57.

173. Itoh, Y., et al., *Scratch regulates neuronal migration onset via an epithelial-mesenchymal transition-like mechanism.* Nat Neurosci, 2013. **16**(4): p. 416-25.

174. Leong, K.G., et al., *Jagged1-mediated Notch activation induces epithelial-to-mesenchymal transition through Slug-induced repression of E-cadherin.* J Exp Med, 2007. **204**(12): p. 2935-48.

175. Sahlgren, C., et al., *Notch signaling mediates hypoxia-induced tumor cell migration and invasion.* Proc Natl Acad Sci U S A, 2008. **105**(17): p. 6392-7.

176. Timmerman, L.A., et al., *Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation.* Genes Dev, 2004. **18**(1): p. 99-115.

177. Louvi, A., S.S. Sisodia, and E.A. Grove, *Presenilin 1 in migration and morphogenesis in the central nervous system.* Development, 2004. **131**(13): p. 3093-105.

178. Wines-Samuelson, M., M. Handler, and J. Shen, *Role of presenilin-1 in cortical lamination and survival of Cajal-Retzius neurons.* Dev Biol, 2005. **277**(2): p. 332-46.

179. Irvin, D.K., et al., *Expression patterns of Notch1, Notch2, and Notch3 suggest multiple functional roles for the Notch-DSL signaling system during brain development.* J Comp Neurol, 2001. **436**(2): p. 167-81.

180. Solecki, D.J., et al., *Activated Notch2 signaling inhibits differentiation of cerebellar granule neuron precursors by maintaining proliferation.* Neuron, 2001. **31**(4): p. 557-68.

181. Hashimoto-Torii, K., et al., *Interaction between Reelin and Notch signaling regulates neuronal migration in the cerebral cortex.* Neuron, 2008. **60**(2): p. 273-84.

182. Cooper, J.A., *Cell biology in neuroscience: mechanisms of cell migration in the nervous system.* J Cell Biol, 2013. **202**(5): p. 725-34.

183. Huang, Z., *Molecular regulation of neuronal migration during neocortical development.* Mol Cell Neurosci, 2009. **42**(1): p. 11-22.

184. Kwan, K.Y., N. Sestan, and E.S. Anton, *Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex.* Development, 2012. **139**(9): p. 1535-46.

185. Singh, S. and D.J. Solecki, *Polarity transitions during neurogenesis and germinal zone exit in the developing central nervous system.* Front Cell Neurosci, 2015. **9**: p. 62.

186. Charoensawan, V., D. Wilson, and S.A. Teichmann, *Lineage-specific expansion of DNA-binding transcription factor families.* Trends Genet, 2010. **26**(9): p. 388-93.

187. Charoensawan, V., D. Wilson, and S.A. Teichmann, *Genomic repertoires of DNA-binding transcription factors across the tree of life.* Nucleic Acids Res, 2010. **38**(21): p. 7364-77.

188. Vaquerizas, J.M., et al., *A census of human transcription factors: function, expression and evolution.* Nat Rev Genet, 2009. **10**(4): p. 252-63.

189. Zhou, Q., et al., *A mouse tissue transcription factor atlas.* Nat Commun, 2017. **8**: p. 15089.

190. Sainsbury, S., C. Bernecky, and P. Cramer, *Structural basis of transcription initiation by RNA polymerase II.* Nat Rev Mol Cell Biol, 2015. **16**(3): p. 129-43.

191. Naef, F. and J. Huelsken, *Cell-type-specific transcriptomics in chimeric models using transcriptome-based masks.* Nucleic Acids Res, 2005. **33**(13): p. e111.

192. Spitz, F. and E.E. Furlong, *Transcription factors: from enhancer binding to developmental control.* Nat Rev Genet, 2012. **13**(9): p. 613-26.

193. Zhang, W., et al., *The functional landscape of mouse gene expression.* J Biol, 2004. **3**(5): p. 21.

194. Weirauch, M.T. and T.R. Hughes, *A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution.* Subcell Biochem, 2011. **52**: p. 25-73.

195. Guertin, M.J. and J.T. Lis, *Mechanisms by which transcription factors gain access to target sequence elements in chromatin.* Curr Opin Genet Dev, 2013. **23**(2): p. 116-23.

196. Villar, D., P. Flicek, and D.T. Odom, *Evolution of transcription factor binding in metazoans - mechanisms and functional implications.* Nat Rev Genet, 2014. **15**(4): p. 221-33.

197. Jolma, A., et al., *DNA-dependent formation of transcription factor pairs alters their binding specificity.* Nature, 2015. **527**(7578): p. 384-8.

198. Ravasi, T., et al., *An atlas of combinatorial transcriptional regulation in mouse and man.* Cell, 2010. **140**(5): p. 744-52.

199. Teng, L., et al., *Discover context-specific combinatorial transcription factor interactions by integrating diverse ChIP-Seq data sets.* Nucleic Acids Res, 2014. **42**(4): p. e24.

200. Badis, G., et al., *Diversity and complexity in DNA recognition by transcription factors.* Science, 2009. **324**(5935): p. 1720-3.

201. Cusanovich, D.A., et al., *The functional consequences of variation in transcription factor binding.* PLoS Genet, 2014. **10**(3): p. e1004226.

202. Todeschini, A.L., A. Georges, and R.A. Veitia, *Transcription factors: specific DNA binding and specific gene regulation.* Trends Genet, 2014. **30**(6): p. 211-9.

203. Macneil, L.T. and A.J. Walhout, *Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression.* Genome Res, 2011. **21**(5): p. 645-57.

204. Slattery, M., et al., *Absence of a simple code: how transcription factors read the genome.* Trends Biochem Sci, 2014. **39**(9): p. 381-99.

205. Wunderlich, Z. and L.A. Mirny, *Different gene regulation strategies revealed by analysis of binding motifs.* Trends Genet, 2009. **25**(10): p. 434-40.

206. White, M.A., et al., *Massively parallel in vivo enhancer assay reveals that highly local features determine the cis-regulatory function of ChIP-seq peaks.* Proc Natl Acad Sci U S A, 2013. **110**(29): p. 11952-7.

207. Schmidt, D., et al., *Five-vertebrate ChIP-seq reveals the evolutionary dynamics of transcription factor binding.* Science, 2010. **328**(5981): p. 1036-40.

208. Stefflova, K., et al., *Cooperativity and rapid evolution of cobound transcription factors in closely related mammals.* Cell, 2013. **154**(3): p. 530-40.

209. Jones, S., *An overview of the basic helix-loop-helix proteins.* Genome Biol, 2004. **5**(6): p. 226.

210. Littlewood, T.D. and G.I. Evan, *Transcription factors 2: helix-loop-helix.* Protein Profile, 1995. **2**(6): p. 621-702.

211. Amoutzias, G.D., et al., *Choose your partners: dimerization in eukaryotic transcription factors.* Trends Biochem Sci, 2008. **33**(5): p. 220-9.

212. Brownlie, P., et al., *The crystal structure of an intact human Max-DNA complex: new insights into mechanisms of transcriptional control.* Structure, 1997. **5**(4): p. 509-20.

213. Amoutzias, G.D., et al., *Convergent evolution of gene networks by single-gene duplications in higher eukaryotes.* EMBO Rep, 2004. **5**(3): p. 274-9.

214. Ledent, V. and M. Vervoort, *The basic helix-loop-helix protein family: comparative genomics and phylogenetic analysis.* Genome Res, 2001. **11**(5): p. 754-70.

215. Luscher, B., *Function and regulation of the transcription factors of the Myc/Max/Mad network.* Gene, 2001. **277**(1-2): p. 1-14.

216. Massari, M.E. and C. Murre, *Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms.* Mol Cell Biol, 2000. **20**(2): p. 429-40.

217. Imayoshi, I. and R. Kageyama, *bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells.* Neuron, 2014. **82**(1): p. 9-23.

218. Rudnicki, M.A. and R. Jaenisch, *The MyoD family of transcription factors and skeletal myogenesis.* Bioessays, 1995. **17**(3): p. 203-9.

219. Lai, H.C., D.M. Meredith, and J.E. Johnson, *bHLH factors in neurogenesis and neuronal subtype specification. In: Patterning and Cell Type Specification in the Developing CNS and PNS. Editors: Rubenstein JLR and Rakic P.* Elsevier, 2013: p. 333–354.

220. Longo, A., G.P. Guanga, and R.B. Rose, *Crystal structure of E47-NeuroD1/beta2 bHLH domain-DNA complex: heterodimer selectivity and DNA recognition.* Biochemistry, 2008. **47**(1): p. 218-29.

221. Cho, J.H. and M.J. Tsai, *The role of BETA2/NeuroD1 in the development of the nervous system.* Mol Neurobiol, 2004. **30**(1): p. 35-47.

222. Emerson, R.O. and J.H. Thomas, *Adaptive evolution in zinc finger transcription factors.* PLoS Genet, 2009. **5**(1): p. e1000325.

223. Stubbs, L., Y. Sun, and D. Caetano-Anolles, *Function and Evolution of C2H2 Zinc Finger Arrays.* Subcell Biochem, 2011. **52**: p. 75-94.

224. Jen, J. and Y.C. Wang, *Zinc finger proteins in cancer progression.* J Biomed Sci, 2016. **23**(1): p. 53.

225. Miller, J., A.D. McLachlan, and A. Klug, *Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes.* EMBO J, 1985. **4**(6): p. 1609-14.

226. Klug, A., *The discovery of zinc fingers and their applications in gene regulation and genome manipulation.* Annu Rev Biochem, 2010. **79**: p. 213-31.

227. Laity, J.H., B.M. Lee, and P.E. Wright, *Zinc finger proteins: new insights into structural and functional diversity.* Curr Opin Struct Biol, 2001. **11**(1): p. 39-46.

228. Wolfe, S.A., L. Nekludova, and C.O. Pabo, *DNA recognition by Cys2His2 zinc finger proteins.* Annu Rev Biophys Biomol Struct, 2000. **29**: p. 183-212.

229. Najafabadi, H.S., et al., *C2H2 zinc finger proteins greatly expand the human regulatory lexicon.* Nat Biotechnol, 2015. **33**(5): p. 555-62.

230. Mackay, J.P. and M. Crossley, *Zinc fingers are sticking together.* Trends Biochem Sci, 1998. **23**(1): p. 1-4.

231. Brayer, K.J. and D.J. Segal, *Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains.* Cell Biochem Biophys, 2008. **50**(3): p. 111-31.

232. Brayer, K.J., S. Kulshreshtha, and D.J. Segal, *The protein-binding potential of C2H2 zinc finger domains.* Cell Biochem Biophys, 2008. **51**(1): p. 9-19.

233. Iuchi, S., *Three classes of C2H2 zinc finger proteins.* Cell Mol Life Sci, 2001. **58**(4): p. 625-35.

234. Pavletich, N.P. and C.O. Pabo, *Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A.* Science, 1991. **252**(5007): p. 809-17.

235. Urrutia, R., *KRAB-containing zinc-finger repressor proteins.* Genome Biol, 2003. **4**(10): p. 231.

236. Thomas, J.H. and S. Schneider, *Coevolution of retroelements and tandem zinc finger genes.* Genome Res, 2011. **21**(11): p. 1800-12.

237. Birtle, Z. and C.P. Ponting, *Meisetz and the birth of the KRAB motif.* Bioinformatics, 2006. **22**(23): p. 2841-5.

238. Bellefroid, E.J., et al., *The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins.* Proc Natl Acad Sci U S A, 1991. **88**(9): p. 3608-12.

239. Margolin, J.F., et al., *Kruppel-associated boxes are potent transcriptional repression domains.* Proc Natl Acad Sci U S A, 1994. **91**(10): p. 4509-13.

240. Witzgall, R., et al., *The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression.* Proc Natl Acad Sci U S A, 1994. **91**(10): p. 4514-8.

241. Friedman, J.R., et al., *KAP-1, a novel corepressor for the highly conserved KRAB repression domain.* Genes Dev, 1996. **10**(16): p. 2067-78.

242. Schultz, D.C., et al., *SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins.* Genes Dev, 2002. **16**(8): p. 919-32.

243. Schultz, D.C., J.R. Friedman, and F.J. Rauscher, 3rd, *Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD.* Genes Dev, 2001. **15**(4): p. 428-43.

244. Wolf, G., D. Greenberg, and T.S. Macfarlan, *Spotting the enemy within: Targeted silencing of foreign DNA in mammalian genomes by the Kruppel-associated box zinc finger protein family.* Mob DNA, 2015. **6**: p. 17.

245. Wolf, D. and S.P. Goff, *Embryonic stem cells use ZFP809 to silence retroviral DNAs.* Nature, 2009. **458**(7242): p. 1201-4.

246. Wolf, G., et al., *The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses.* Genes Dev, 2015. **29**(5): p. 538-54.

247. Friedli, M. and D. Trono, *The developmental control of transposable elements and the evolution of higher species.* Annu Rev Cell Dev Biol, 2015. **31**: p. 429-51.

248. Feschotte, C. and C. Gilbert, *Endogenous viruses: insights into viral evolution and impact on host biology.* Nat Rev Genet, 2012. **13**(4): p. 283-96.

249. Rowe, H.M. and D. Trono, *Dynamic control of endogenous retroviruses during development.* Virology, 2011. **411**(2): p. 273-87.

250. Matsui, T., et al., *Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET.* Nature, 2010. **464**(7290): p. 927-31.

251. Rowe, H.M., et al., *KAP1 controls endogenous retroviruses in embryonic stem cells.* Nature, 2010. **463**(7278): p. 237-40.

252. Turelli, P., et al., *Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements.* Genome Res, 2014. **24**(8): p. 1260-70.

253. Barde, I., et al., *A KRAB/KAP1-miRNA cascade regulates erythropoiesis through stage-specific control of mitophagy.* Science, 2013. **340**(6130): p. 350-3.

254. Corsinotti, A., et al., *Global and stage specific patterns of Kruppel-associated-box zinc finger protein gene expression in murine early embryonic cells.* PLoS One, 2013. **8**(2): p. e56721.

255. Lizio, M., et al., *Gateways to the FANTOM5 promoter level mammalian expression atlas.* Genome Biol, 2015. **16**: p. 22.

256. Ecco, G., et al., *Transposable Elements and Their KRAB-ZFP Controllers Regulate Gene Expression in Adult Tissues.* Dev Cell, 2016. **36**(6): p. 611-23.

257. Huntley, S., et al., *A comprehensive catalog of human KRAB-associated zinc finger genes: insights into the evolutionary history of a large family of transcriptional repressors.* Genome Res, 2006. **16**(5): p. 669-77.

258. Nowick, K., et al., *Rapid sequence and expression divergence suggest selection for novel function in primate-specific KRAB-ZNF genes.* Mol Biol Evol, 2010. **27**(11): p. 2606-17.

259. Mouse Genome Sequencing, C., et al., *Initial sequencing and comparative analysis of the mouse genome.* Nature, 2002. **420**(6915): p. 520-62.

260. Guenet, J.L., *The mouse genome.* Genome Res, 2005. **15**(12): p. 1729-40.

261. Gnerre, S., et al., *High-quality draft assemblies of mammalian genomes from massively parallel sequence data.* Proc Natl Acad Sci U S A, 2011. **108**(4): p. 1513-8.

262. Adams, D.J., et al., *The Mouse Genomes Project: a repository of inbred laboratory mouse strain genomes.* Mamm Genome, 2015. **26**(9-10): p. 403-12.

263. Yue, F., et al., *A comparative encyclopedia of DNA elements in the mouse genome.* Nature, 2014. **515**(7527): p. 355-64.

264. Levo, M. and E. Segal, *In pursuit of design principles of regulatory sequences.* Nat Rev Genet, 2014. **15**(7): p. 453-68.

265. Shen, Y., et al., *A map of the cis-regulatory sequences in the mouse genome.* Nature, 2012. **488**(7409): p. 116-20.

266. De Santa, F., et al., *A large fraction of extragenic RNA pol II transcription sites overlap enhancers.* PLoS Biol, 2010. **8**(5): p. e1000384.

267. Kim, T.K. and R. Shiekhattar, *Architectural and Functional Commonalities between Enhancers and Promoters.* Cell, 2015. **162**(5): p. 948-59.

268. Li, W., M.T. Lam, and D. Notani, *Enhancer RNAs.* Cell Cycle, 2014. **13**(20): p. 3151-2.

269. Lam, M.T., et al., *Enhancer RNAs and regulated transcriptional programs.* Trends Biochem Sci, 2014. **39**(4): p. 170-82.

270. Natoli, G. and J.C. Andrau, *Noncoding transcription at enhancers: general principles and functional models.* Annu Rev Genet, 2012. **46**: p. 1-19.

271. Vernimmen, D. and W.A. Bickmore, *The Hierarchy of Transcriptional Activation: From Enhancer to Promoter.* Trends Genet, 2015. **31**(12): p. 696-708.

272. Zabidi, M.A., et al., *Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation.* Nature, 2015. **518**(7540): p. 556-9.

273. Zabidi, M.A. and A. Stark, *Regulatory Enhancer-Core-Promoter Communication via Transcription Factors and Cofactors.* Trends Genet, 2016. **32**(12): p. 801-814.

274. Stergachis, A.B., et al., *Conservation of trans-acting circuitry during mammalian regulatory evolution.* Nature, 2014. **515**(7527): p. 365-70.

275. Cheng, Y., et al., *Principles of regulatory information conservation between mouse and human.* Nature, 2014. **515**(7527): p. 371-375.

276. Pope, B.D., et al., *Topologically associating domains are stable units of replication-timing regulation.* Nature, 2014. **515**(7527): p. 402-5.

277. Carninci, P., *Genomics: mice in the ENCODE spotlight.* Nature, 2014. **515**(7527): p. 346-7.

278. Vierstra, J., et al., *Mouse regulatory DNA landscapes reveal global principles of cis-regulatory evolution.* Science, 2014. **346**(6212): p. 1007-12.

279. Rada-Iglesias, A., et al., *A unique chromatin signature uncovers early developmental enhancers in humans.* Nature, 2011. **470**(7333): p. 279-83.

280. Thurman, R.E., et al., *The accessible chromatin landscape of the human genome.* Nature, 2012. **489**(7414): p. 75-82.

281. Muller, F. and L. Tora, *Chromatin and DNA sequences in defining promoters for transcription initiation.* Biochim Biophys Acta, 2014. **1839**(3): p. 118-28.

282. Vierstra, J., et al., *Coupling transcription factor occupancy to nucleosome architecture with DNase-FLASH.* Nat Methods, 2014. **11**(1): p. 66-72.

283. Andersson, R., *Promoter or enhancer, what's the difference? Deconstruction of established distinctions and presentation of a unifying model.* Bioessays, 2015. **37**(3): p. 314-23.

284. Andersson, R., A. Sandelin, and C.G. Danko, *A unified architecture of transcriptional regulatory elements.* Trends Genet, 2015. **31**(8): p. 426-33.

285. Kowalczyk, M.S., et al., *Intragenic enhancers act as alternative promoters.* Mol Cell, 2012. **45**(4): p. 447-58.

286. Biemont, C., *A brief history of the status of transposable elements: from junk DNA to major players in evolution.* Genetics, 2010. **186**(4): p. 1085-93.

287. Elbarbary, R.A., B.A. Lucas, and L.E. Maquat, *Retrotransposons as regulators of gene expression.* Science, 2016. **351**(6274): p. aac7247.

288. Deveson, I.W., et al., *The Dimensions, Dynamics, and Relevance of the Mammalian Noncoding Transcriptome.* Trends Genet, 2017. **33**(7): p. 464-478.

289. Gaiti, F., et al., *Origin and evolution of the metazoan non-coding regulatory genome.* Dev Biol, 2017. **427**(2): p. 193-202.

290. Smale, S.T. and J.T. Kadonaga, *The RNA polymerase II core promoter.* Annu Rev Biochem, 2003. **72**: p. 449-79.

291. Lenhard, B., A. Sandelin, and P. Carninci, *Metazoan promoters: emerging characteristics and insights into transcriptional regulation.* Nat Rev Genet, 2012. **13**(4): p. 233-45.

292. Juven-Gershon, T. and J.T. Kadonaga, *Regulation of gene expression via the core promoter and the basal transcriptional machinery.* Dev Biol, 2010. **339**(2): p. 225-9.

293. Haberle, V. and B. Lenhard, *Promoter architectures and developmental gene regulation.* Semin Cell Dev Biol, 2016. **57**: p. 11-23.

294. Carninci, P., et al., *Genome-wide analysis of mammalian promoter architecture and evolution.* Nat Genet, 2006. **38**(6): p. 626-35.

295. Cooper, S.J., et al., *Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome.* Genome Res, 2006. **16**(1): p. 1-10.

296. Yang, C., et al., *Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters.* Gene, 2007. **389**(1): p. 52-65.

297. Sandelin, A., et al., *Mammalian RNA polymerase II core promoters: insights from genome-wide studies.* Nat Rev Genet, 2007. **8**(6): p. 424-36.

298. Valen, E. and A. Sandelin, *Genomic and chromatin signals underlying transcription start-site selection.* Trends Genet, 2011. **27**(11): p. 475-85.

299. Akbari, O.S., et al., *A novel promoter-tethering element regulates enhancer-driven gene expression at the bithorax complex in the Drosophila embryo.* Development, 2008. **135**(1): p. 123-31.

300. Butler, J.E. and J.T. Kadonaga, *Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs.* Genes Dev, 2001. **15**(19): p. 2515-9.

301. Aerts, S., *Computational strategies for the genome-wide identification of cis-regulatory elements and transcriptional targets.* Curr Top Dev Biol, 2012. **98**: p. 121-45.

302. Visel, A., et al., *ChIP-seq accurately predicts tissue-specific activity of enhancers.* Nature, 2009. **457**(7231): p. 854-8.

303. Ong, C.T. and V.G. Corces, *Enhancer function: new insights into the regulation of tissue-specific gene expression.* Nat Rev Genet, 2011. **12**(4): p. 283-93.

304. Banerji, J., S. Rusconi, and W. Schaffner, *Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences.* Cell, 1981. **27**(2 Pt 1): p. 299-308.

305. Consortium, E.P., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. **489**(7414): p. 57-74.

306. Schaffner, W., *Enhancers, enhancers - from their discovery to today's universe of transcription enhancers.* Biol Chem, 2015. **396**(4): p. 311-27.

307. Ghanem, N., et al., *Distinct cis-regulatory elements from the Dlx1/Dlx2 locus mark different progenitor cell populations in the ganglionic eminences and different subtypes of adult cortical interneurons.* J Neurosci, 2007. **27**(19): p. 5012-22.

308. Visel, A., et al., *A high-resolution enhancer atlas of the developing telencephalon.* Cell, 2013. **152**(4): p. 895-908.

309. Zerucha, T., et al., *A highly conserved enhancer in the Dlx5/Dlx6 intergenic region is the site of cross-regulatory interactions between Dlx genes in the embryonic forebrain.* J Neurosci, 2000. **20**(2): p. 709-21.

310. Panne, D., *The enhanceosome.* Curr Opin Struct Biol, 2008. **18**(2): p. 236-42.

311. Buecker, C. and J. Wysocka, *Enhancers as information integration hubs in development: lessons from genomics.* Trends Genet, 2012. **28**(6): p. 276-84.

312. Long, H.K., S.L. Prescott, and J. Wysocka, *Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution.* Cell, 2016. **167**(5): p. 1170-1187.

313. Bender, M.A., et al., *The hypersensitive sites of the murine beta-globin locus control region act independently to affect nuclear localization and transcriptional elongation.* Blood, 2012. **119**(16): p. 3820-7.

314. Sawado, T., et al., *The beta -globin locus control region (LCR) functions primarily by enhancing the transition from transcription initiation to elongation.* Genes Dev, 2003. **17**(8): p. 1009-18.

315. Schaukowitch, K., et al., *Enhancer RNA facilitates NELF release from immediate early genes.* Mol Cell, 2014. **56**(1): p. 29-42.

316. Spicuglia, S., et al., *Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes.* Mol Cell, 2002. **10**(6): p. 1479-87.

317. Ho, Y., et al., *A defined locus control region determinant links chromatin domain acetylation with long-range gene activation.* Mol Cell, 2002. **9**(2): p. 291-302.

318. Ho, Y., et al., *Locus control region transcription plays an active role in long-range gene activation.* Mol Cell, 2006. **23**(3): p. 365-75.

319. Lin, C., et al., *The RNA Pol II elongation factor Ell3 marks enhancers in ES cells and primes future gene activation.* Cell, 2013. **152**(1-2): p. 144-56.

320. Vernimmen, D., *Uncovering enhancer functions using the alpha-globin locus.* PLoS Genet, 2014. **10**(10): p. e1004668.

321. Vernimmen, D., et al., *Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression.* EMBO J, 2007. **26**(8): p. 2041-51.

322. Vernimmen, D., et al., *Chromosome looping at the human alpha-globin locus is mediated via the major upstream regulatory element (HS -40).* Blood, 2009. **114**(19): p. 4253-60.

323. Guo, Y., et al., *CTCF/cohesin-mediated DNA looping is required for protocadherin alpha promoter choice.* Proc Natl Acad Sci U S A, 2012. **109**(51): p. 21081-6.

324. Kagey, M.H., et al., *Mediator and cohesin connect gene expression and chromatin architecture.* Nature, 2010. **467**(7314): p. 430-5.

325. Plank, J.L. and A. Dean, *Enhancer function: mechanistic and genome-wide insights come together.* Mol Cell, 2014. **55**(1): p. 5-14.

326. Stadhouders, R., et al., *Transcription regulation by distal enhancers: who's in the loop?* Transcription, 2012. **3**(4): p. 181-6.

327. Zhang, Y., et al., *Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations.* Nature, 2013. **504**(7479): p. 306-310.

328. DeMare, L.E., et al., *The genomic landscape of cohesin-associated chromatin interactions.* Genome Res, 2013. **23**(8): p. 1224-34.

329. Kieffer-Kwon, K.R., et al., *Interactome maps of mouse gene regulatory domains reveal basic principles of transcriptional regulation.* Cell, 2013. **155**(7): p. 1507-20.

330. Daniel, B., G. Nagy, and L. Nagy, *The intriguing complexities of mammalian gene regulation: how to link enhancers to regulated genes. Are we there yet?* FEBS Lett, 2014. **588**(15): p. 2379-91.

331. Visel, A., et al., *VISTA Enhancer Browser--a database of tissue-specific human enhancers.* Nucleic Acids Res, 2007. **35**(Database issue): p. D88-92.

332. de Laat, W. and D. Duboule, *Topology of mammalian developmental enhancers and their regulatory landscapes.* Nature, 2013. **502**(7472): p. 499-506.

333. Maston, G.A., et al., *Characterization of enhancer function from genome-wide analyses.* Annu Rev Genomics Hum Genet, 2012. **13**: p. 29-57.

334. Alexander, R.P., et al., *Annotating non-coding regions of the genome.* Nat Rev Genet, 2010. **11**(8): p. 559-71.

335. Castillo-Davis, C.I., *The evolution of noncoding DNA: how much junk, how much func?* Trends Genet, 2005. **21**(10): p. 533-6.

336. Cavalier-Smith, T., *The Evolution of Genome Size.* New York: John Wiley and Sons, 1985.

337. Gregory, T.R., *The bigger the C-value, the larger the cell: genome size and red blood cell size in vertebrates.* Blood Cells Mol Dis, 2001. **27**(5): p. 830-43.

338. Gregory, T.R., *Genome size and developmental complexity.* Genetica, 2002. **115**(1): p. 131-46.

339. Patrushev, L.I. and I.G. Minkevich, *The problem of the eukaryotic genome size.* Biochemistry (Mosc), 2008. **73**(13): p. 1519-52.

340. Liu, G., J.S. Mattick, and R.J. Taft, *A meta-analysis of the genomic and transcriptomic composition of complex life.* Cell Cycle, 2013. **12**(13): p. 2061-72.

341. Kidwell, M.G., *Transposable elements and the evolution of genome size in eukaryotes.* Genetica, 2002. **115**(1): p. 49-63.

342. Chenais, B., et al., *The impact of transposable elements on eukaryotic genomes: from genome size increase to genetic adaptation to stressful environments.* Gene, 2012. **509**(1): p. 7-15.

343. Habibi, L., et al., *Mobile DNA Elements: The Seeds of Organic Complexity on Earth.* DNA Cell Biol, 2015. **34**(10): p. 597-609.

344. Nelson, C.E., B.M. Hersh, and S.B. Carroll, *The regulatory content of intergenic DNA shapes genome architecture.* Genome Biol, 2004. **5**(4): p. R25.

345. Reichmann, J., et al., *Microarray analysis of LTR retrotransposon silencing identifies Hdac1 as a regulator of retrotransposon expression in mouse embryonic stem cells.* PLoS Comput Biol, 2012. **8**(4): p. e1002486.

346. Nellaker, C., et al., *The genomic landscape shaped by selection on transposable elements across 18 mouse strains.* Genome Biol, 2012. **13**(6): p. R45.

347. Levin, H.L. and J.V. Moran, *Dynamic interactions between transposable elements and their hosts.* Nat Rev Genet, 2011. **12**(9): p. 615-27.

348. Munoz-Lopez, M., et al., *Study of Transposable Elements and Their Genomic Impact.* Methods Mol Biol, 2016. **1400**: p. 1-19.

349. Chuong, E.B., N.C. Elde, and C. Feschotte, *Regulatory activities of transposable elements: from conflicts to benefits.* Nat Rev Genet, 2017. **18**(2): p. 71-86.

350. Feschotte, C., *Transposable elements and the evolution of regulatory networks.* Nat Rev Genet, 2008. **9**(5): p. 397-405.

351. Kidwell, M.G. and D.R. Lisch, *Perspective: transposable elements, parasitic DNA, and genome evolution.* Evolution, 2001. **55**(1): p. 1-24.

352. Rebollo, R., M.T. Romanish, and D.L. Mager, *Transposable elements: an abundant and natural source of regulatory sequences for host genes.* Annu Rev Genet, 2012. **46**: p. 21-42.

353. Richardson, S.R., et al., *The Influence of LINE-1 and SINE Retrotransposons on Mammalian Genomes.* Microbiol Spectr, 2015. **3**(2): p. MDNA3-0061-2014.

354. Mager, D.L. and J.P. Stoye, *Mammalian Endogenous Retroviruses.* Microbiol Spectr, 2015. **3**(1): p. MDNA3-0009-2014.

355. Stocking, C. and C.A. Kozak, *Murine endogenous retroviruses.* Cell Mol Life Sci, 2008. **65**(21): p. 3383-98.

356. Thompson, P.J., T.S. Macfarlan, and M.C. Lorincz, *Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire.* Mol Cell, 2016. **62**(5): p. 766-76.

357. Sundaram, V., et al., *Widespread contribution of transposable elements to the innovation of gene regulatory networks.* Genome Res, 2014. **24**(12): p. 1963-76.

358. Jacques, P.E., J. Jeyakani, and G. Bourque, *The majority of primate-specific regulatory sequences are derived from transposable elements.* PLoS Genet, 2013. **9**(5): p. e1003504.

359. Bourque, G., et al., *Evolution of the mammalian transcription factor binding repertoire via transposable elements.* Genome Res, 2008. **18**(11): p. 1752-62.

360. Kunarso, G., et al., *Transposable elements have rewired the core regulatory network of human embryonic stem cells.* Nat Genet, 2010. **42**(7): p. 631-4.

361. Faulkner, G.J., et al., *The regulated retrotransposon transcriptome of mammalian cells.* Nat Genet, 2009. **41**(5): p. 563-71.

362. Chuong, E.B., et al., *Endogenous retroviruses function as species-specific enhancer elements in the placenta.* Nat Genet, 2013. **45**(3): p. 325-9.

363. Gaudet, F., et al., *Dnmt1 expression in pre- and postimplantation embryogenesis and the maintenance of IAP silencing.* Mol Cell Biol, 2004. **24**(4): p. 1640-8.

364. Dewannieux, M. and T. Heidmann, *Endogenous retroviruses: acquisition, amplification and taming of genome invaders.* Curr Opin Virol, 2013. **3**(6): p. 646-56.

365. Holmes, E.C., *The evolution of endogenous viral elements.* Cell Host Microbe, 2011. **10**(4): p. 368-77.

366. Stoye, J.P., *Studies of endogenous retroviruses reveal a continuing evolutionary saga.* Nat Rev Microbiol, 2012. **10**(6): p. 395-406.

367. Maksakova, I.A., et al., *Retroviral elements and their hosts: insertional mutagenesis in the mouse germ line.* PLoS Genet, 2006. **2**(1): p. e2.

368. Rowe, H.M., et al., *De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET.* Development, 2013. **140**(3): p. 519-29.

369. Ecco, G., M. Imbeault, and D. Trono, *KRAB zinc finger proteins.* Development, 2017. **144**(15): p. 2719-2729.

370. Crick, F., *Central dogma of molecular biology.* Nature, 1970. **227**(5258): p. 561-3.

371. Crick, F.H., *On protein synthesis.* Symp Soc Exp Biol, 1958. **12**: p. 138-63.

372. Morris, K.V. and J.S. Mattick, *The rise of regulatory RNA.* Nat Rev Genet, 2014. **15**(6): p. 423-37.

373. Cech, T.R. and J.A. Steitz, *The noncoding RNA revolution-trashing old rules to forge new ones.* Cell, 2014. **157**(1): p. 77-94.

374. Carter, R. and G. Drouin, *Structural differentiation of the three eukaryotic RNA polymerases.* Genomics, 2009. **94**(6): p. 388-96.

375. Vannini, A. and P. Cramer, *Conservation between the RNA polymerase I, II, and III transcription initiation machineries.* Mol Cell, 2012. **45**(4): p. 439-46.

376. Darnell, J.E., Jr., *Reflections on the history of pre-mRNA processing and highlights of current knowledge: a unified picture.* RNA, 2013. **19**(4): p. 443-60.

377. Mignone, F., et al., *Untranslated regions of mRNAs.* Genome Biol, 2002. **3**(3): p. REVIEWS0004.

378. Matoulkova, E., et al., *The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells.* RNA Biol, 2012. **9**(5): p. 563-76.

379. Lee, Y. and D.C. Rio, *Mechanisms and Regulation of Alternative Pre-mRNA Splicing.* Annu Rev Biochem, 2015. **84**: p. 291-323.

380. Allis, C.D., et al., *Epigenetics, 2nd edn.* Cold Spring Harbor Laboratory Press, 2015.

381. Coulon, A., et al., *Eukaryotic transcriptional dynamics: from single molecules to cell populations.* Nat Rev Genet, 2013. **14**(8): p. 572-84.

382. Levine, M., C. Cattoglio, and R. Tjian, *Looping back to leap forward: transcription enters a new era.* Cell, 2014. **157**(1): p. 13-25.

383. Weake, V.M. and J.L. Workman, *Inducible gene expression: diverse regulatory mechanisms.* Nat Rev Genet, 2010. **11**(6): p. 426-37.

384. Roeder, R.G., *Transcriptional regulation and the role of diverse coactivators in animal cells.* FEBS Lett, 2005. **579**(4): p. 909-15.

385. Farnham, P.J., *Insights from genomic profiling of transcription factors.* Nat Rev Genet, 2009. **10**(9): p. 605-16.

386. Koch, F., et al., *Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters.* Nat Struct Mol Biol, 2011. **18**(8): p. 956-63.

387. Zentner, G.E., P.J. Tesar, and P.C. Scacheri, *Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions.* Genome Res, 2011. **21**(8): p. 1273-83.

388. Zhang, W., et al., *Bromodomain-containing protein 4 (BRD4) regulates RNA polymerase II serine 2 phosphorylation in human CD4+ T cells.* J Biol Chem, 2012. **287**(51): p. 43137-55.

389. Kwak, H. and J.T. Lis, *Control of transcriptional elongation.* Annu Rev Genet, 2013. **47**: p. 483-508.

390. Iwafuchi-Doi, M. and K.S. Zaret, *Pioneer transcription factors in cell reprogramming.* Genes Dev, 2014. **28**(24): p. 2679-92.

391. Zaret, K.S. and J.S. Carroll, *Pioneer transcription factors: establishing competence for gene expression.* Genes Dev, 2011. **25**(21): p. 2227-41.

392. Edelman, L.B. and P. Fraser, *Transcription factories: genetic programming in three dimensions.* Curr Opin Genet Dev, 2012. **22**(2): p. 110-4.

393. Yosef, N. and A. Regev, *Impulse control: temporal dynamics in gene transcription.* Cell, 2011. **144**(6): p. 886-96.

394. Lelli, K.M., M. Slattery, and R.S. Mann, *Disentangling the many layers of eukaryotic transcriptional regulation.* Annu Rev Genet, 2012. **46**: p. 43-68.

395. Waddington, C.H., *Canalization of development and the inheritance of acquired characters.* Nature, 1942. **150**: p. 563–565.

396. Waddington, C.H., *The epigenotype.* Endeavour, 1942. **1**: p. 18–20.

397. Waddington, C.H., *The Strategy of the Genes; a Discussion of Some Aspects of Theoretical Biology.* Allen & Unwin, 1957.

398. Stricker, S.H., A. Koferle, and S. Beck, *From profiles to function in epigenomics.* Nat Rev Genet, 2017. **18**(1): p. 51-66.

399. Berger, S.L., et al., *An operational definition of epigenetics.* Genes Dev, 2009. **23**(7): p. 781-3.

400. Deans, C. and K.A. Maggert, *What do you mean, "epigenetic"?* Genetics, 2015. **199**(4): p. 887-96.

401. Johnson, D.G. and S.Y. Dent, *Chromatin: receiver and quarterback for cellular signals.* Cell, 2013. **152**(4): p. 685-9.

402. Lawrence, M., S. Daujat, and R. Schneider, *Lateral Thinking: How Histone Modifications Regulate Gene Expression.* Trends Genet, 2016. **32**(1): p. 42-56.

403. Mueller-Planitz, F., H. Klinker, and P.B. Becker, *Nucleosome sliding mechanisms: new twists in a looped history.* Nat Struct Mol Biol, 2013. **20**(9): p. 1026-32.

404. Narlikar, G.J., R. Sundaramoorthy, and T. Owen-Hughes, *Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes.* Cell, 2013. **154**(3): p. 490-503.

405. Heitz, E., *Das Heterochromatin der Moose: I* Jahrbücher für wissenschaftliche Botanik., 1928. **69**: p. 762–818.

406. Bentley, G.A., et al., *Crystal structure of the nucleosome core particle at 16 A resolution.* J Mol Biol, 1984. **176**(1): p. 55-75.

407. Kornberg, R.D. and J.O. Thomas, *Chromatin structure; oligomers of the histones.* Science, 1974. **184**(4139): p. 865-8.

408. Luger, K., et al., *Crystal structure of the nucleosome core particle at 2.8 A resolution.* Nature, 1997. **389**(6648): p. 251-60.

409. Olins, D.E. and A.L. Olins, *Chromatin history: our view from the bridge.* Nat Rev Mol Cell Biol, 2003. **4**(10): p. 809-14.

410. Allan, J., et al., *The structure of histone H1 and its location in chromatin.* Nature, 1980. **288**(5792): p. 675-9.

411. Kornberg, R.D., *Chromatin structure: a repeating unit of histones and DNA.* Science, 1974. **184**(4139): p. 868-71.

412. Luger, K., M.L. Dechassa, and D.J. Tremethick, *New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?* Nat Rev Mol Cell Biol, 2012. **13**(7): p. 436-47.

413. Andrews, A.J., et al., *The histone chaperone Nap1 promotes nucleosome assembly by eliminating nonnucleosomal histone DNA interactions.* Mol Cell, 2010. **37**(6): p. 834-42.

414. Lee, D.Y., et al., *A positive role for histone acetylation in transcription factor access to nucleosomal DNA.* Cell, 1993. **72**(1): p. 73-84.

415. Wasylyk, B. and P. Chambon, *Transcription by eukaryotic RNA polymerases A and B of chromatin assembled in vitro.* Eur J Biochem, 1979. **98**(2): p. 317-27.

416. Teves, S.S., C.M. Weber, and S. Henikoff, *Transcribing through the nucleosome.* Trends Biochem Sci, 2014. **39**(12): p. 577-86.

417. Bickmore, W.A. and B. van Steensel, *Genome architecture: domain organization of interphase chromosomes.* Cell, 2013. **152**(6): p. 1270-84.

418. Dekker, J. and T. Misteli, *Long-Range Chromatin Interactions.* Cold Spring Harb Perspect Biol, 2015. **7**(10): p. a019356.

419. Sexton, T. and G. Cavalli, *The role of chromosome domains in shaping the functional genome.* Cell, 2015. **160**(6): p. 1049-59.

420. Tang, Z., et al., *CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription.* Cell, 2015. **163**(7): p. 1611-27.

421. Hotchkiss, R.D., *The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography.* J Biol Chem, 1948. **175**(1): p. 315-32.

422. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.

423. Bird, A., et al., *A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA.* Cell, 1985. **40**(1): p. 91-9.

424. Holliday, R. and J.E. Pugh, *DNA modification mechanisms and gene activity during development.* Science, 1975. **187**(4173): p. 226-32.

425. Razin, A. and A.D. Riggs, *DNA methylation and gene function.* Science, 1980. **210**(4470): p. 604-10.

426. Riggs, A.D., *X inactivation, differentiation, and DNA methylation.* Cytogenet Cell Genet, 1975. **14**(1): p. 9-25.

427. Jones, P.A., *Functions of DNA methylation: islands, start sites, gene bodies and beyond.* Nat Rev Genet, 2012. **13**(7): p. 484-92.

428. Goll, M.G. and T.H. Bestor, *Eukaryotic cytosine methyltransferases.* Annu Rev Biochem, 2005. **74**: p. 481-514.

429. Jurkowska, R.Z., T.P. Jurkowski, and A. Jeltsch, *Structure and function of mammalian DNA methyltransferases.* Chembiochem, 2011. **12**(2): p. 206-22.

430. Okano, M., S. Xie, and E. Li, *Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases.* Nat Genet, 1998. **19**(3): p. 219-20.

431. Law, J.A. and S.E. Jacobsen, *Establishing, maintaining and modifying DNA methylation patterns in plants and animals.* Nat Rev Genet, 2010. **11**(3): p. 204-20.

432. Smith, Z.D. and A. Meissner, *DNA methylation: roles in mammalian development.* Nat Rev Genet, 2013. **14**(3): p. 204-20.

433. Bird, A., *CpG-rich islands and the function of DNA methylation.* Nature, 1986. **321**(6067): p. 209-13.

434. Doskocil, J. and F. Sorm, *Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids.* Biochim Biophys Acta, 1962. **55**(6): p. 953-959.

435. Bergman, Y. and H. Cedar, *DNA methylation dynamics in health and disease.* Nat Struct Mol Biol, 2013. **20**(3): p. 274-81.

436. Deaton, A.M. and A. Bird, *CpG islands and the regulation of transcription.* Genes Dev, 2011. **25**(10): p. 1010-22.

437. Shukla, S., et al., *CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing.* Nature, 2011. **479**(7371): p. 74-9.

438. Spruijt, C.G. and M. Vermeulen, *DNA methylation: old dog, new tricks?* Nat Struct Mol Biol, 2014. **21**(11): p. 949-54.

439. Tahiliani, M., et al., *Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1.* Science, 2009. **324**(5929): p. 930-5.

440. Ito, S., et al., *Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification.* Nature, 2010. **466**(7310): p. 1129-33.

441. Ito, S., et al., *Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine.* Science, 2011. **333**(6047): p. 1300-3.

442. He, Y.F., et al., *Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA.* Science, 2011. **333**(6047): p. 1303-7.

443. Wu, H. and Y. Zhang, *Reversing DNA methylation: mechanisms, genomics, and biological functions.* Cell, 2014. **156**(1-2): p. 45-68.

444. Zhu, H., G. Wang, and J. Qian, *Transcription factors as readers and effectors of DNA methylation.* Nat Rev Genet, 2016. **17**(9): p. 551-65.

445. Meehan, R.R., et al., *Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs.* Cell, 1989. **58**(3): p. 499-507.

446. Hendrich, B. and A. Bird, *Identification and characterization of a family of mammalian methyl-CpG binding proteins.* Mol Cell Biol, 1998. **18**(11): p. 6538-47.

447. Domcke, S., et al., *Competition between DNA methylation and transcription factors determines binding of NRF1.* Nature, 2015. **528**(7583): p. 575-9.

448. Breiling, A. and F. Lyko, *Epigenetic regulatory functions of DNA modifications: 5-methylcytosine and beyond.* Epigenetics Chromatin, 2015. **8**: p. 24.

449. Franchini, D.M., K.M. Schmitz, and S.K. Petersen-Mahrt, *5-Methylcytosine DNA demethylation: more than losing a methyl group.* Annu Rev Genet, 2012. **46**: p. 419-41.

450. Jaenisch, R. and A. Bird, *Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals.* Nat Genet, 2003. **33 Suppl**: p. 245-54.

451. Panning, B. and R. Jaenisch, *DNA hypomethylation can activate Xist expression and silence X-linked genes.* Genes Dev, 1996. **10**(16): p. 1991-2002.

452. Li, E., et al., *DNA methylation, genomic imprinting, and mammalian development.* Cold Spring Harb Symp Quant Biol, 1993. **58**: p. 297-305.

453. Li, E., C. Beard, and R. Jaenisch, *Role for DNA methylation in genomic imprinting.* Nature, 1993. **366**(6453): p. 362-5.

454. Li, E., T.H. Bestor, and R. Jaenisch, *Targeted mutation of the DNA methyltransferase gene results in embryonic lethality.* Cell, 1992. **69**(6): p. 915-26.

455. Walsh, C.P., J.R. Chaillet, and T.H. Bestor, *Transcription of IAP endogenous retroviruses is constrained by cytosine methylation.* Nat Genet, 1998. **20**(2): p. 116-7.

456. Luger, K. and T.J. Richmond, *The histone tails of the nucleosome.* Curr Opin Genet Dev, 1998. **8**(2): p. 140-6.

457. Richmond, T.J. and C.A. Davey, *The structure of DNA in the nucleosome core.* Nature, 2003. **423**(6936): p. 145-50.

458. Kebede, A.F., R. Schneider, and S. Daujat, *Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest.* FEBS J, 2015. **282**(9): p. 1658-74.

459. Kouzarides, T., *Chromatin modifications and their function.* Cell, 2007. **128**(4): p. 693-705.

460. Zhao, Y. and B.A. Garcia, *Comprehensive Catalog of Currently Documented Histone Modifications.* Cold Spring Harb Perspect Biol, 2015. **7**(9): p. a025064.

461. Turner, B.M., *Decoding the nucleosome.* Cell, 1993. **75**(1): p. 5-8.

462. Strahl, B.D. and C.D. Allis, *The language of covalent histone modifications.* Nature, 2000. **403**(6765): p. 41-5.

463. Jenuwein, T. and C.D. Allis, *Translating the histone code.* Science, 2001. **293**(5532): p. 1074-80.

464. van Rossum, B., W. Fischle, and P. Selenko, *Asymmetrically modified nucleosomes expand the histone code.* Nat Struct Mol Biol, 2012. **19**(11): p. 1064-6.

465. Voigt, P., et al., *Asymmetrically modified nucleosomes.* Cell, 2012. **151**(1): p. 181-93.

466. Norton, V.G., et al., *Histone acetylation reduces nucleosome core particle linking number change.* Cell, 1989. **57**(3): p. 449-57.

467. Rice, J.C. and C.D. Allis, *Histone methylation versus histone acetylation: new insights into epigenetic regulation.* Curr Opin Cell Biol, 2001. **13**(3): p. 263-73.

468. Wang, X., et al., *Acetylation increases the alpha-helical content of the histone tails of the nucleosome.* J Biol Chem, 2000. **275**(45): p. 35013-20.

469. Arrowsmith, C.H., et al., *Epigenetic protein families: a new frontier for drug discovery.* Nat Rev Drug Discov, 2012. **11**(5): p. 384-400.

470. Badeaux, A.I. and Y. Shi, *Emerging roles for chromatin as a signal integration and storage platform.* Nat Rev Mol Cell Biol, 2013. **14**(4): p. 211-24.

471. Yun, M., et al., *Readers of histone modifications.* Cell Res, 2011. **21**(4): p. 564-78.

472. Eberl, H.C., et al., *A map of general and specialized chromatin readers in mouse tissues generated by label-free interaction proteomics.* Mol Cell, 2013. **49**(2): p. 368-78.

473. Taverna, S.D., et al., *How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers.* Nat Struct Mol Biol, 2007. **14**(11): p. 1025-1040.

474. Liokatis, S., et al., *Phosphorylation of histone H3 Ser10 establishes a hierarchy for subsequent intramolecular modification events.* Nat Struct Mol Biol, 2012. **19**(8): p. 819-23.

475. Patel, D.J. and Z. Wang, *Readout of epigenetic modifications.* Annu Rev Biochem, 2013. **82**: p. 81-118.

476. Cho, H., et al., *A human RNA polymerase II complex containing factors that modify chromatin structure.* Mol Cell Biol, 1998. **18**(9): p. 5355-63.

477. Gerber, M. and A. Shilatifard, *Transcriptional elongation by RNA polymerase II and histone methylation.* J Biol Chem, 2003. **278**(29): p. 26303-6.

478. Cao, R. and Y. Zhang, *The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3.* Curr Opin Genet Dev, 2004. **14**(2): p. 155-64.

479. Muller, J., et al., *Histone methyltransferase activity of a Drosophila Polycomb group repressor complex.* Cell, 2002. **111**(2): p. 197-208.

480. Pasini, D., et al., *Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity.* EMBO J, 2004. **23**(20): p. 4061-71.

481. Schuettengruber, B., et al., *Genome regulation by polycomb and trithorax proteins.* Cell, 2007. **128**(4): p. 735-45.

482. Brower-Toland, B., et al., *Multiple SET methyltransferases are required to maintain normal heterochromatin domains in the genome of Drosophila melanogaster.* Genetics, 2009. **181**(4): p. 1303-19.

483. Dodge, J.E., et al., *Histone H3-K9 methyltransferase ESET is essential for early development.* Mol Cell Biol, 2004. **24**(6): p. 2478-86.

484. Fritsch, L., et al., *A subset of the histone H3 lysine 9 methyltransferases Suv39h1, G9a, GLP, and SETDB1 participate in a multimeric complex.* Mol Cell, 2010. **37**(1): p. 46-56.

485. Harte, P.J., et al., *Assignment of a novel bifurcated SET domain gene, SETDB1, to human chromosome band 1q21 by in situ hybridization and radiation hybrids.* Cytogenet Cell Genet, 1999. **84**(1-2): p. 83-6.

486. Lachner, M., et al., *Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins.* Nature, 2001. **410**(6824): p. 116-20.

487. Loyola, A., et al., *The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin.* EMBO Rep, 2009. **10**(7): p. 769-75.

488. Melcher, M., et al., *Structure-function analysis of SUV39H1 reveals a dominant role in heterochromatin organization, chromosome segregation, and mitotic progression.* Mol Cell Biol, 2000. **20**(10): p. 3728-41.

489. Rea, S., et al., *Regulation of chromatin structure by site-specific histone H3 methyltransferases.* Nature, 2000. **406**(6796): p. 593-9.

490. Wang, H., et al., *mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression.* Mol Cell, 2003. **12**(2): p. 475-87.

491. Smith, B.C. and J.M. Denu, *Chemical mechanisms of histone lysine and arginine modifications.* Biochim Biophys Acta, 2009. **1789**(1): p. 45-57.

492. Di Lorenzo, A. and M.T. Bedford, *Histone arginine methylation.* FEBS Lett, 2011. **585**(13): p. 2024-31.

493. Bedford, M.T. and S. Richard, *Arginine methylation an emerging regulator of protein function.* Mol Cell, 2005. **18**(3): p. 263-72.

494. Accari, S.L. and P.R. Fisher, *Emerging Roles of JmjC Domain-Containing Proteins.* Int Rev Cell Mol Biol, 2015. **319**: p. 165-220.

495. Shi, Y., et al., *Histone demethylation mediated by the nuclear amine oxidase homolog LSD1.* Cell, 2004. **119**(7): p. 941-53.

496. Whetstine, J.R., et al., *Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases.* Cell, 2006. **125**(3): p. 467-81.

497. Fodor, B.D., et al., *Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells.* Genes Dev, 2006. **20**(12): p. 1557-62.

498. Brownell, J.E. and C.D. Allis, *Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation.* Curr Opin Genet Dev, 1996. **6**(2): p. 176-84.

499. Brownell, J.E., et al., *Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation.* Cell, 1996. **84**(6): p. 843-51.

500. Grunstein, M., *Histone acetylation in chromatin structure and transcription.* Nature, 1997. **389**(6649): p. 349-52.

501. Kuo, M.H., et al., *Transcription-linked acetylation by Gcn5p of histones H3 and H4 at specific lysines.* Nature, 1996. **383**(6597): p. 269-72.

502. Shahbazian, M.D. and M. Grunstein, *Functions of site-specific histone acetylation and deacetylation.* Annu Rev Biochem, 2007. **76**: p. 75-100.

503. Struhl, K., *Histone acetylation and transcriptional regulatory mechanisms.* Genes Dev, 1998. **12**(5): p. 599-606.

504. Bannister, A.J. and T. Kouzarides, *The CBP co-activator is a histone acetyltransferase.* Nature, 1996. **384**(6610): p. 641-3.

505. Ogryzko, V.V., et al., *The transcriptional coactivators p300 and CBP are histone acetyltransferases.* Cell, 1996. **87**(5): p. 953-9.

506. Tie, F., et al., *CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing.* Development, 2009. **136**(18): p. 3131-41.

507. Marmorstein, R. and M.M. Zhou, *Writers and readers of histone acetylation: structure, mechanism, and inhibition.* Cold Spring Harb Perspect Biol, 2014. **6**(7): p. a018762.

508. Mizzen, C.A., et al., *The TAF(II)250 subunit of TFIID has histone acetyltransferase activity.* Cell, 1996. **87**(7): p. 1261-70.

509. Taunton, J., C.A. Hassig, and S.L. Schreiber, *A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p.* Science, 1996. **272**(5260): p. 408-11.

510. Allfrey, V.G., R. Faulkner, and A.E. Mirsky, *Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis.* Proc Natl Acad Sci U S A, 1964. **51**: p. 786-94.

511. Henikoff, S. and M.M. Smith, *Histone variants and epigenetics.* Cold Spring Harb Perspect Biol, 2015. **7**(1): p. a019364.

512. Talbert, P.B. and S. Henikoff, *Histone variants on the move: substrates for chromatin dynamics.* Nat Rev Mol Cell Biol, 2017. **18**(2): p. 115-126.

513. Ahmad, K. and S. Henikoff, *Histone H3 variants specify modes of chromatin assembly.* Proc Natl Acad Sci U S A, 2002. **99 Suppl 4**: p. 16477-84.

514. Campos, E.I. and D. Reinberg, *Histones: annotating chromatin.* Annu Rev Genet, 2009. **43**: p. 559-99.

515. Marzluff, W.F., et al., *The human and mouse replication-dependent histone genes.* Genomics, 2002. **80**(5): p. 487-98.

516. Talbert, P.B. and S. Henikoff, *Histone variants--ancient wrap artists of the epigenome.* Nat Rev Mol Cell Biol, 2010. **11**(4): p. 264-75.

517. Tagami, H., et al., *Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis.* Cell, 2004. **116**(1): p. 51-61.

518. Cheema, M.S. and J. Ausio, *The Structural Determinants behind the Epigenetic Role of Histone Variants.* Genes (Basel), 2015. **6**(3): p. 685-713.

519. Talbert, P.B. and S. Henikoff, *Environmental responses mediated by histone variants.* Trends Cell Biol, 2014. **24**(11): p. 642-50.

520. Weber, C.M. and S. Henikoff, *Histone variants: dynamic punctuation in transcription.* Genes Dev, 2014. **28**(7): p. 672-82.

521. Creyghton, M.P., et al., *H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment.* Cell, 2008. **135**(4): p. 649-61.

522. Gevry, N., et al., *p21 transcription is regulated by differential localization of histone H2A.Z.* Genes Dev, 2007. **21**(15): p. 1869-81.

523. Gallant-Behm, C.L., et al., *DeltaNp63alpha represses anti-proliferative genes via H2A.Z deposition.* Genes Dev, 2012. **26**(20): p. 2325-36.

524. Hu, G., et al., *H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation.* Cell Stem Cell, 2013. **12**(2): p. 180-92.

525. Li, Z., et al., *Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation.* Cell, 2012. **151**(7): p. 1608-16.

526. Buschbeck, M., et al., *The histone variant macroH2A is an epigenetic regulator of key developmental genes.* Nat Struct Mol Biol, 2009. **16**(10): p. 1074-9.

527. Chang, E.Y., et al., *MacroH2A allows ATP-dependent chromatin remodeling by SWI/SNF and ACF complexes but specifically reduces recruitment of SWI/SNF.* Biochemistry, 2008. **47**(51): p. 13726-32.

528. Chakravarthy, S., et al., *Structural characterization of the histone variant macroH2A.* Mol Cell Biol, 2005. **25**(17): p. 7616-24.

529. Maze, I., et al., *Every amino acid matters: essential contributions of histone variants to mammalian development and disease.* Nat Rev Genet, 2014. **15**(4): p. 259-71.

530. Santoro, S.W. and C. Dulac, *Histone variants and cellular plasticity.* Trends Genet, 2015. **31**(9): p. 516-27.

531. Bell, O., et al., *Determinants and dynamics of genome accessibility.* Nat Rev Genet, 2011. **12**(8): p. 554-64.

532. Becker, P.B. and J.L. Workman, *Nucleosome remodeling and epigenetics.* Cold Spring Harb Perspect Biol, 2013. **5**(9).

533. Ho, L. and G.R. Crabtree, *Chromatin remodelling during development.* Nature, 2010. **463**(7280): p. 474-84.

534. Clapier, C.R., et al., *Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes.* Nat Rev Mol Cell Biol, 2017. **18**(7): p. 407-422.

535. Cote, J., et al., *Stimulation of GAL4 derivative binding to nucleosomal DNA by the yeast SWI/SNF complex.* Science, 1994. **265**(5168): p. 53-60.

536. Hirschhorn, J.N., et al., *Evidence that SNF2/SWI2 and SNF5 activate transcription in yeast by altering chromatin structure.* Genes Dev, 1992. **6**(12A): p. 2288-98.

537. Ito, T., et al., *ACF, an ISWI-containing and ATP-utilizing chromatin assembly and remodeling factor.* Cell, 1997. **90**(1): p. 145-55.

538. Kwon, H., et al., *Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex.* Nature, 1994. **370**(6489): p. 477-81.

539. Tsukiyama, T. and C. Wu, *Purification and properties of an ATP-dependent nucleosome remodeling factor.* Cell, 1995. **83**(6): p. 1011-20.

540. Varga-Weisz, P.D., et al., *Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II.* Nature, 1997. **388**(6642): p. 598-602.

541. Flaus, A., et al., *Identification of multiple distinct Snf2 subfamilies with conserved structural motifs.* Nucleic Acids Res, 2006. **34**(10): p. 2887-905.

542. Clapier, C.R. and B.R. Cairns, *The biology of chromatin remodeling complexes.* Annu Rev Biochem, 2009. **78**: p. 273-304.

543. Bao, Y. and X. Shen, *SnapShot: chromatin remodeling complexes.* Cell, 2007. **129**(3): p. 632.

544. Lessard, J.A. and G.R. Crabtree, *Chromatin regulatory mechanisms in pluripotency.* Annu Rev Cell Dev Biol, 2010. **26**: p. 503-32.

545. Chatterjee, N., et al., *Histone H3 tail acetylation modulates ATP-dependent remodeling through multiple mechanisms.* Nucleic Acids Res, 2011. **39**(19): p. 8378-91.

546. Gutierrez, J.L., et al., *Activation domains drive nucleosome eviction by SWI/SNF.* EMBO J, 2007. **26**(3): p. 730-40.

547. Li, M., et al., *Dynamic regulation of transcription factors by nucleosome remodeling.* Elife, 2015. **4**.

548. Owen-Hughes, T. and J.L. Workman, *Remodeling the chromatin structure of a nucleosome array by transcription factor-targeted trans-displacement of histones.* EMBO J, 1996. **15**(17): p. 4702-12.

549. Pollard, K.J. and C.L. Peterson, *Chromatin remodeling: a marriage between two families?* Bioessays, 1998. **20**(9): p. 771-80.

550. Suganuma, T. and J.L. Workman, *Signals and combinatorial functions of histone modifications.* Annu Rev Biochem, 2011. **80**: p. 473-99.

551. Yudkovsky, N., et al., *Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators.* Genes Dev, 1999. **13**(18): p. 2369-74.

552. Hirose, T., Y. Mishima, and Y. Tomari, *Elements and machinery of non-coding RNAs: toward their taxonomy.* EMBO Rep, 2014. **15**(5): p. 489-507.

553. Mattick, J.S., *Non-coding RNAs: the architects of eukaryotic complexity.* EMBO Rep, 2001. **2**(11): p. 986-91.

554. Rinn, J.L. and H.Y. Chang, *Genome regulation by long noncoding RNAs.* Annu Rev Biochem, 2012. **81**: p. 145-66.

555. Wery, M., M. Kwapisz, and A. Morillon, *Noncoding RNAs in gene regulation.* Wiley Interdiscip Rev Syst Biol Med, 2011. **3**(6): p. 728-38.

556. Chen, J., et al., *Over 20% of human transcripts might form sense-antisense pairs.* Nucleic Acids Res, 2004. **32**(16): p. 4812-20.

557. Li, K. and R. Ramchandran, *Natural antisense transcript: a concomitant engagement with protein-coding transcript.* Oncotarget, 2010. **1**(6): p. 447-52.

558. Mattick, J.S., *The genetic signatures of noncoding RNAs.* PLoS Genet, 2009. **5**(4): p. e1000459.

559. Pelechano, V. and L.M. Steinmetz, *Gene regulation by antisense transcription.* Nat Rev Genet, 2013. **14**(12): p. 880-93.

560. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-11.

561. Waterhouse, P.M., M.W. Graham, and M.B. Wang, *Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA.* Proc Natl Acad Sci U S A, 1998. **95**(23): p. 13959-64.

562. Basyuk, E., et al., *Human let-7 stem-loop precursors harbor features of RNase III cleavage products.* Nucleic Acids Res, 2003. **31**(22): p. 6593-7.

563. Bernstein, E., et al., *Role for a bidentate ribonuclease in the initiation step of RNA interference.* Nature, 2001. **409**(6818): p. 363-6.

564. Berezikov, E., *Evolution of microRNA diversity and regulation in animals.* Nat Rev Genet, 2011. **12**(12): p. 846-60.

565. Castel, S.E. and R.A. Martienssen, *RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond.* Nat Rev Genet, 2013. **14**(2): p. 100-12.

566. Doi, N., et al., *Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors.* Curr Biol, 2003. **13**(1): p. 41-6.

567. Lee, Y., et al., *The nuclear RNase III Drosha initiates microRNA processing.* Nature, 2003. **425**(6956): p. 415-9.

568. Zeng, Y., R. Yi, and B.R. Cullen, *MicroRNAs and small interfering RNAs can inhibit mRNA expression by similar mechanisms.* Proc Natl Acad Sci U S A, 2003. **100**(17): p. 9779-84.

569. Batista, P.J. and H.Y. Chang, *Long noncoding RNAs: cellular address codes in development and disease.* Cell, 2013. **152**(6): p. 1298-307.

570. Bonasio, R. and R. Shiekhattar, *Regulation of transcription by long noncoding RNAs.* Annu Rev Genet, 2014. **48**: p. 433-55.

571. Engreitz, J.M., N. Ollikainen, and M. Guttman, *Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression.* Nat Rev Mol Cell Biol, 2016. **17**(12): p. 756-770.

572. Flynn, R.A. and H.Y. Chang, *Long noncoding RNAs in cell-fate programming and reprogramming.* Cell Stem Cell, 2014. **14**(6): p. 752-61.

573. Guttman, M. and J.L. Rinn, *Modular regulatory principles of large non-coding RNAs.* Nature, 2012. **482**(7385): p. 339-46.

574. Hu, W., J.R. Alvarez-Dominguez, and H.F. Lodish, *Regulation of mammalian cell differentiation by long non-coding RNAs.* EMBO Rep, 2012. **13**(11): p. 971-83.

575. Marques, A.C. and C.P. Ponting, *Catalogues of mammalian long noncoding RNAs: modest conservation and incompleteness.* Genome Biol, 2009. **10**(11): p. R124.

576. Mercer, T.R. and J.S. Mattick, *Structure and function of long noncoding RNAs in epigenetic regulation.* Nat Struct Mol Biol, 2013. **20**(3): p. 300-7.

577. Nagano, T. and P. Fraser, *No-nonsense functions for long noncoding RNAs.* Cell, 2011. **145**(2): p. 178-81.

578. Wang, K.C. and H.Y. Chang, *Molecular mechanisms of long noncoding RNAs.* Mol Cell, 2011. **43**(6): p. 904-14.

579. Boyle, A.P., et al., *High-resolution mapping and characterization of open chromatin across the genome.* Cell, 2008. **132**(2): p. 311-22.

580. Orlando, V., H. Strutt, and R. Paro, *Analysis of chromatin structure by in vivo formaldehyde cross-linking.* Methods, 1997. **11**(2): p. 205-14.

581. Robertson, G., et al., *Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing.* Nat Methods, 2007. **4**(8): p. 651-7.

582. Giresi, P.G., et al., *FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin.* Genome Res, 2007. **17**(6): p. 877-85.

583. Simon, J.M., et al., *Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA.* Nat Protoc, 2012. **7**(2): p. 256-67.

584. Buenrostro, J.D., et al., *ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide.* Curr Protoc Mol Biol, 2015. **109**: p. 21 29 1-9.

585. Bernstein, B.E., et al., *The NIH Roadmap Epigenomics Mapping Consortium.* Nat Biotechnol, 2010. **28**(10): p. 1045-8.

586. Bogdanovic, O., et al., *Dynamics of enhancer chromatin signatures mark the transition from pluripotency to cell specification during embryogenesis.* Genome Res, 2012. **22**(10): p. 2043-53.

587. Bonn, S., et al., *Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development.* Nat Genet, 2012. **44**(2): p. 148-56.

588. Ernst, J., et al., *Mapping and analysis of chromatin state dynamics in nine human cell types.* Nature, 2011. **473**(7345): p. 43-9.

589. Roadmap Epigenomics, C., et al., *Integrative analysis of 111 reference human epigenomes.* Nature, 2015. **518**(7539): p. 317-30.

590. Garraway, L.A. and E.S. Lander, *Lessons from the cancer genome.* Cell, 2013. **153**(1): p. 17-37.

591. Perino, M. and G.J. Veenstra, *Chromatin Control of Developmental Dynamics and Plasticity.* Dev Cell, 2016. **38**(6): p. 610-20.

592. Azuara, V., et al., *Chromatin signatures of pluripotent cell lines.* Nat Cell Biol, 2006. **8**(5): p. 532-8.

593. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. **125**(2): p. 315-26.

594. Boyer, L.A., et al., *Polycomb complexes repress developmental regulators in murine embryonic stem cells.* Nature, 2006. **441**(7091): p. 349-53.

595. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome.* Cell, 2007. **129**(4): p. 823-37.

596. Mikkelsen, T.S., et al., *Genome-wide maps of chromatin state in pluripotent and lineage-committed cells.* Nature, 2007. **448**(7153): p. 553-60.

597. Vastenhouw, N.L. and A.F. Schier, *Bivalent histone modifications in early embryogenesis.* Curr Opin Cell Biol, 2012. **24**(3): p. 374-86.

598. Creyghton, M.P., et al., *Histone H3K27ac separates active from poised enhancers and predicts developmental state.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21931-6.

599. Shlyueva, D., G. Stampfel, and A. Stark, *Transcriptional enhancers: from properties to genome-wide predictions.* Nat Rev Genet, 2014. **15**(4): p. 272-86.

600. May, D., et al., *Large-scale discovery of enhancers from human heart tissue.* Nat Genet, 2011. **44**(1): p. 89-93.

601. Steiner, F.A., et al., *Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling.* Genome Res, 2012. **22**(4): p. 766-77.

602. Goldberg, A.D., et al., *Distinct factors control histone variant H3.3 localization at specific genomic regions.* Cell, 2010. **140**(5): p. 678-91.

603. Jin, C., et al., *H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions.* Nat Genet, 2009. **41**(8): p. 941-5.

604. Martens, J.H., et al., *The profile of repeat-associated histone lysine methylation states in the mouse epigenome.* EMBO J, 2005. **24**(4): p. 800-12.

605. Saksouk, N., E. Simboeck, and J. Dejardin, *Constitutive heterochromatin formation and transcription in mammals.* Epigenetics Chromatin, 2015. **8**: p. 3.

606. Breiling, A., et al., *General transcription factors bind promoters repressed by Polycomb group proteins.* Nature, 2001. **412**(6847): p. 651-5.

607. Dellino, G.I., et al., *Polycomb silencing blocks transcription initiation.* Mol Cell, 2004. **13**(6): p. 887-93.

608. Soufi, A., G. Donahue, and K.S. Zaret, *Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome.* Cell, 2012. **151**(5): p. 994-1004.

609. Becker, J.S., D. Nicetto, and K.S. Zaret, *H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes.* Trends Genet, 2016. **32**(1): p. 29-41.

610. Beisel, C. and R. Paro, *Silencing chromatin: comparing modes and mechanisms.* Nat Rev Genet, 2011. **12**(2): p. 123-35.

611. Margueron, R. and D. Reinberg, *The Polycomb complex PRC2 and its mark in life.* Nature, 2011. **469**(7330): p. 343-9.

612. Trojer, P. and D. Reinberg, *Facultative heterochromatin: is there a distinctive molecular signature?* Mol Cell, 2007. **28**(1): p. 1-13.

613. Hawkins, R.D., et al., *Distinct epigenomic landscapes of pluripotent and lineage-committed human cells.* Cell Stem Cell, 2010. **6**(5): p. 479-91.

614. O'Geen, H., et al., *Genome-wide analysis of KAP1 binding suggests autoregulation of KRAB-ZNFs.* PLoS Genet, 2007. **3**(6): p. e89.

615. Vogel, M.J., et al., *Human heterochromatin proteins form large domains containing KRAB-ZNF genes.* Genome Res, 2006. **16**(12): p. 1493-504.

616. Zhu, J., et al., *Genome-wide chromatin state transitions associated with developmental and environmental cues.* Cell, 2013. **152**(3): p. 642-54.

617. Bahar Halpern, K., T. Vana, and M.D. Walker, *Paradoxical role of DNA methylation in activation of FoxA2 gene expression during endoderm development.* J Biol Chem, 2014. **289**(34): p. 23882-92.

618. Charlet, J., et al., *Bivalent Regions of Cytosine Methylation and H3K27 Acetylation Suggest an Active Role for DNA Methylation at Enhancers.* Mol Cell, 2016. **62**(3): p. 422-31.

619. Hantusch, B., et al., *Sp1/Sp3 and DNA-methylation contribute to basal transcriptional activation of human podoplanin in MG63 versus Saos-2 osteoblastic cells.* BMC Mol Biol, 2007. **8**: p. 20.

620. Niesen, M.I., et al., *Activation of a methylated promoter mediated by a sequence-specific DNA-binding protein, RFX.* J Biol Chem, 2005. **280**(47): p. 38914-22.

621. Zhu, Y., et al., *Predicting enhancer transcription and activity from chromatin modifications.* Nucleic Acids Res, 2013. **41**(22): p. 10032-43.

622. Song, C.X., et al., *Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming.* Cell, 2013. **153**(3): p. 678-91.

623. Sun, Z., et al., *A sensitive approach to map genome-wide 5-hydroxymethylcytosine and 5-formylcytosine at single-base resolution.* Mol Cell, 2015. **57**(4): p. 750-61.

624. Yu, M., et al., *Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome.* Cell, 2012. **149**(6): p. 1368-80.

625. Xia, B., et al., *Bisulfite-free, base-resolution analysis of 5-formylcytosine at the genome scale.* Nat Methods, 2015. **12**(11): p. 1047-50.

626. Battich, N., T. Stoeger, and L. Pelkmans, *Control of Transcript Variability in Single Mammalian Cells.* Cell, 2015. **163**(7): p. 1596-610.

627. Grun, D., et al., *Single-cell messenger RNA sequencing reveals rare intestinal cell types.* Nature, 2015. **525**(7568): p. 251-5.

628. Grun, D. and A. van Oudenaarden, *Design and Analysis of Single-Cell Sequencing Experiments.* Cell, 2015. **163**(4): p. 799-810.

629. Zaret, K.S. and S.E. Mango, *Pioneer transcription factors, chromatin dynamics, and cell fate control.* Curr Opin Genet Dev, 2016. **37**: p. 76-81.

630. Hsu, H.T., et al., *TRANSCRIPTION. Recruitment of RNA polymerase II by the pioneer transcription factor PHA-4.* Science, 2015. **348**(6241): p. 1372-6.

631. Drouin, J., *Minireview: pioneer transcription factors in cell fate specification.* Mol Endocrinol, 2014. **28**(7): p. 989-98.

632. Magnani, L., J. Eeckhoute, and M. Lupien, *Pioneer factors: directing transcriptional regulators within the chromatin environment.* Trends Genet, 2011. **27**(11): p. 465-74.

633. Cirillo, L.A., et al., *Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome.* EMBO J, 1998. **17**(1): p. 244-54.

634. Clark, K.L., et al., *Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5.* Nature, 1993. **364**(6436): p. 412-20.

635. Ramakrishnan, V., et al., *Crystal structure of globular domain of histone H5 and its implications for nucleosome binding.* Nature, 1993. **362**(6417): p. 219-23.

636. Iwafuchi-Doi, M., et al., *The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation.* Mol Cell, 2016. **62**(1): p. 79-91.

637. Zaret, K.S., J. Lerner, and M. Iwafuchi-Doi, *Chromatin Scanning by Dynamic Binding of Pioneer Factors.* Mol Cell, 2016. **62**(5): p. 665-7.

638. Soufi, A., et al., *Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming.* Cell, 2015. **161**(3): p. 555-568.

639. Davis, R.L., H. Weintraub, and A.B. Lassar, *Expression of a single transfected cDNA converts fibroblasts to myoblasts.* Cell, 1987. **51**(6): p. 987-1000.

640. Lassar, A.B., B.M. Paterson, and H. Weintraub, *Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts.* Cell, 1986. **47**(5): p. 649-56.

641. Montserrat, N., et al., *Reprogramming of human fibroblasts to pluripotency with lineage specifiers.* Cell Stem Cell, 2013. **13**(3): p. 341-50.

642. Papp, B. and K. Plath, *Epigenetics of reprogramming to induced pluripotency.* Cell, 2013. **152**(6): p. 1324-43.

643. Qin, H., et al., *Epigenetic Control of Reprogramming and Transdifferentiation by Histone Modifications.* Stem Cell Rev, 2016. **12**(6): p. 708-720.

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

645. Takahashi, K. and S. Yamanaka, *A decade of transcription factor-mediated reprogramming to pluripotency.* Nat Rev Mol Cell Biol, 2016. **17**(3): p. 183-93.

646. Vierbuchen, T. and M. Wernig, *Direct lineage conversions: unnatural but useful?* Nat Biotechnol, 2011. **29**(10): p. 892-907.

647. Vierbuchen, T. and M. Wernig, *Molecular roadblocks for cellular reprogramming.* Mol Cell, 2012. **47**(6): p. 827-38.

648. Amamoto, R. and P. Arlotta, *Development-inspired reprogramming of the mammalian central nervous system.* Science, 2014. **343**(6170): p. 1239882.

649. Arlotta, P. and B. Berninger, *Brains in metamorphosis: reprogramming cell identity within the central nervous system.* Curr Opin Neurobiol, 2014. **27**: p. 208-14.

650. Vierbuchen, T., et al., *Direct conversion of fibroblasts to functional neurons by defined factors.* Nature, 2010. **463**(7284): p. 1035-41.

651. Yang, N., et al., *Induced neuronal cells: how to make and define a neuron.* Cell Stem Cell, 2011. **9**(6): p. 517-25.

652. Chen, J., et al., *H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs.* Nat Genet, 2013. **45**(1): p. 34-42.

653. Matoba, S., et al., *Embryonic development following somatic cell nuclear transfer impeded by persisting histone methylation.* Cell, 2014. **159**(4): p. 884-95.

654. Sridharan, R., et al., *Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1gamma in reprogramming to pluripotency.* Nat Cell Biol, 2013. **15**(7): p. 872-82.

655. Adefuin, A.M., et al., *Epigenetic mechanisms regulating differentiation of neural stem/precursor cells.* Epigenomics, 2014. **6**(6): p. 637-49.

656. Nord, A.S., et al., *Genomic perspectives of transcriptional regulation in forebrain development.* Neuron, 2015. **85**(1): p. 27-47.

657. Pataskar, A., J. Jung, and V.K. Tiwari, *Deciphering the epigenetic code of neurogenesis.* Reviews in Cell Biology and Molecular Medicine 2016. **2**: p. 168–220.

658. Shibata, M., F.O. Gulden, and N. Sestan, *From trans to cis: transcriptional regulatory networks in neocortical development.* Trends Genet, 2015. **31**(2): p. 77-87.

659. Yao, B., et al., *Epigenetic mechanisms in neurogenesis.* Nat Rev Neurosci, 2016. **17**(9): p. 537-49.

660. Yao, B. and P. Jin, *Unlocking epigenetic codes in neurogenesis.* Genes Dev, 2014. **28**(12): p. 1253-71.

661. Osumi, N., et al., *Concise review: Pax6 transcription factor contributes to both embryonic and adult neurogenesis as a multifunctional regulator.* Stem Cells, 2008. **26**(7): p. 1663-72.

662. Georgala, P.A., C.B. Carr, and D.J. Price, *The role of Pax6 in forebrain development.* Dev Neurobiol, 2011. **71**(8): p. 690-709.

663. Shaham, O., et al., *Pax6: a multi-level regulator of ocular development.* Prog Retin Eye Res, 2012. **31**(5): p. 351-76.

664. Ypsilanti, A.R. and J.L. Rubenstein, *Transcriptional and epigenetic mechanisms of early cortical development: An examination of how Pax6 coordinates cortical development.* J Comp Neurol, 2016. **524**(3): p. 609-29.

665. Stoykova, A., et al., *Forebrain patterning defects in Small eye mutant mice.* Development, 1996. **122**(11): p. 3453-65.

666. Stoykova, A., et al., *Pax6 modulates the dorsoventral patterning of the mammalian telencephalon.* J Neurosci, 2000. **20**(21): p. 8042-50.

667. Toresson, H., S.S. Potter, and K. Campbell, *Genetic control of dorsal-ventral identity in the telencephalon: opposing roles for Pax6 and Gsh2.* Development, 2000. **127**(20): p. 4361-71.

668. Yun, K., S. Potter, and J.L. Rubenstein, *Gsh2 and Pax6 play complementary roles in dorsoventral patterning of the mammalian telencephalon.* Development, 2001. **128**(2): p. 193-205.

669. Schmahl, W., et al., *Defects of neuronal migration and the pathogenesis of cortical malformations are associated with Small eye (Sey) in the mouse, a point mutation at the Pax-6-locus.* Acta Neuropathol, 1993. **86**(2): p. 126-35.

670. Chapouton, P., A. Gartner, and M. Gotz, *The role of Pax6 in restricting cell migration between developing cortex and basal ganglia.* Development, 1999. **126**(24): p. 5569-79.

671. Nomura, T. and N. Osumi, *Misrouting of mitral cell progenitors in the Pax6/small eye rat telencephalon.* Development, 2004. **131**(4): p. 787-96.

672. Gotz, M., A. Stoykova, and P. Gruss, *Pax6 controls radial glia differentiation in the cerebral cortex.* Neuron, 1998. **21**(5): p. 1031-44.

673. Hack, M.A., et al., *Regionalization and fate specification in neurospheres: the role of Olig2 and Pax6.* Mol Cell Neurosci, 2004. **25**(4): p. 664-78.

674. Haubst, N., et al., *Molecular dissection of Pax6 function: the specific roles of the paired domain and homeodomain in brain development.* Development, 2004. **131**(24): p. 6131-40.

675. Tarabykin, V., et al., *Cortical upper layer neurons derive from the subventricular zone as indicated by Svet1 gene expression.* Development, 2001. **128**(11): p. 1983-93.

676. Tuoc, T.C., et al., *Selective cortical layering abnormalities and behavioral deficits in cortex-specific Pax6 knock-out mice.* J Neurosci, 2009. **29**(26): p. 8335-49.

677. Quinn, J.C., et al., *Pax6 controls cerebral cortical cell number by regulating exit from the cell cycle and specifies cortical cell identity by a cell autonomous mechanism.* Dev Biol, 2007. **302**(1): p. 50-65.

678. Hevner, R.F., et al., *Tbr1 regulates differentiation of the preplate and layer 6.* Neuron, 2001. **29**(2): p. 353-66.

679. Holm, P.C., et al., *Loss- and gain-of-function analyses reveal targets of Pax6 in the developing mouse telencephalon.* Mol Cell Neurosci, 2007. **34**(1): p. 99-119.

680. Lee, J.K., et al., *Expression of neuroD/BETA2 in mitotic and postmitotic neuronal cells during the development of nervous system.* Dev Dyn, 2000. **217**(4): p. 361-7.

681. Georgala, P.A., M. Manuel, and D.J. Price, *The generation of superficial cortical layers is regulated by levels of the transcription factor Pax6.* Cereb Cortex, 2011. **21**(1): p. 81-94.

682. Manuel, M.N., et al., *Regulation of cerebral cortical neurogenesis by the Pax6 transcription factor.* Front Cell Neurosci, 2015. **9**: p. 70.

683. Sansom, S.N., et al., *The level of the transcription factor Pax6 is essential for controlling the balance between neural stem cell self-renewal and neurogenesis.* PLoS Genet, 2009. **5**(6): p. e1000511.

684. Scardigli, R., et al., *Direct and concentration-dependent regulation of the proneural gene Neurogenin2 by Pax6.* Development, 2003. **130**(14): p. 3269-81.

685. Walcher, T., et al., *Functional dissection of the paired domain of Pax6 reveals molecular mechanisms of coordinating neurogenesis and proliferation.* Development, 2013. **140**(5): p. 1123-36.

686. Sarkar, A. and K. Hochedlinger, *The sox family of transcription factors: versatile regulators of stem and progenitor cell fate.* Cell Stem Cell, 2013. **12**(1): p. 15-30.

687. Wegner, M. and C.C. Stolt, *From stem cells to neurons and glia: a Soxist's view of neural development.* Trends Neurosci, 2005. **28**(11): p. 583-8.

688. Bergsland, M., et al., *Sequentially acting Sox transcription factors in neural lineage development.* Genes Dev, 2011. **25**(23): p. 2453-64.

689. Wegner, M., *SOX after SOX: SOXession regulates neurogenesis.* Genes Dev, 2011. **25**(23): p. 2423-8.

690. Gulisano, M., et al., *Emx1 and Emx2 show different patterns of expression during proliferation and differentiation of the developing cerebral cortex in the mouse.* Eur J Neurosci, 1996. **8**(5): p. 1037-50.

691. Liu, Q., N.D. Dwyer, and D.D. O'Leary, *Differential expression of COUP-TFI, CHL1, and two novel genes in developing neocortex identified by differential display PCR.* J Neurosci, 2000. **20**(20): p. 7682-90.

692. Sahara, S., et al., *Sp8 exhibits reciprocal induction with Fgf8 but has an opposing effect on anterior-posterior cortical area patterning.* Neural Dev, 2007. **2**: p. 10.

693. Waclaw, R.R., et al., *The zinc finger transcription factor Sp8 regulates the generation and diversity of olfactory bulb interneurons.* Neuron, 2006. **49**(4): p. 503-16.

694. Zhou, C., S.Y. Tsai, and M.J. Tsai, *COUP-TFI: an intrinsic factor for early regionalization of the neocortex.* Genes Dev, 2001. **15**(16): p. 2054-9.

695. Chou, S.J. and D.D. O'Leary, *Role for Lhx2 in corticogenesis through regulation of progenitor differentiation.* Mol Cell Neurosci, 2013. **56**: p. 1-9.

696. Hanashima, C., et al., *Brain factor-1 controls the proliferation and differentiation of neocortical progenitor cells through independent mechanisms.* J Neurosci, 2002. **22**(15): p. 6526-36.

697. Imayoshi, I., et al., *Essential roles of Notch signaling in maintenance of neural stem cells in developing and adult brains.* J Neurosci, 2010. **30**(9): p. 3489-98.

698. Porter, F.D., et al., *Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive erythrocyte development.* Development, 1997. **124**(15): p. 2935-44.

699. Siegenthaler, J.A., B.A. Tremper-Wells, and M.W. Miller, *Foxg1 haploinsufficiency reduces the population of cortical intermediate progenitor cells: effect of increased p21 expression.* Cereb Cortex, 2008. **18**(8): p. 1865-75.

700. Yoshida, M., et al., *Emx1 and Emx2 functions in development of dorsal telencephalon.* Development, 1997. **124**(1): p. 101-11.

701. Cubelos, B., et al., *Cux-1 and Cux-2 control the development of Reelin expressing cortical interneurons.* Dev Neurobiol, 2008. **68**(7): p. 917-25.

702. Schuurmans, C., et al., *Sequential phases of cortical specification involve Neurogenin-dependent and -independent pathways.* EMBO J, 2004. **23**(14): p. 2892-902.

703. Sugitani, Y., et al., *Brn-1 and Brn-2 share crucial roles in the production and positioning of mouse neocortical neurons.* Genes Dev, 2002. **16**(14): p. 1760-5.

704. Guo, C., et al., *Fezf2 expression identifies a multipotent progenitor for neocortical projection neurons, astrocytes, and oligodendrocytes.* Neuron, 2013. **80**(5): p. 1167-74.

705. Yun, K., et al., *Id4 regulates neural progenitor proliferation and differentiation in vivo.* Development, 2004. **131**(21): p. 5441-8.

706. Britz, O., et al., *A role for proneural genes in the maturation of cortical progenitor cells.* Cereb Cortex, 2006. **16 Suppl 1**: p. i138-51.

707. Farkas, L.M., et al., *Insulinoma-associated 1 has a panneurogenic role and promotes the generation and expansion of basal progenitors in the developing mouse neocortex.* Neuron, 2008. **60**(1): p. 40-55.

708. Gaiano, N., J.S. Nye, and G. Fishell, *Radial glial identity is promoted by Notch1 signaling in the murine forebrain.* Neuron, 2000. **26**(2): p. 395-404.

709. Kageyama, R. and T. Ohtsuka, *The Notch-Hes pathway in mammalian neural development.* Cell Res, 1999. **9**(3): p. 179-88.

710. Kageyama, R., et al., *Roles of bHLH genes in neural stem cell differentiation.* Exp Cell Res, 2005. **306**(2): p. 343-8.

711. Ohata, S., et al., *Dual roles of Notch in regulation of apically restricted mitosis and apicobasal polarity of neuroepithelial cells.* Neuron, 2011. **69**(2): p. 215-30.

712. Ohtsuka, T., et al., *Hes1 and Hes5 as notch effectors in mammalian neuronal differentiation.* EMBO J, 1999. **18**(8): p. 2196-207.

713. Hatakeyama, J., et al., *Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation.* Development, 2004. **131**(22): p. 5539-50.

714. Lyden, D., et al., *Id1 and Id3 are required for neurogenesis, angiogenesis and vascularization of tumour xenografts.* Nature, 1999. **401**(6754): p. 670-7.

715. Hirata, H., et al., *Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop.* Science, 2002. **298**(5594): p. 840-3.

716. Shimojo, H., T. Ohtsuka, and R. Kageyama, *Oscillations in notch signaling regulate maintenance of neural progenitors.* Neuron, 2008. **58**(1): p. 52-64.

717. Shimojo, H., T. Ohtsuka, and R. Kageyama, *Dynamic expression of notch signaling genes in neural stem/progenitor cells.* Front Neurosci, 2011. **5**: p. 78.

718. Imayoshi, I., et al., *Oscillatory control of factors determining multipotency and fate in mouse neural progenitors.* Science, 2013. **342**(6163): p. 1203-8.

719. Bertrand, N., D.S. Castro, and F. Guillemot, *Proneural genes and the specification of neural cell types.* Nat Rev Neurosci, 2002. **3**(7): p. 517-30.

720. Powell, L.M. and A.P. Jarman, *Context dependence of proneural bHLH proteins.* Curr Opin Genet Dev, 2008. **18**(5): p. 411-7.

721. Wilkinson, G., D. Dennis, and C. Schuurmans, *Proneural genes in neocortical development.* Neuroscience, 2013. **253**: p. 256-73.

722. Nieto, M., et al., *Neural bHLH genes control the neuronal versus glial fate decision in cortical progenitors.* Neuron, 2001. **29**(2): p. 401-13.

723. Tomita, K., et al., *Mammalian achaete-scute and atonal homologs regulate neuronal versus glial fate determination in the central nervous system.* EMBO J, 2000. **19**(20): p. 5460-72.

724. Heinrich, C., et al., *Directing astroglia from the cerebral cortex into subtype specific functional neurons.* PLoS Biol, 2010. **8**(5): p. e1000373.

725. Marro, S. and N. Yang, *Transdifferentiation of mouse fibroblasts and hepatocytes to functional neurons.* Methods Mol Biol, 2014. **1150**: p. 237-46.

726. Pang, Z.P., et al., *Induction of human neuronal cells by defined transcription factors.* Nature, 2011. **476**(7359): p. 220-3.

727. Kovach, C., et al., *Neurog2 simultaneously activates and represses alternative gene expression programs in the developing neocortex.* Cereb Cortex, 2013. **23**(8): p. 1884-900.

728. Mattar, P., et al., *A screen for downstream effectors of Neurogenin2 in the embryonic neocortex.* Dev Biol, 2004. **273**(2): p. 373-89.

729. Mattar, P., et al., *Basic helix-loop-helix transcription factors cooperate to specify a cortical projection neuron identity.* Mol Cell Biol, 2008. **28**(5): p. 1456-69.

730. Sun, Y., et al., *Neurogenin promotes neurogenesis and inhibits glial differentiation by independent mechanisms.* Cell, 2001. **104**(3): p. 365-76.

731. Martynoga, B., D. Drechsel, and F. Guillemot, *Molecular control of neurogenesis: a view from the mammalian cerebral cortex.* Cold Spring Harb Perspect Biol, 2012. **4**(10).

732. MacDonald, J.L., et al., *Specification of cortical projection neurons. In: Patterning and Cell Type Specification in the Developing CNS and PNS. Editors: Rubenstein JLR and Rakic P.* Elsevier, 2013: p. 475-502.

733. Rubenstein, J.L.R. and K. Campbell, *Neurogenesis in the Basal Ganglia. In: Patterning and Cell Type Specification in the Developing CNS and PNS. Editors: Rubenstein JLR and Rakic P.* Elsevier, 2013: p. 455-474.

734. Belgard, T.G., et al., *A transcriptomic atlas of mouse neocortical layers.* Neuron, 2011. **71**(4): p. 605-16.

735. Molyneaux, B.J., et al., *DeCoN: genome-wide analysis of in vivo transcriptional dynamics during pyramidal neuron fate selection in neocortex.* Neuron, 2015. **85**(2): p. 275-288.

736. Thompson, C.L., et al., *A high-resolution spatiotemporal atlas of gene expression of the developing mouse brain.* Neuron, 2014. **83**(2): p. 309-323.

737. Hamby, M.E., V. Coskun, and Y.E. Sun, *Transcriptional regulation of neuronal differentiation: the epigenetic layer of complexity.* Biochim Biophys Acta, 2008. **1779**(8): p. 432-7.

738. Jobe, E.M., A.L. McQuate, and X. Zhao, *Crosstalk among Epigenetic Pathways Regulates Neurogenesis.* Front Neurosci, 2012. **6**: p. 59.

739. LaSalle, J.M., W.T. Powell, and D.H. Yasui, *Epigenetic layers and players underlying neurodevelopment.* Trends Neurosci, 2013. **36**(8): p. 460-70.

740. Shin, J., G.L. Ming, and H. Song, *Decoding neural transcriptomes and epigenomes via high-throughput sequencing.* Nat Neurosci, 2014. **17**(11): p. 1463-75.

741. Shin, J., G.L. Ming, and H. Song, *Seeking a roadmap toward neuroepigenetics.* Neuron, 2015. **86**(1): p. 12-5.

742. Sweatt, J.D., *The emerging field of neuroepigenetics.* Neuron, 2013. **80**(3): p. 624-32.

743. Hu, X.L., Y. Wang, and Q. Shen, *Epigenetic control on cell fate choice in neural stem cells.* Protein Cell, 2012. **3**(4): p. 278-90.

744. Wen, S., H. Li, and J. Liu, *Epigenetic background of neuronal fate determination.* Prog Neurobiol, 2009. **87**(2): p. 98-117.

745. Emera, D., et al., *Origin and evolution of developmental enhancers in the mammalian neocortex.* Proc Natl Acad Sci U S A, 2016. **113**(19): p. E2617-26.

746. Nord, A.S., et al., *Rapid and pervasive changes in genome-wide enhancer usage during mammalian development.* Cell, 2013. **155**(7): p. 1521-31.

747. Lister, R., et al., *Global epigenomic reconfiguration during mammalian brain development.* Science, 2013. **341**(6146): p. 1237905.

748. Blow, M.J., et al., *ChIP-Seq identification of weakly conserved heart enhancers.* Nat Genet, 2010. **42**(9): p. 806-10.

749. Pattabiraman, K., et al., *Transcriptional regulation of enhancers active in protodomains of the developing cerebral cortex.* Neuron, 2014. **82**(5): p. 989-1003.

750. Pennacchio, L.A., et al., *In vivo enhancer analysis of human conserved non-coding sequences.* Nature, 2006. **444**(7118): p. 499-502.

751. Stergachis, A.B., et al., *Developmental fate and cellular maturity encoded in human regulatory DNA landscapes.* Cell, 2013. **154**(4): p. 888-903.

752. Berger, J., et al., *Conditional activation of Pax6 in the developing cortex of transgenic mice causes progenitor apoptosis.* Development, 2007. **134**(7): p. 1311-22.

753. Shimomura, A. and E. Hashin, *Epigenetic Regulation of Neural Differentiation from Embryonic Stem Cells.* Trends in Cell Signaling Pathways in Neuronal Fate Decision. Sabine Wislet-Gendebien (Ed.), InTech., 2013.

754. Albert, M., et al., *Epigenome profiling and editing of neocortical progenitor cells during development.* EMBO J, 2017. **36**(17): p. 2642-2658.

755. Broccoli, V., et al., *Histone modifications controlling native and induced neural stem cell identity.* Curr Opin Genet Dev, 2015. **34**: p. 95-101.

756. Pereira, J.D., et al., *Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in the cerebral cortex.* Proc Natl Acad Sci U S A, 2010. **107**(36): p. 15957-62.

757. Jepsen, K., et al., *SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron.* Nature, 2007. **450**(7168): p. 415-9.

758. Burgold, T., et al., *The histone H3 lysine 27-specific demethylase Jmjd3 is required for neural commitment.* PLoS One, 2008. **3**(8): p. e3034.

759. Schmitz, S.U., et al., *Jarid1b targets genes regulating development and is involved in neural differentiation.* EMBO J, 2011. **30**(22): p. 4586-600.

760. Sun, G., et al., *Histone demethylase LSD1 regulates neural stem cell proliferation.* Mol Cell Biol, 2010. **30**(8): p. 1997-2005.

761. Tan, S.L., et al., *Essential roles of the histone methyltransferase ESET in the epigenetic control of neural progenitor cells during development.* Development, 2012. **139**(20): p. 3806-16.

762. Thomas, T. and A.K. Voss, *Querkopf, a histone acetyltransferase, is essential for embryonic neurogenesis.* Front Biosci, 2004. **9**: p. 24-31.

763. Koyano-Nakagawa, N., D. Wettstein, and C. Kintner, *Activation of Xenopus genes required for lateral inhibition and neuronal differentiation during primary neurogenesis.* Mol Cell Neurosci, 1999. **14**(4-5): p. 327-39.

764. Lee, S., et al., *Retinoid signaling and neurogenin2 function are coupled for the specification of spinal motor neurons through a chromatin modifier CBP.* Neuron, 2009. **62**(5): p. 641-54.

765. Wang, J., et al., *CBP histone acetyltransferase activity regulates embryonic neural differentiation in the normal and Rubinstein-Taybi syndrome brain.* Dev Cell, 2010. **18**(1): p. 114-25.

766. MacDonald, J.L. and A.J. Roskams, *Histone deacetylases 1 and 2 are expressed at distinct stages of neuro-glial development.* Dev Dyn, 2008. **237**(8): p. 2256-67.

767. Sun, G., et al., *Histone deacetylases in neural stem cells and induced pluripotent stem cells.* J Biomed Biotechnol, 2011. **2011**: p. 835968.

768. Goto, K., et al., *Expression of DNA methyltransferase gene in mature and immature neurons as well as proliferating cells in mice.* Differentiation, 1994. **56**(1-2): p. 39-44.

769. Feng, J., et al., *Dynamic expression of de novo DNA methyltransferases Dnmt3a and Dnmt3b in the central nervous system.* J Neurosci Res, 2005. **79**(6): p. 734-46.

770. Fan, G., et al., *DNA methylation controls the timing of astrogliogenesis through regulation of JAK-STAT signaling.* Development, 2005. **132**(15): p. 3345-56.

771. Globisch, D., et al., *Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates.* PLoS One, 2010. **5**(12): p. e15367.

772. Hahn, M.A., et al., *Dynamics of 5-hydroxymethylcytosine and chromatin marks in Mammalian neurogenesis.* Cell Rep, 2013. **3**(2): p. 291-300.

773. Kriaucionis, S. and N. Heintz, *The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain.* Science, 2009. **324**(5929): p. 929-30.

774. Szulwach, K.E., et al., *5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging.* Nat Neurosci, 2011. **14**(12): p. 1607-16.

775. Yoo, A.S. and G.R. Crabtree, *ATP-dependent chromatin remodeling in neural development.* Curr Opin Neurobiol, 2009. **19**(2): p. 120-6.

776. Vestin, A. and A.A. Mills, *The tumor suppressor Chd5 is induced during neuronal differentiation in the developing mouse brain.* Gene Expr Patterns, 2013. **13**(8): p. 482-9.

777. Potts, R.C., et al., *CHD5, a brain-specific paralog of Mi2 chromatin remodeling enzymes, regulates expression of neuronal genes.* PLoS One, 2011. **6**(9): p. e24515.

778. Egan, C.M., et al., *CHD5 is required for neurogenesis and has a dual role in facilitating gene expression and polycomb gene repression.* Dev Cell, 2013. **26**(3): p. 223-36.

779. Ho, L., et al., *An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network.* Proc Natl Acad Sci U S A, 2009. **106**(13): p. 5187-91.

780. Lessard, J., et al., *An essential switch in subunit composition of a chromatin remodeling complex during neural development.* Neuron, 2007. **55**(2): p. 201-15.

781. Wu, J.I., et al., *Regulation of dendritic development by neuron-specific chromatin remodeling complexes.* Neuron, 2007. **56**(1): p. 94-108.

782. Yan, Z., et al., *BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells.* Stem Cells, 2008. **26**(5): p. 1155-65.

783. Lamba, D.A., et al., *Baf60c is a component of the neural progenitor-specific BAF complex in developing retina.* Dev Dyn, 2008. **237**(10): p. 3016-23.

784. Ninkovic, J., et al., *The BAF complex interacts with Pax6 in adult neural progenitors to establish a neurogenic cross-regulatory transcriptional network.* Cell Stem Cell, 2013. **13**(4): p. 403-18.

785. Takeuchi, J.K., et al., *Baf60c is a nuclear Notch signaling component required for the establishment of left-right asymmetry.* Proc Natl Acad Sci U S A, 2007. **104**(3): p. 846-51.

786. Tuoc, T.C., et al., *Chromatin regulation by BAF170 controls cerebral cortical size and thickness.* Dev Cell, 2013. **25**(3): p. 256-69.

787. Tuoc, T.C., R. Narayanan, and A. Stoykova, *BAF chromatin remodeling complex: cortical size regulation and beyond.* Cell Cycle, 2013. **12**(18): p. 2953-9.

788. Staahl, B.T. and G.R. Crabtree, *Creating a neural specific chromatin landscape by npBAF and nBAF complexes.* Curr Opin Neurobiol, 2013. **23**(6): p. 903-13.

789. Ooi, L., et al., *BRG1 chromatin remodeling activity is required for efficient chromatin binding by repressor element 1-silencing transcription factor (REST) and facilitates REST-mediated repression.* J Biol Chem, 2006. **281**(51): p. 38974-80.

790. Ballas, N., et al., *REST and its corepressors mediate plasticity of neuronal gene chromatin throughout neurogenesis.* Cell, 2005. **121**(4): p. 645-57.

791. Lunyak, V.V., et al., *Corepressor-dependent silencing of chromosomal regions encoding neuronal genes.* Science, 2002. **298**(5599): p. 1747-52.

792. Schoenherr, C.J. and D.J. Anderson, *The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes.* Science, 1995. **267**(5202): p. 1360-3.

793. Chen, Z.F., A.J. Paquette, and D.J. Anderson, *NRSF/REST is required in vivo for repression of multiple neuronal target genes during embryogenesis.* Nat Genet, 1998. **20**(2): p. 136-42.

794. Jorgensen, H.F., et al., *REST selectively represses a subset of RE1-containing neuronal genes in mouse embryonic stem cells.* Development, 2009. **136**(5): p. 715-21.

795. Chong, J.A., et al., *REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons.* Cell, 1995. **80**(6): p. 949-57.

796. Westbrook, T.F., et al., *SCFbeta-TRCP controls oncogenic transformation and neural differentiation through REST degradation.* Nature, 2008. **452**(7185): p. 370-4.

797. Yoo, A.S., et al., *MicroRNA-mediated switching of chromatin-remodelling complexes in neural development.* Nature, 2009. **460**(7255): p. 642-6.

798. Shen, T., et al., *CHD2 is Required for Embryonic Neurogenesis in the Developing Cerebral Cortex.* Stem Cells, 2015. **33**(6): p. 1794-806.

799. Alexander, J.M. and S. Lomvardas, *Nuclear architecture as an epigenetic regulator of neural development and function.* Neuroscience, 2014. **264**: p. 39-50.

800. Fraser, J., et al., *Hierarchical folding and reorganization of chromosomes are linked to transcriptional changes in cellular differentiation.* Mol Syst Biol, 2015. **11**(12): p. 852.

801. Shchuka, V.M., et al., *Chromatin Dynamics in Lineage Commitment and Cellular Reprogramming.* Genes (Basel), 2015. **6**(3): p. 641-61.

802. Clark, B.S. and S. Blackshaw, *Long non-coding RNA-dependent transcriptional regulation in neuronal development and disease.* Front Genet, 2014. **5**: p. 164.

803. Mercer, T.R., et al., *Specific expression of long noncoding RNAs in the mouse brain.* Proc Natl Acad Sci U S A, 2008. **105**(2): p. 716-21.

804. Mercer, T.R., et al., *Long noncoding RNAs in neuronal-glial fate specification and oligodendrocyte lineage maturation.* BMC Neurosci, 2010. **11**: p. 14.

805. Lin, N., et al., *An evolutionarily conserved long noncoding RNA TUNA controls pluripotency and neural lineage commitment.* Mol Cell, 2014. **53**(6): p. 1005-19.

806. Ramos, A.D., et al., *The long noncoding RNA Pnky regulates neuronal differentiation of embryonic and postnatal neural stem cells.* Cell Stem Cell, 2015. **16**(4): p. 439-447.

807. Tochitani, S. and Y. Hayashizaki, *Nkx2.2 antisense RNA overexpression enhanced oligodendrocytic differentiation.* Biochem Biophys Res Commun, 2008. **372**(4): p. 691-6.

808. Vance, K.W., et al., *The long non-coding RNA Paupar regulates the expression of both local and distal genes.* EMBO J, 2014. **33**(4): p. 296-311.

809. Fineberg, S.K., K.S. Kosik, and B.L. Davidson, *MicroRNAs potentiate neural development.* Neuron, 2009. **64**(3): p. 303-9.

810. Kawahara, H., T. Imai, and H. Okano, *MicroRNAs in Neural Stem Cells and Neurogenesis.* Front Neurosci, 2012. **6**: p. 30.

811. Kosik, K.S., *The neuronal microRNA system.* Nat Rev Neurosci, 2006. **7**(12): p. 911-20.

812. Li, X. and P. Jin, *Roles of small regulatory RNAs in determining neuronal identity.* Nat Rev Neurosci, 2010. **11**(5): p. 329-38.

813. Petri, R., et al., *miRNAs in brain development.* Exp Cell Res, 2014. **321**(1): p. 84-9.

814. Hirabayashi, Y., et al., *Polycomb limits the neurogenic competence of neural precursor cells to promote astrogenic fate transition.* Neuron, 2009. **63**(5): p. 600-13.

815. Roman-Trufero, M., et al., *Maintenance of undifferentiated state and self-renewal of embryonic neural stem cells by Polycomb protein Ring1B.* Stem Cells, 2009. **27**(7): p. 1559-70.

816. Fuentealba, L.C., et al., *Embryonic Origin of Postnatal Neural Stem Cells.* Cell, 2015. **161**(7): p. 1644-55.

817. Furutachi, S., et al., *Slowly dividing neural progenitors are an embryonic origin of adult neural stem cells.* Nat Neurosci, 2015. **18**(5): p. 657-65.

818. Jakovcevski, M. and S. Akbarian, *Epigenetic mechanisms in neurological disease.* Nat Med, 2012. **18**(8): p. 1194-204.

819. Ronan, J.L., W. Wu, and G.R. Crabtree, *From neural development to cognition: unexpected roles for chromatin.* Nat Rev Genet, 2013. **14**(5): p. 347-59.

820. Vermunt, M.W. and M.P. Creyghton, *Transcriptional Dynamics at Brain Enhancers: from Functional Specialization to Neurodegeneration.* Curr Neurol Neurosci Rep, 2016. **16**(10): p. 94.

821. Gifford, W.D., S.L. Pfaff, and T.S. Macfarlan, *Transposable elements as genetic regulatory substrates in early development.* Trends Cell Biol, 2013. **23**(5): p. 218-26.

822. Rebollo, R., et al., *Epigenetic interplay between mouse endogenous retroviruses and host genes.* Genome Biol, 2012. **13**(10): p. R89.

823. Flemr, M., et al., *A retrotransposon-driven dicer isoform directs endogenous small interfering RNA production in mouse oocytes.* Cell, 2013. **155**(4): p. 807-16.

824. Fort, A., et al., *Deep transcriptome profiling of mammalian stem cells supports a regulatory role for retrotransposons in pluripotency maintenance.* Nat Genet, 2014. **46**(6): p. 558-66.

825. Gerdes, P., et al., *Transposable elements in the mammalian embryo: pioneers surviving through stealth and service.* Genome Biol, 2016. **17**: p. 100.

826. Macfarlan, T.S., et al., *Embryonic stem cell potency fluctuates with endogenous retrovirus activity.* Nature, 2012. **487**(7405): p. 57-63.

827. Herquel, B., et al., *Trim24-repressed VL30 retrotransposons regulate gene expression by producing noncoding RNA.* Nat Struct Mol Biol, 2013. **20**(3): p. 339-46.

828. Lu, X., et al., *The retrovirus HERVH is a long noncoding RNA required for human embryonic stem cell identity.* Nat Struct Mol Biol, 2014. **21**(4): p. 423-5.

829. Elsasser, S.J., et al., *Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells.* Nature, 2015. **522**(7555): p. 240-244.

830. Quenneville, S., et al., *The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development.* Cell Rep, 2012. **2**(4): p. 766-73.

831. Rowe, H.M., et al., *TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells.* Genome Res, 2013. **23**(3): p. 452-61.

832. Li, X., et al., *A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints.* Dev Cell, 2008. **15**(4): p. 547-57.

833. Quenneville, S., et al., *In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions.* Mol Cell, 2011. **44**(3): p. 361-72.

834. Fasching, L., et al., *TRIM28 represses transcription of endogenous retroviruses in neural progenitor cells.* Cell Rep, 2015. **10**(1): p. 20-8.

835. Franchini, L.F., et al., *Convergent evolution of two mammalian neuronal enhancers by sequential exaptation of unrelated retroposons.* Proc Natl Acad Sci U S A, 2011. **108**(37): p. 15270-5.

836. Lam, D.D., et al., *Partially redundant enhancers cooperatively maintain Mammalian pomc expression above a critical functional threshold.* PLoS Genet, 2015. **11**(2): p. e1004935.

837. Notwell, J.H., et al., *A family of transposable elements co-opted into developmental enhancers in the mouse neocortex.* Nat Commun, 2015. **6**: p. 6644.

838. Jakobsson, J., et al., *KAP1-mediated epigenetic repression in the forebrain modulates behavioral vulnerability to stress.* Neuron, 2008. **60**(5): p. 818-31.

839. Douville, R., et al., *Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis.* Ann Neurol, 2011. **69**(1): p. 141-51.

840. Hancks, D.C. and H.H. Kazazian, Jr., *Roles for retrotransposon insertions in human disease.* Mob DNA, 2016. **7**: p. 9.

841. Kremer, D., et al., *Human endogenous retrovirus type W envelope protein inhibits oligodendroglial precursor cell differentiation.* Ann Neurol, 2013. **74**(5): p. 721-32.

842. Li, W., et al., *Human endogenous retrovirus-K contributes to motor neuron disease.* Sci Transl Med, 2015. **7**(307): p. 307ra153.

843. Manghera, M., J. Ferguson, and R. Douville, *Endogenous retrovirus-K and nervous system diseases.* Curr Neurol Neurosci Rep, 2014. **14**(10): p. 488.

844. Reilly, M.T., et al., *The role of transposable elements in health and diseases of the central nervous system.* J Neurosci, 2013. **33**(45): p. 17577-86.

845. Baillie, J.K., et al., *Somatic retrotransposition alters the genetic landscape of the human brain.* Nature, 2011. **479**(7374): p. 534-7.

846. Coufal, N.G., et al., *L1 retrotransposition in human neural progenitor cells.* Nature, 2009. **460**(7259): p. 1127-31.

847. Erwin, J.A., M.C. Marchetto, and F.H. Gage, *Mobile DNA elements in the generation of diversity and complexity in the brain.* Nat Rev Neurosci, 2014. **15**(8): p. 497-506.

848. Evrony, G.D., et al., *Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain.* Cell, 2012. **151**(3): p. 483-96.

849. Evrony, G.D., et al., *Cell lineage analysis in human brain using endogenous retroelements.* Neuron, 2015. **85**(1): p. 49-59.

850. Muotri, A.R., et al., *Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition.* Nature, 2005. **435**(7044): p. 903-10.

851. Muotri, A.R., et al., *L1 retrotransposition in neurons is modulated by MeCP2.* Nature, 2010. **468**(7322): p. 443-6.

852. Lupski, J.R., *Genetics. Genome mosaicism--one human, multiple genomes.* Science, 2013. **341**(6144): p. 358-9.

853. Richardson, S.R., S. Morell, and G.J. Faulkner, *L1 retrotransposons and somatic mosaicism in the brain.* Annu Rev Genet, 2014. **48**: p. 1-27.

854. Thomas, C.A., A.C. Paquola, and A.R. Muotri, *LINE-1 retrotransposition in the nervous system.* Annu Rev Cell Dev Biol, 2012. **28**: p. 555-73.

2 Aim of this study

The aim of this thesis was to functionally characterise the interplay of specific transcription factors with the chromatin landscape during neuronal cell fate acquisition.

The following objectives were addressed:

- Unravel whether there are unexplored transcription factors expressed in a neuron-specific context.
- Uncover if these nuclear factors function by directly modulating the chromatin landscape.
- Identify if these factors play a functional role during neuronal differentiation by contributing to the differentiation-associated transcriptional dynamics.

3 Results Part A - NeuroD1

NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program

The following section is based on the publication of Pataskar & Jung et al. (2016) [1]:

Pataskar A#, Jung J#, Smialowski P, Noack F, Calegari F, Straub T, Tiwari VK (2016) NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. The EMBO journal 35: 24-45.

The # indicates equal contribution. Computational analyses were carried out by Abhijeet Pataskar (AP) and Pawel Smialowski (PS). *In utero* electroporations and FAC-sorting of cortical cells were performed by Florian Noack (FN). All other experiments were carried out by myself (JJ). Each figure legend states which panels are derived from this publication and which data was not derived by me.

In mammals, a plethora of transcription factors (TFs) act on numerous levels to orchestrate a spatio-temporally regulated gene expression program which ensures embryonic development [2, 3]. One family of such key players are the basic helix-loop-helix (bHLH) TFs, which contribute to lineage commitment as well as terminal differentiation of various cell types throughout development [4]. NeuroD1 has been described as an important bHLH factor during neuronal differentiation [5, 6]. This protein has recently also been used to reprogram other somatic cell types into neurons: Pang and colleagues successfully used NeuroD1 in combination with Brn2, Ascl1, and Myt1l to reprogram fetal and postnatal fibroblasts into neurons [7]. Furthermore, NeuroD1 alone was able to convert human astrocytes into glutamatergic neurons as well as reactive glial cells into functional neurons *in vivo* [8]. These discoveries imply that this bHLH TF is a highly potent factor that promotes neuronal fate. However, no comprehensive investigation has been performed to reveal the gene regulatory network through which NeuroD1 mediates neuronal fate specification during development. In addition, the direct genomewide targets of NeuroD1 during neurogenesis remain elusive. Despite increasing knowledge that cell fate specification involves reprogramming of the epigenome [9-12], very little is known about whether remodeling of chromatin is involved in NeuroD1's function at its target sites. It is still an open question whether a transient action of such cell fate determining factors is able to induce long-term epigenetic memory during differentiation.

3.1 Generation of transgenic embryonic stem cells for inducible expression of NeuroD1

Embryonic stem cells are due to their pluripotent nature frequently used to study differentiation processes into several cell types from different lineages [13, 14]. Especially in the context of neuronal differentiation, an *in vitro* protocol published by Bibel and colleagues [15-17] enables the generation of a nearly homogenous population of neuronal cells from murine embryonic stem cells (mESCs). Furthermore, as cellular differentiation relies on changes of transcriptional programs, a conditional (over-)expression of

cellular proteins like TFs provides a powerful tool to study cellular processes directly or indirectly regulated by such proteins. This further allows to uncover key regulators of cell fate changes and to elucidate their mode of action. Interestingly, Iacovino and colleagues published a transgenic mESCs line (A2lox.Cre) which contains a doxycycline inducible gene expression locus [18]. These mESCs were reported to efficiently integrate transgenes into this genomic locus, to differentiate into several other cell types like skeletal muscle cells and furthermore to give rise to transgenic mice [18]. Therefore, in order to combine both, the differentiation potential of stem cells as well as the possibility of an ectopic overexpression of NeuroD1, the doxycycline inducible A2lox mESCs were adopted in this thesis to study NeuroD1's downstream effects on chromatin, gene regulation and concomitant cell fate.

3.1.1 Design and principle of NeuroD1-inducible A2lox embryonic stem cells

The A2lox mESCs combine two molecular biology tools to ensure a highly efficient site-specific integration of DNA into as well as a reliable conditional gene expression from a single genomic recombination locus (Diploma thesis Johannes Jung [19]): First, they utilize the Cre-LoxP site-specific recombination system to integrate a sequence of interest such as the *NeuroD1* CDS into the genome [20-22]. This system originates from the bacteriophage P1 and consists of the Cre recombinase enzyme (*Cre*) and its 34 bp target sequences (loxP sites) for site-specific recombination [23]. A typical loxP site has a directionality due to the asymmetric 8 bp spacer sequence which is flanked by 13 bp inverted repeats [24]. Therefore, the Cre recombinase can only mediate the recombination between identical loxP sites due to the spacer sequence [25-27]. The A2lox system encompasses a wild-type loxP as well as a mutant lox2272 site, which is characterised by two transversions in the spacer sequence (G to C and T to A), preventing a cross-recombination between these two sites [24]. This combination of these heterologous loxP sites assures the locus-specific and directional recombination in the A2lox mESCs system.

Second, the A2lox mESCs harbor a tetracycline-dependent transcriptional activator as well as a tetracycline-responsive promoter (tetracycline responsive element, TRE) at the recombination locus to facilitate gene expression in an inducible fashion [28]. Tetracycline-responsive promoters have been widely adapted for inducible gene expression in mammalian cells after their initial discovery in the tetracycline-resistance operon of *E. coli* [28]. This mechanism is based on a reverse tetracycline-controlled transactivator protein (rtTA), which recognizes its specific DNA target sequence, the Tet operator (tetO), only in the presence of tetracycline to induce transcriptional activation of the associated gene(s) [29]. The reverse Tet system is frequently used in mammalian cells to express genes of interest under the control of a TRE by supplementing the culture medium with the tetracycline derivate doxycycline (Dox) [29].

The male A2Lox.Cre mESCs are derived from A17 mESCs [30], which contain the rtTA in the *Rosa26* locus [31-33]. This locus has been shown to guarantee constant gene expression during development [34]. Iacovino and colleagues additionally modified these cells by inserting an inducible cassette exchange (ICE) target locus upstream of the housekeeping gene *Hprt* on the single X-chromosome to generate the A2lox.Cre mESCs [18]: The ICE combines the doxycycline (tetracycline)-responsive promoter (TRE)-2loxP-Δneo construct [35] with the coding sequence for Cre recombinase [36] under the control of this TRE. Hence,

the A2lox.Cre mESCs are able to express the Cre recombinase from the X-chromosome after supplementing the culture medium with doxycycline as the rtTA is constitutively expressed from the *Rosa26* locus on chromosome 6. The A2Lox.Cre mESCs can therefore be used for recombinase-mediated cassette exchange (RMCE) to introduce a gene of interest into the genome under the control of the TRE. The p2lox vector containing the gene of interest flanked by a wildtype loxP and a mutant lox2272 site has to be provided to the cells after the Cre recombinase has been expressed. The principle for generating A2lox.NeuroD1 mESCs using the A2lox.Cre cells in combination with the p2loxHA-NeuroD1 vector is illustrated in Figure 31.

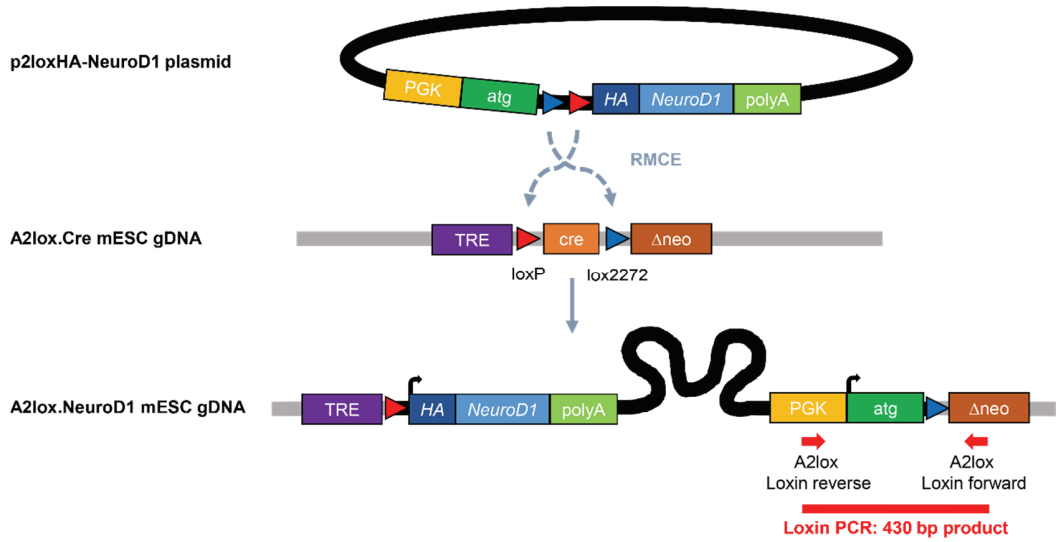


Figure 31: Strategy for recombinase-mediated cassette exchange in A2Lox.Cre embryonic stem cells to generate A2lox.NeuroD1 mESCs.
The Cre recombinase in A2Lox.Cre mESCs is expressed from the TRE promoter after supplying doxycycline to the culture medium. After transfecting the induced cells with the p2loxHA-NeuroD1 vector, the Cre recombinase mediates the recombinase-mediated cassette exchange (RMCE) between the heterologous loxP sites (blue and red triangles) in the genomic and vector DNA. Thereupon, the whole vector DNA, except the sequence between the two loxP sites, is integrated into the target locus by replacing the *cre* gene. The loss of the *cre* gene prevents further recombination events during future inductions. Hence, the HA-NeuroD1 sequence is under the control of the inducible TRE promoter. Furthermore, the neomycin resistance gene in the A2lox genome lacking a start codon as well as a promoter (Δneo) becomes expressed by the upstream insertion of the PGK promoter (PGK) and a start codon (atg) from the p2loxHA-NeuroD1 vector. This functional antibiotic resistance can concomitantly be used for the selection of mESC clones where RMCE took place successfully. An additional PCR-based screening for positive clones combines two primers, one binding in the PGK promoter (A2lox Loxin reverse) and one in the Δneo gene (A2lox Loxin forward). This Loxin PCR only generates a 430 bp product if the RMCE led to a correctly oriented p2lox vector insertion into the ICE locus. The two heterologous genomic loxP sites stay intact.
Figure information: Cre, Cre recombinase; Δneo, genomic neomycin resistance gene lacking the start codon as well as a promoter; gDNA, genomic DNA; HA, human influenza hemagglutinin tag; ICE, inducible cassette exchange; PGK, phosphoglycerate kinase; polyA, polyadenylation sequence; TRE, doxycycline (tetracycline)-responsive element. Modified from [18] by JJ.

In conclusion, introducing the HA-tagged *NeuroD1* sequence stably into the tetracycline inducible gene expression locus of A2lox mESCs enables an ectopic HA-NeuroD1 overexpression directly in these stem cells as well as during their differentiation into neurons *in vitro* [16, 17]. From here onwards the term A2lox.NeuroD1 refers to A2lox mESCs which contain inducible HA-NeuroD1.

3.1.2 Generation and validation of A2Lox.NeuroD1 embryonic stem cells for overexpression of HA-tagged NeuroD1

3.1.2.1 Identification of A2Lox.NeuroD1 clones expressing HA-NeuroD1

The A2lox mESCs system was established and validated using the p2lox.EGFP vector and hence the inducible expression of the EGFP protein (Diploma thesis Johannes Jung [19]). Based on such pilot experiments, a doxycycline concentration of 500 ng/ml was used throughout this thesis to induce the TRE promoter if not stated otherwise. In order to generate the p2loxHA-NeuroD1 vector, the coding sequence (CDS) for NeuroD1 was retrieved from NCBI (NM_010894.2) and custom synthesized at GeneArt. Next, the CDS was cloned into the p2loxHA-Nanog vector (data not shown), leading to the N-terminal fusion of a HA-tag with the NeuroD1 CDS. This p2loxHA-NeuroD1 vector was electroporated into A2lox.Cre mESCs expressing the Cre recombinase. After antibiotic selection, twelve of the obtained single stem cell clones were chosen for gDNA preparation and the Loxin PCR was carried out (Figure 32). This PCR screening revealed the successful generation of A2lox.NeuroD1 mESCs and three clones (clone 1, 6 and 9) were chosen for further validation.

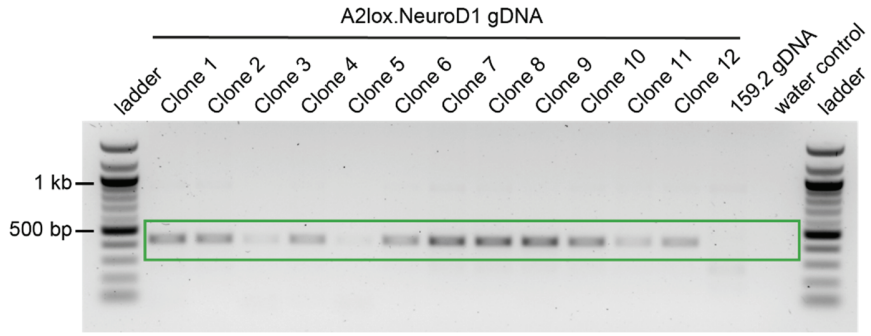


Figure 32: Loxin PCR reveals successful integration of the p2loxHA-NeuroD1 plasmid in selected A2lox.NeuroD1 mESC clones.
The Loxin PCR was performed on 100 ng gDNA of the 12 selected A2lox.NeuroD1 clones. The gel electrophoresis revealed PCR products at the expected size of 430 bp (green box) in the 1.2% agarose gel as identified with the 100 bp ladder. The PCR products for clone 3 and clone 5 were weak. The specificity of the PCR screen is evident as no PCR product was observed with the gDNA of 159.2 mESCs as a template (negative control, no RMCE site) or the water control. Data derived by JJ.

To confirm the conditional expression of HA-tagged NeuroD1 in these transgenic mESCs, a doxycycline (Dox) induction for 48 h was carried out in all three clones. Measuring the RNA level of conditionally expressed *HA-NeuroD1* (ectopic *NeuroD1*) by RT-qPCR revealed that the expression was specific to induced cells (Figure 33A). The overall expression of *NeuroD1*, measured with a different qPCR primer pair detecting the ectopically induced as well as endogenous *NeuroD1* RNA levels (total *NeuroD1*), increased concomitantly (Figure 33A). To ensure the specificity of *NeuroD1* expression also on protein level, a western blot analysis probing for the HA-tag was performed on A2lox.NeuroD1 mESCs comparing the induced (+Dox) and non-induced (-Dox) conditions of all three selected clones. In line with RNA expression, ectopic NeuroD1 was only detectable on protein level specifically after Dox induction at the expected protein size

(Figure 33B). The protein was slightly bigger than expected for endogenous NeuroD1 due to the fusion with the N-terminal HA-tag and possibly due to post-translational modifications. These three validated A2lox.NeuroD1 mESC clones were used as replicates for further experiments in this thesis. HA-NeuroD1 is referred to as ectopic NeuroD1 from here onwards.

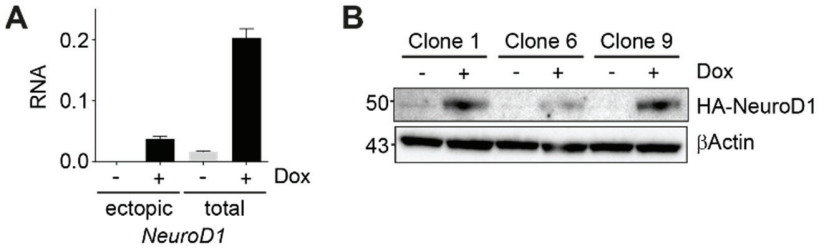


Figure 33: Ectopic NeuroD1 is detectable on RNA as well as protein level in A2lox.NeuroD1 mESCs 48 h post Dox induction.
(A) RT-qPCR analysis detected ectopic *NeuroD1* only after Dox induction (+Dox). Although a basal level of *NeuroD1* was detected in non-induced (-Dox) cells, the level of total *NeuroD1* increased after Dox treatment. RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*). The error bars reflect standard error of the mean from three independent A2lox.NeuroD1 mESC clones. (B) Western blot assay detected HA-NeuroD1 only after Dox treatment in three independent A2lox.NeuroD1 mESC clones. Equal loading was shown by probing for βActin. The protein size in the 10% acrylamide gel is shown in kDa on the left.
Figure information: A2lox.NeuroD1 mESCs were treated for 48 h with doxycycline to induce the expression of ectopic NeuroD1 (+Dox) or were left untreated (-Dox) as control. Data derived by JJ.

3.1.2.2 Expression kinetics of A2lox.NeuroD1 mESCs reveals a time-dependent induction of ectopic NeuroD1 on RNA as well as protein level

The previous induction experiments have been carried out as an endpoint analysis after 48 h. In order to investigate on the one hand after which duration ectopic NeuroD1 can be detected and on the other hand if there is a time dependent fluctuation in overall expression levels, a time course analysis was carried out in A2lox.NeuroD1 mESCs. Therefore, the cells were treated with doxycycline for several durations (1 h, 2 h, 6 h, 12 h, 24 h and 48 h) and first analysed for *HA-NeuroD1* expression on RNA level (Figure 34A). The RNA expression exhibited a time dependency: RNA of ectopic *NeuroD1* was detectable within 1 h of induction. The expression level increased until 12 h, accompanied by an increase of the expression level for total *NeuroD1* (Figure 34A). Although RNA levels of *NeuroD1* decreased until 48 h, the induction was still eminent (Figure 34A) in line with the previous results (Figure 33A). Following the RNA time course analysis, a western blot assay was performed to investigate the expression of HA-NeuroD1 at the same time points (Figure 34B). As the translation of RNA into protein occurs post-transcriptionally in the cytoplasm, it was expected that the HA-tagged protein would be detectable after RNA levels increased. Indeed, the HA-western blot showed that the ectopic NeuroD1 protein was specifically detectable only after 6 h and started to decrease between 24 h and 48 h (Figure 34B). This analysis suggested that chromatin immunoprecipitation experiments for ectopic NeuroD1 binding should be carried out during the induction phase and not at the endpoint after 48 h.

In conclusion, the A2lox system was validated to overexpress ectopic NeuroD1 in a conditional fashion. Therefore, the generated transgenic mESCs were used to further investigate the functional consequence of the increased ectopic NeuroD1 levels in stem cells as well as during their differentiation along the neuronal lineage.

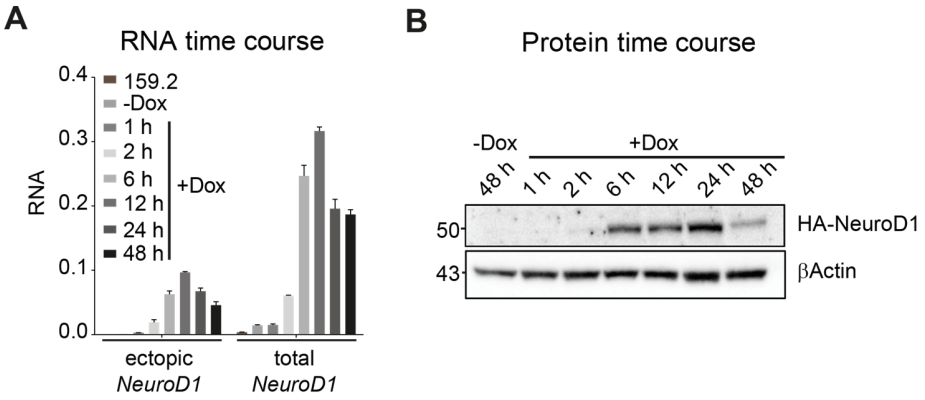


Figure 34: Ectopic NeuroD1 expression exhibits a time-dependent induction on RNA as well as protein level in A2lox.NeuroD1 embryonic stem cells.
(A) Time course RT-qPCR analysis of the doxycycline-induced expression dynamics of ectopic as well as total *NeuroD1*. RT-qPCRs were performed in biological replicates after the indicated duration of doxycycline treatment. Non-induced A2lox.NeuroD1 (-Dox; 48 h) as well as 159.2 mESCs were taken as negative controls. Ectopic *NeuroD1* was detectable after 1h of induction and increased until 12 h concomitant the total *NeuroD1* RNA levels. RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*), and the error bars indicate the standard error of the mean of two biological replicates. (B) Western blot time course assay for HA-NeuroD1 as well as βActin after Dox induction for different durations in A2lox.NeuroD1 mESCs. HA-NeuroD1 was detectable on protein level after 6 h and decreased after 24 h Dox treatment. The βActin plot served as loading control. The protein size in the 10% acrylamide gel is shown in kDa on the left. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.2 Ectopic expression of NeuroD1 is sufficient to induce the neuronal development program

3.2.1 Analysis of *NeuroD1* expression during *in vivo* and *in vitro* neurogenesis

Previous studies have shown that the expression of *NeuroD1* is endogenously induced at the onset of neurogenesis [37-39]. In order to get further insights into the expression profile of *NeuroD1*, several RNA-seq datasets of murine embryonic tissues representative of all three lineages including several layers of the embryonic cortex [40-42] were analysed. The cellular architecture of the murine embryonic cortex can be described by distinguishing three distinct layers spanning from the ventricle to the pial surface: the ventricular as well as subventricular zone (VZ/SVZ), which harbor proliferative progenitors, as well as the cortical plate (CP), which consists of newborn and maturating neurons [43-45]. Interestingly, the transcriptome analyses of these layers revealed that *NeuroD1* was highly expressed in the VZ and even several fold upregulated in the SVZ (Figure 35A). Its expression decreased in the cortical plate where

neurons mature. Additionally, *NeuroD1* was not expressed in any of the other investigated tissue except pancreas, where it was transcribed at low levels (Figure 35A).

The next question addressed was if similar *NeuroD1* expression dynamics could be observed during the neuronal differentiation of mESCs *in vitro*. To this end, a previously described differentiation protocol was adapted that generates around 95% pure neuronal progenitor (NP) cells ("radial glial-like" cells) from mESCs [16, 17]. These progenitors can subsequently become terminally differentiated pyramidal neurons (TN) and can be kept in culture for several days. This differentiation protocol has been demonstrated to exhibit highly synchronous and reproducible changes in the transcriptome as well as the epigenome during neuronal differentiation [9, 46-50]. Furthermore, this differentiation system can serve as a model to mimic *in vivo* development, as the behavior of individual genes and genomewide analyses of murine primary cortical neurons have been found to be in good agreement with this *in vitro* system in terms of the transcriptome as well as the epigenome [9, 49, 50]. Therefore, an expression analysis time course for *NeuroD1* was performed during *in vitro* neuronal differentiation of mESCs using a RT-qPCR assay (Figure 35B). This analysis showed that *NeuroD1* RNA levels were peaking immediately following the onset of neurogenesis also *in vitro*. In conclusion, *NeuroD1* transcription seems to be a specific feature of neurogenic progenitors and newborn neurons *in vivo* as well as *in vitro*.

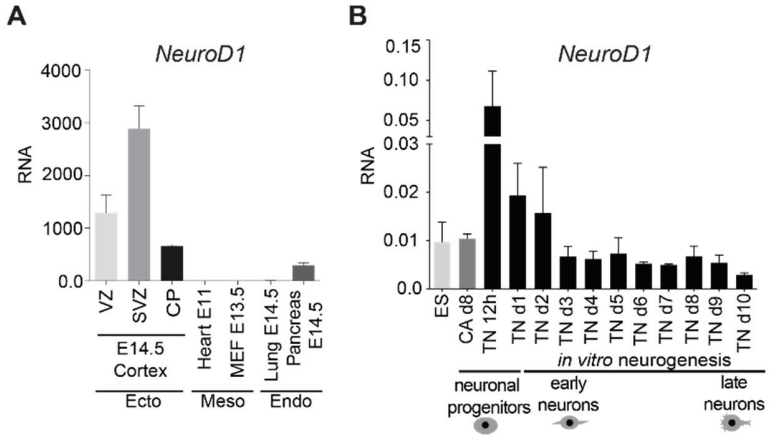


Figure 35: *NeuroD1* is highly expressed during the onset of neurogenesis *in vitro* as well as *in vivo*.
(A) Several RNA-seq datasets from *in vivo* embryonic samples of all three lineages were analysed for *NeuroD1* expression. NeuroD1 was only expressed in the embryonic cortex and at very low levels in the pancreas. RNA represents the normalized tag counts from biological replicates. (B) RT-qPCR analysis for *NeuroD1* expression in an *in vitro* neuronal differentiation time course. The bHLH factor *NeuroD1* was endogenously induced during the onset of neurogenesis. RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*) of three biological replicates.
Figure information: Error bars reflect standard error of the mean. VZ, ventricular zone; SVZ, subventricular zone; CP, cortical plate; Ecto, ectoderm; Meso, mesoderm; Endo, endoderm; ES, embryonic stem cells; CA, cellular aggregates; TN, terminally differentiated neurons. RNA-seq analysis performed by AP, *in vitro* data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.2.2 Ectopic NeuroD1 induction in embryonic stem cells leads to the formation of TUJ1-positive cells

Given the specific expression of NeuroD1 during the onset of neurogenesis and its demonstrated ability to reprogram cells towards neurons [7, 8, 51, 52], it was tempting to investigate whether NeuroD1 can differentiate pluripotent mESCs directly into neurons *in vitro*. In order to address this question, NeuroD1 was ectopically induced in the previously established A2lox.NeuroD1 mESCs. Strikingly, a immunocytochemistry analysis revealed that many cells expressed the neuron-specific protein TUJ1 and exhibited additionally a neuron-like morphology within 48 h of ectopic NeuroD1 induction (Figure 36A). The generation of neuron-like cells was further supported by a significant transcriptional increase of several neuronal marker genes, including TUJ1 (*Tubb3*) and Synaptophysin (*Syp*) specifically in cells ectopically expressing NeuroD1 (+Dox) in comparison to non-induced control cells (-Dox) as measured by RT-qPCR (Figure 36B). This induction of neuronal marker genes as a hallmark of neuronal differentiation suggested an exit from pluripotent state of mESCs. To address this, a RT-qPCR analysis of hallmark genes promoting and maintaining pluripotency (*Oct4*, *Sox2*, *Nanog* as well as *Klf4* [13, 53-58]) was carried out in the same system 48 h post ectopic NeuroD1 induction (Figure 36C). Strikingly, a downregulation of *Oct4*, *Nanog*, and *Klf4*, but not *Sox2* was observed. Taken together, these results suggested that NeuroD1 is able to induce neuronal differentiation of murine pluripotent stem cells.

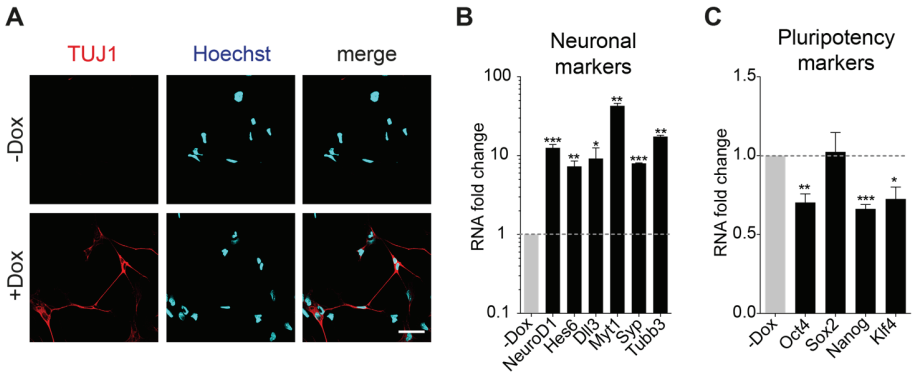


Figure 36: Ectopic expression of NeuroD1 is sufficient to initiate the generation of TUJ1-positive cells along with induction of neuronal marker gene expression.
(A) Immunocytochemistry of A2lox.NeuroD1 cells for TUJ1 (neuronal marker) and Hoechst (nucleus) after 48 h of ectopic NeuroD1 expression (+Dox) as well as for non-induced control cells (-Dox). TUJ1 expression and a neuron-like morphology were only found after doxycycline treatment. However, not all cells expressed this neuronal marker after 48 h. Scale bar: 10 µm. (B, C) RT-qPCR analysis for the expression of neuronal markers (B) and hallmark pluripotency genes (C) after 48 h of ectopic NeuroD1 induction in A2lox.NeuroD1 mESCs. The increase of neuronal markers upon ectopic NeuroD1 expression was accompanied by a reduction of the pluripotency markers *Oct4*, *Nanog* as well as *Klf4*, but not *Sox2*.
Figure information: RNA fold change reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*) plotted as the fold change of induced (+Dox) versus non-induced (-Dox) condition. The y-axis in (B) is plotted in a log10 scale. Error bars reflect standard error of the mean from three biological replicates. Significance was determined by t-test with * P < 0.05, ** P < 0.01 and *** P < 0.001. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.2.3 Global transcriptome profiling after NeuroD1 overexpression reveals induction of neuronal program

Tempted by the previous observations, the next step was to uncover the global gene regulatory circuitry through which NeuroD1 functions to promote neuronal fate. To this end, RNA-seq was performed on A2lox.NeuroD1 cells, both the 48 h NeuroD1-induced (+Dox) and non-induced (-Dox) population, in order to unravel the global changes in gene expression initiated either directly or indirectly by enhanced levels of ectopic NeuroD1. The RNA-seq analysis revealed 2,209 upregulated genes as well as 1,699 downregulated genes after applying stringent criteria for the significance and fold change (FC) cut-off values (FDR ≤ 0.05, FC ≥ 1.5) to the Next-Generation Sequencing (NGS) data (Figure 37A). From here onwards throughout the thesis the upregulated genes as well as downregulated genes are referred to as URG (see 8.1) and DRG (see 8.2), respectively. To get further insights into the nature of these differentially expressed genes and in which biological processes they are involved, an enrichment analysis for gene ontologies (GO terms) was performed. This type of analysis represents properties of the gene products in a given gene set, for example in which biological process (the ontology) they have already been described. As a result this analysis reveals which ontologies are over-represented or under-represented in a given gene set using annotation databases. Interestingly, the GO term analysis for the URG exhibited enrichments exclusively for biological processes related to a neuronal context like transmission of nerve impulse, synaptogenesis, neuronal differentiation or nervous system development (Figure 37B). These results were further supported by a similar, independent analysis for URG using the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway database, which also showed and enrichment for neurogenesis related pathways (Figure 37C).

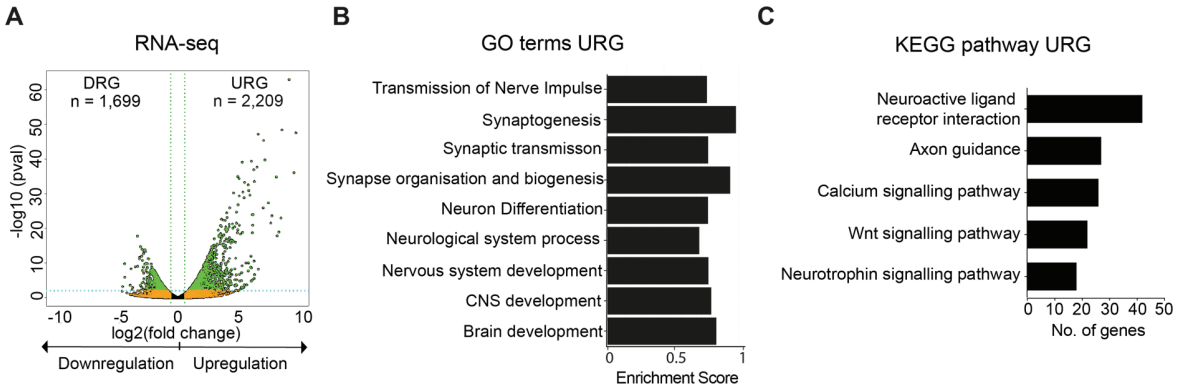


Figure 37: Ectopic NeuroD1 induction for 48 h induces neuronal developmental program.
(A) Volcano plot [x-axis: log2 (fold change +Dox versus -Dox), y-axis: -log10 (P-value)] depicting differential gene expression after 48 h of ectopic NeuroD1 induction in A2lox.NeuroD1 mESCs. In total 2,209 genes were upregulated (URG) and 1,699 genes were downregulated (DRG) following ectopic NeuroD1 expression. Data points marked in green represent genes crossing the significance cut-off for differential expression: abs(log2(fold change)) ≥ 0.58 (green dotted line) as well as P-value < 0.05 (blue dotted line). (B) Bar plot for the top GO terms for URG (n = 2,209) from (A). The plot showed enrichment for biological processes related to neuronal lineage. The bar length is determined by the enrichment score calculated by GSEA. (C) Bar plot showing that the KEGG pathways enriched in URG (n = 2,209) from (A) were linked to neurogenesis related context. Bar lengths represents number of genes contributing to pathway enrichment.

Figure information: DRG, downregulated genes; URG, upregulated genes; GSEA, Gene Set Enrichment Analysis. KEGG, Kyoto Encyclopedia of Genes and Genomes. Computational analyses performed by AP, data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

Furthermore, the most significant contributors to the neurogenesis-related GO terms were found to be also the highest upregulated genes (data not shown). Taken together, these observations suggested that NeuroD1 function is sufficient to override the pluripotent state of mESCs and promote their neuronal commitment. This bHLH factor induced the neuronal program despite the absence of neuron-promoting culture media and the presence of differentiation inhibiting factors (ES media promoting pluripotency).

This protruding enrichment of neurogenesis related genes among the ectopic NeuroD1-induced transcripts (URG) raised the question, whether the same genes are also endogenously induced during neuronal differentiation *in vitro* as well as *in vivo*. To this end, a RNA-seq analysis was performed on an *in vitro* time course where embryonic stem cells were differentiated towards neuronal lineage according to Bibel et al. [16, 17], as shown for the RT-qPCR measurement of *NeuroD1* in Figure 35B. Strikingly, this transcriptome analysis revealed that the identified URG were also mostly endogenously upregulated during the onset of neurogenesis *in vitro* (Figure 38A and B). Similarly, their transcription profile in the previously analysed *in vivo* tissues (see Figure 35A), showed also mostly an upregulation during the transition from neuronal progenitors to neurons. Additionally, these URG were considerably more expressed in the examined cortical layers than in the probed embryonic tissues of other lineages (Figure 38C and D). These data together showed that the expression of a large number of URG remained high in terminally differentiated neurons (Figure 38A - D). This observation suggested that genes upregulated upon NeuroD1 induction in mESCs are also essential for differentiation along neuronal lineage *in vitro* as well as *in vivo*, implicating NeuroD1 as a key driver of neuronal program. Furthermore, the NeuroD1-induced transcriptional state persisted following the transitory period of NeuroD1's peaked expression during neuronal development *in vitro* as well as *in vivo*, suggesting the maintenance of such state by other players or epigenetic memory.

As a huge number of genes were also downregulated upon NeuroD1 induction in mESCs (DRG), the question arose if this expression profile of DRG mimics their endogenous expression during neuronal development. The analysis of the previously used NGS data revealed that DRG were indeed endogenously downregulated during *in vitro* neurogenesis (Figure 38E and F) and were almost exclusively expressed in non-neuronal lineages *in vivo* (Figure 38G and H). This gene set was mostly enriched for non-neurogenesis related GO terms such as cell adhesion and metabolic processes (data not shown). Taken together, these observations showed that the ectopic expression of NeuroD1 is sufficient to induce a neuronal differentiation program that closely recapitulates neuronal development *in vivo*.

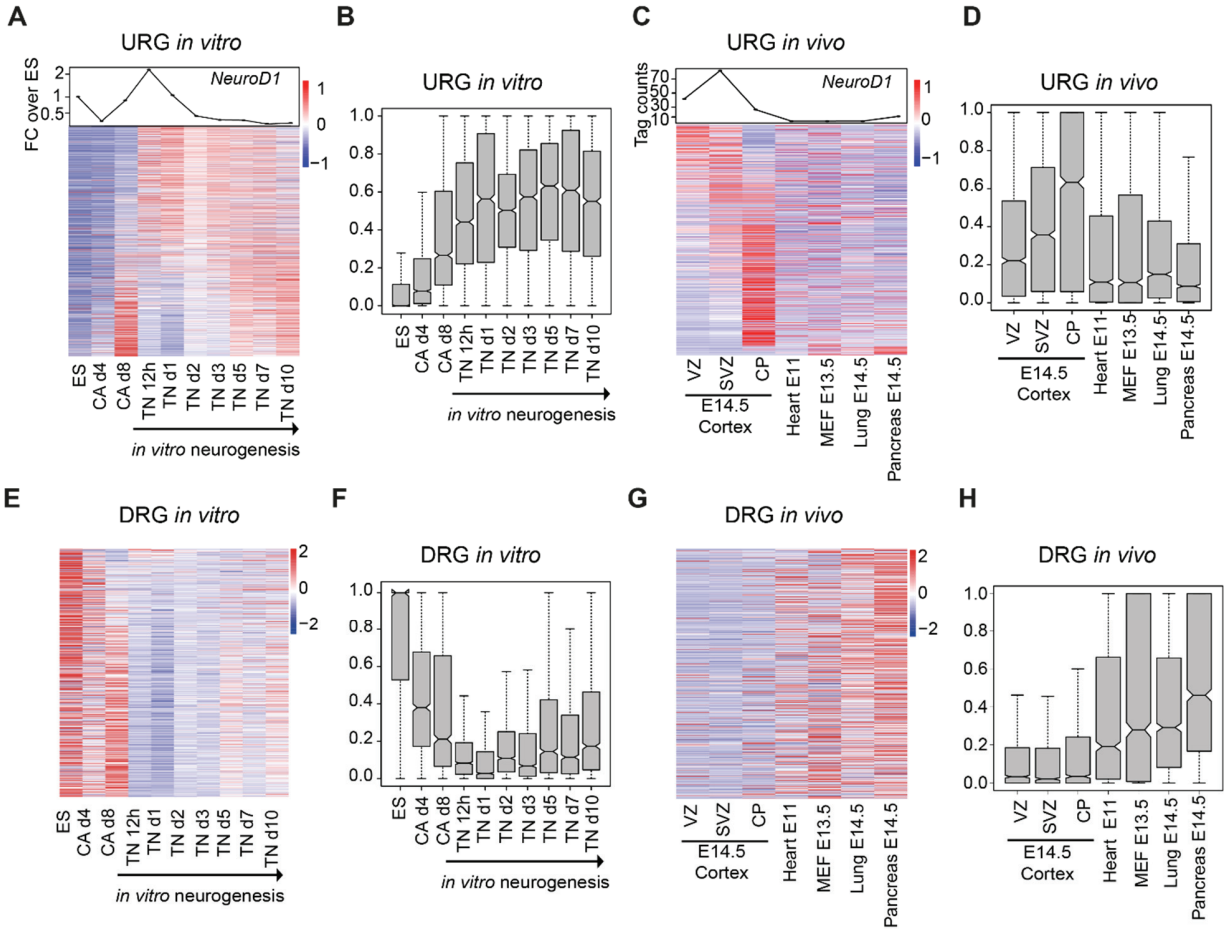


Figure 38: Differentially expressed genes upon ectopic NeuroD1 induction in mESCs recapitulate the endogenous transcription profile during neurogenesis *in vitro* as well as *in vivo*.
(A) Line plot showing the expression of *NeuroD1* as fold change over ES during *in vitro* neurogenesis (A, upper panel). The heatmap (A, lower panel) depicts the expression of URG in an *in vitro* neuronal RNA-seq time course using biological triplicates. Most URG were endogenously upregulated during neurogenesis *in vitro*. (B) Same information as in (A) but represented as a boxplot. (C) Line plot showing *NeuroD1* expression in various embryonic tissues (C, upper panel). The heatmap (C, lower panel) depicts the expression of URG in these embryonic tissues. Most URG were endogenously upregulated exclusively during *in vivo* neurogenesis from VZ to CP. (D) The same information as in (C) but represented as a boxplot. (E) Heatmap depicting the expression of DRG in an *in vitro* neuronal RNA-seq time course using biological triplicates. Most DRG were endogenously downregulated during *in vitro* neurogenesis. (F) Same information as in (E) but represented as a boxplot. (G) Heatmap depicting the RNA-seq expression of DRG in various embryonic tissues. Most DRG were not expressed in the embryonic cortex but in tissues of other lineages (H) Same information as in (G) but represented as a boxplot.
Figure information: Each row in the heatmaps (A, C, E and G) represents a gene where the RNA-seq expression value is scaled from blue (low expression) to red (high expression). The boxplots depicting RNA-seq data (B, D, F and H) contain expression values scaled between 0 and 1 on the y-axis. VZ, ventricular zone; SVZ, subventricular zone; CP, cortical plate; ES, embryonic stem cells; CA, cellular aggregates; TN, terminally differentiated neurons; DRG, downregulated genes; and URG, upregulated genes. Computational analyses performed by AP, *in vitro* data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.3 NeuroD1 directly targets regulatory elements of genes critical for neuronal development to induce their expression

The induction of the neuronal developmental program in mESCs by the overexpression of NeuroD1 raised the question to which extend such dramatic induction of cell fate change is directly linked to its DNA binding

ability. In order to identify genomewide targets of NeuroD1, a chromatin immunoprecipitation (ChIP) assay in combination with NGS (ChIP-seq) was performed. NeuroD1 was ectopically expressed in A2lox.NeuroD1 mESCs for 24 h and the ChIP assay was executed with an antibody directed against its N-terminally fused HA-tag. The precipitated material was subjected to NGS followed by computational analyses.

3.3.1 Ectopic NeuroD1 directly binds to regulatory elements

As a first quality control, the browser tracks of the NeuroD1 ChIP-seq data were visually inspected for the appearance of regions with ectopic NeuroD1 enrichment ("peaks"), meaning that NeuroD1 is binding to these genomic regions. This data revealed that on the one hand NeuroD1 was directly targeting to gene promoters (Figure 39A). On the other hand, also distinct genomic sites including intergenic regions were found to be enriched for ectopic NeuroD1 (Figure 39B). This first manual inspection of the NeuroD1 ChIP-seq data was further extended to a global computational analysis which quantified that NeuroD1 binding occurred predominantly at both, the promoter and non-promoter regions (in total n = 2,409) (Figure 39C). Given the small size promoters constitute in the entire genome (around 2%), the peak distribution was normalized to the genomic size each particular genomic class (promoters and the three non-promoter classes of exons, introns and intergenic regions) exhibits in the genome. This analysis demonstrated that ectopic NeuroD1 was preferentially targeting to gene promoters (around 39%, n = 341) (Figure 39C).

To further study the function of NeuroD1 on chromatin level, the advanced analyses were focused on genomic target sites encompassing promoters as well as enhancers, as these are prominent regulatory elements in the genome. It has been well described in previous studies that the acetylation of lysine 27 at histone H3 (H3K27ac) is a hallmark of active enhancers [59-63]. In order to determine potential non-promoter regulatory regions (potential enhancers) which are bound by NeuroD1 but also linked to neurogenesis, an additional selection criteria of enrichment for H3K27ac during the onset of neurogenesis was applied to the already identified non-promoter peaks (see Figure 39C). Therefore, H3K27ac ChIP was performed at an early stage of the *in vitro* neuronal differentiation system of mESCs (TN d1, a time point immediate to the highest *NeuroD1* expression *in vitro*). The precipitated material was subjected to NGS and further analysed for identified NeuroD1-bound non-promoter sites which are also enriched for the H3K27ac modification in this additional data. To get insight into how many of the upregulated genes upon NeuroD1 induction in mESCs (URG) were also targeted by NeuroD1 at a potential regulatory element, all non-promoter peaks were first associated to the gene in their closest vicinity (nearest gene approach). The following comparison of the 2,209 URG (see Figure 37A) with the genes associated to non-promoter target sites as well as NeuroD1-bound promoters revealed that a significant number of genes induced upon ectopic NeuroD1 induction in mESCs were directly bound by NeuroD1 at their potential regulatory elements (~25%) (data not shown). Additionally, the identified non-promoter NeuroD1 regulatory elements were furthermore subdivided into exonic, intronic, and intergenic enhancers (data not shown). Interestingly, NeuroD1-bound intronic and exonic enhancers were associated with URG which were annotated with neurogenesis related GO terms and exhibited transcriptional induction upon neuronal differentiation *in vitro* as well as *in vivo* (data not shown). However, further functional analyses of non-promoter regulatory elements were focused

exclusively on intergenic regulatory regions in order to avoid a potential influence of the genic chromatin landscape and other transcriptional regulatory events such as transcriptional elongation that occur in gene bodies. These intergenic non-promoter NeuroD1 peaks enriched for H3K27ac *in vitro* are referred to as "enhancers" from here onwards. Taking the NeuroD1-bound promoters as well as the genes associated to NeuroD1-bound enhancers and overlapping them with the URG (see Figure 37A) revealed, that a significant number of URG were targeted by NeuroD1 at their promoters (n = 83), enhancers (n = 107), or both elements (n = 5) (Figure 39D). These 195 upregulated target genes of ectopic NeuroD1 are referred to as URT (see 8.3 and 8.4) from here onwards. Interestingly, these URT as direct NeuroD1 targets where significantly higher upregulated in comparison to the transcriptional induction of URG overall after 48 h of ectopic NeuroD1 expression in mESCs (data not shown). As the NeuroD1 target enhancers had been identified by *in vitro* data so far, they were compared to a dataset describing murine E14.5 cortex- or brain-specific enhancers [41]. Strikingly, these correlation revealed that a significant fraction of these enhancers were indeed already identified *in vivo* (data not shown), supporting the relevance of the previous observations. This was further strengthened by the expression profile of URT during both, *in vitro* (Figure 39E and F) as well as *in vivo* neurogenesis (Figure 39G and H), as the majority of these identified NeuroD1 associated targets were transcriptionally upregulated in this context and almost exclusively expressed in cortical layers but not in tissues of other lineages (Figure 39G and H).

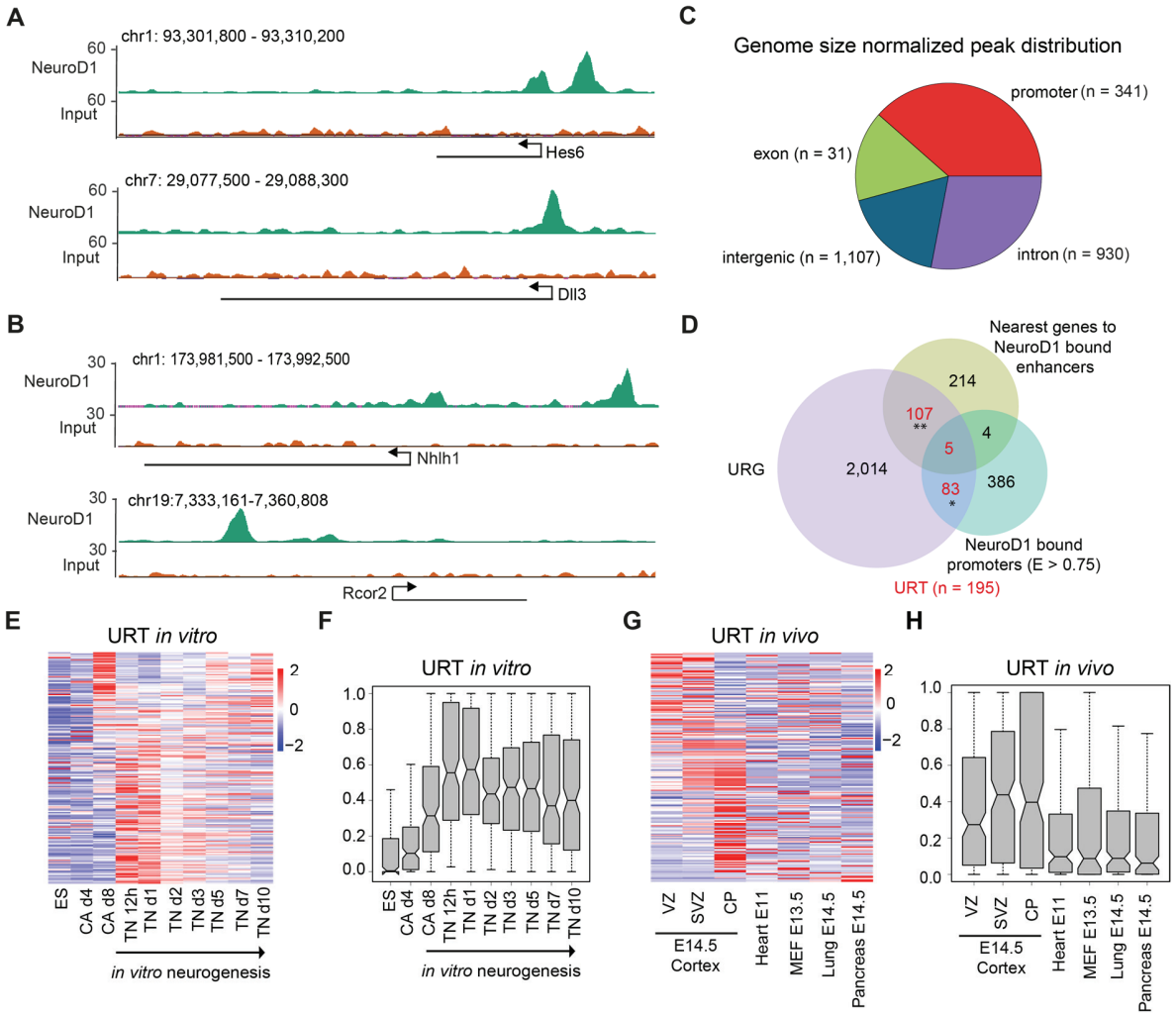


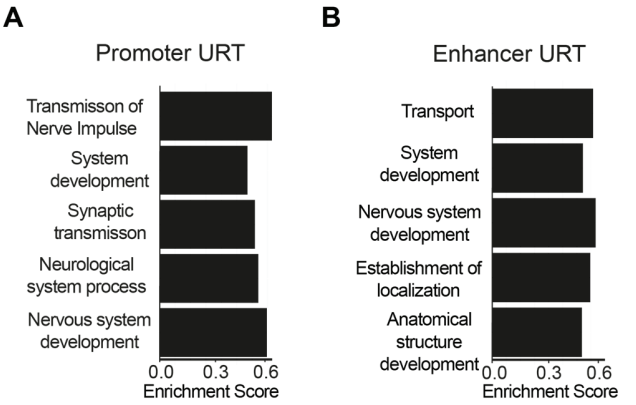
Figure 39: NeuroD1 directly binds to *cis*-regulatory elements of genes critical for neuronal development.
(A, B) Screenshots of the UCSC genome browser for the HA-NeuroD1 ChIP-seq (ChIP and Input) in A2lox.NeuroD1 cells 24 h post doxycycline induction (+Dox). The plots showed that ectopic NeuroD1 was enriched at representative target promoters (A) or target enhancers (B) as visible from the eminent green wave-like shaped structures ("peaks"). The merged wiggle files were generated from two biological replicates and the baseline on y-axis represents "0" values. The highlighted genes are indicated as black arrows representing the direction of transcription and therefore the location of the associated promoter with the TSS. (C) Pie chart depicting the distribution of NeuroD1 peaks (n = 2,409) reproducibly called in two biological replicates. The peaks were distributed for particular genomic classes (promoters, intergenic regions, exons, and introns) and normalized by genome size. The absolute number of peaks is shown in the inset. After genome size normalization, ~39% (absolute number = 341) of total peaks were called at promoters, ~28% (absolute number = 930) at intronic, ~18% (absolute number = 1,107) at intergenic, and ~16% at exonic (absolute number = 31) regions. (D) Venn diagram showing the intersection of URG (n = 2,209) with NeuroD1-bound promoters (E > 0.75, n = 478) as well as the genes (n = 330) associated with NeuroD1-bound enhancers (top 500 enriched sites). In total 195 URG were bound by ectopic NeuroD1 at their regulatory elements (referred to as upregulated target genes, URT). Significance was determined by Fischer's test with * P < 0.05 and ** P < 0.01. (E) Heatmap depicting the expression of URT in an *in vitro* neuronal RNA-seq time course using biological triplicates. Most URT were endogenously upregulated during neurogenesis. (F) Same information as in (E) but represented as a boxplot. (G) Heatmap depicting the expression of URT in various embryonic tissues. Most URG were endogenously upregulated exclusively during *in vivo* neurogenesis from VZ to SVZ/CP. (H) The same information as in (G) but represented as a boxplot.
Figure information: Each row in the heatmaps (E, G) represents one gene whose RNA-seq expression value is scaled from blue (low expression) to red (high expression). Boxplots depicting RNA-seq data (F and H) contain expression values scaled between 0 and 1 on the y-axis. VZ, ventricular zone; SVZ, subventricular zone; CP, cortical plate; ES, embryonic stem cells; CA, cellular

aggregates; TN, terminally differentiated neurons; URT, upregulated target; TSS, transcriptional start site. Computational analyses performed by AP, *in vitro* data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.3.2 Ectopic NeuroD1 binds to neurogenesis related genomic elements in a sequence specific manner

The previous observations were further supported by a GO term analysis for the promoter and enhancer URT as the observed ontologies were enrichment for neurogenesis and development related biological processes (Figure 40A and B). The next question addressed was to which extent NeuroD1 targets genomic loci in a sequence specific manner. Towards this, a motif enrichment analysis of genomic sequences underlying the observed NeuroD1 peaks was carried out. Interestingly, the NeuroD1 motif was among the top three enriched motifs at the target gene promoters (Figure 40C) as well as target enhancer elements (Figure 40D), suggesting a sequence-dependent targeting of ectopic NeuroD1. Furthermore, a *de novo* motif prediction revealed that most ectopic NeuroD1 peaks (~95%) exhibit an E-box motif directly at the peak summit (data not shown), which has been described to be associated with typical members of the bHLH protein family [4].

Taken together, these observations suggested that ectopic NeuroD1 directly targets promoters as well as enhancers linked to neurogenesis in a sequence specific manner.



C

Top motifs at promoter targets

Rank	Motif	Name	p-val	% of targets
1		Atoh1	1e-4	77.78
2		NeuroD1	1e-3	62.90
3		Olig2	1e-3	85.19

D

Top motifs at target enhancer sites

Rank	Motif	Name	p-val	% of targets
1		Atoh1	1e-8	63.5
2		Olig2	1e-8	71.5
3		NeuroD1	1e-7	61.0

Figure 40: Ectopic NeuroD1 binds to neurogenesis related target elements in a sequence specific manner.
(A, B) Bar plots showing the top five GO terms for URT regulated at promoters (A) or enhancers (B). The analysis showed an enrichment for development- and neurogenesis-related GO terms among URT. The bar length is based on the enrichment score determined by GSEA. (C, D) Tables depicting the top three enriched motifs at ectopic NeuroD1-bound promoter (C) and enhancer (D) sites.
Figure information: URT, upregulated target genes. GSEA, Gene Set Enrichment Analysis. Computational analyses performed by AP. All panels are taken from Pataskar & Jung et al., 2016.

3.4 NeuroD1 induces the expression of transcription factors involved in neuronal development and cellular migration

The RNA-seq profiling of A2lox.NeuroD1 cells +/- ectopic NeuroD1 expression suggested that although a large number of neuronal genes are induced upon doxycycline treatment (see Figure 37), many of them are not directly occupied by ectopic NeuroD1 at their associated regulatory elements based on the applied cut-offs (see Figure 39D). This led to the question whether NeuroD1 might induce the expression of additional transcriptional regulators that could then possibly mediate the observed expression response of non-target genes. To address this question each gene either coding for a transcription factor (TFs) or an epigenetic regulator (EPI: exhibiting a protein domain known to act on chromatin apart from DNA binding) was considered for the ontology of transcriptional regulators.

3.4.1 NeuroD1 induces the expression of several transcriptional regulators naturally upregulated during neurogenesis

A deeper investigation of the 195 URTs (see Figure 39D) indeed exposed a number of induced transcriptional regulators being either occupied by ectopic NeuroD1 at promoter (promoter URT: n = 27/88) or associated enhancer elements (enhancer URT: n = 30/112). Interestingly, the TF/EPI promoter URTs were induced during *in vitro* (Figure 41A) as well as *in vivo* neurogenesis and were largely repressed in non-neuronal lineages (Figure 41B). Similar results were observed for the TF/EPI enhancer URTs *in vitro* (Figure 41C) as well as *in vivo* (Figure 41D). These identified candidate genes contained several regulators which were already described in the context of neurogenesis like *Pou3f2* or *Sox11* (see 1.5.2). This observation indicated that NeuroD1 was indeed inducing a network of transcriptional regulators that could further enhance or maintain the initiated neuronal developmental program. Remarkably, among the identified candidate TF/EPIs were a number of factors which have not been explored in the context of neurogenesis (e.g. *Zfand5*, *Rnf182*, or *Aff3*).

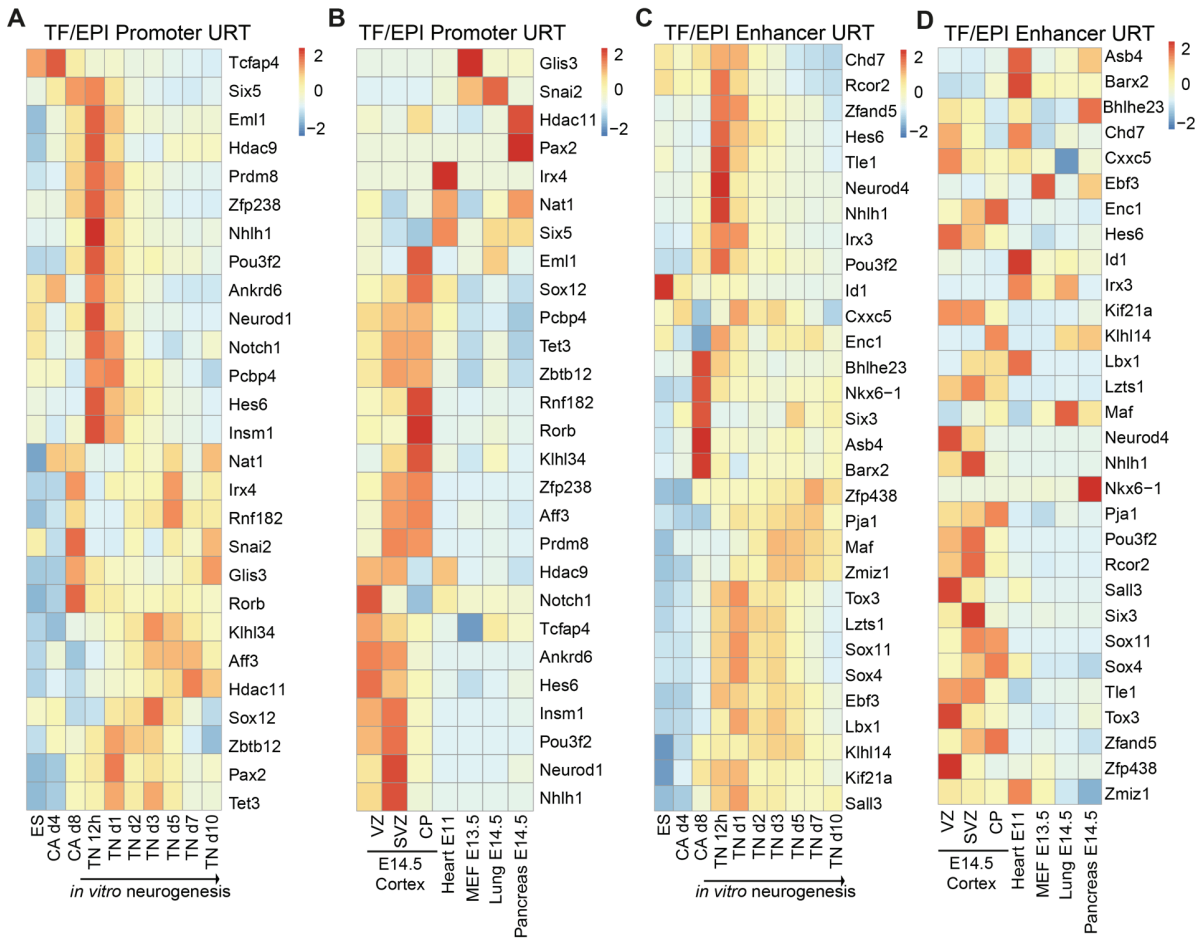


Figure 41: NeuroD1 targets encompass several transcriptional regulators induced during neurogenesis.
(A - D) Heatmaps depicting the RNA-seq expression of transcriptional regulators which belong to the group of NeuroD1-induced and regulated transcripts (URT, see Figure 39D). These induced regulators were divided into promoter URT, which were directly

enriched for ectopic NeuroD1 at their promoters, and enhancer URT, which were enriched for ectopic NeuroD1 at an associated intergenic enhancer element. The probed transcriptional regulators were mostly induced during neurogenesis *in vitro* as well as *in vivo* and were largely repressed in non-neuronal lineages. (A) Heatmap depicting the expression of TF/EPI promoter URT during *in vitro* neurogenesis. (B) Heatmap depicting the expression of TF/EPI promoter URT in various embryonic tissues. (C) Heatmap depicting the expression of TF/EPI enhancer URT during *in vitro* neurogenesis. (C) Heatmap depicting the expression of TF/EPI enhancer URT during *in vitro* neurogenesis. (D) Heatmap depicting the expression of TF/EPI enhancer URT in various embryonic tissues.
Figure information: RNA-seq data was derived from three (A, C) or two (B, D) biological replicates. Each row of the heatmaps represents one gene whose RNA-seq expression value is scaled from blue (low expression) to red (high expression). EPI, epigenetic regulator; TF, transcription factor; URT, upregulated targets. VZ, ventricular zone; SVZ, subventricular zone; CP, cortical plate; ES, embryonic stem cells; CA, cellular aggregates; TN, terminally differentiated neurons. Computational analyses performed by AP, *in vitro* data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.4.2 NeuroD1 induces expression of neurogenesis and epithelial-to-mesenchymal transition associated genes.

The analysis on NeuroD1-induced targets was further extended to independently validate their expression profile in the developing murine cortex by examining *in situ* hybridization (ISH) images that were derived independently of the analysed RNA-seq data. The ISH methodology enables to visualize the localization of a certain transcript of choice directly in a section of primary tissue due to the combination of microscopy imaging and labeled nucleic acid probes [64]. The Allen Brain Atlas (http://brain-map.org) provides a publicly available database containing e.g. ISH images for the murine and human brain. This resource was used to exemplify the neurogenesis related expression pattern of many identified NeuroD1 targets (Figure 42A – F). To further address the question whether the expression profile of these NeuroD1 target genes correlates with the localization of the *NeuroD1* transcript itself during cortical development, available cortical ISH images of several developmental stages from the Allen Brain Atlas were investigated. Interestingly, the available data for *NeuroD1* and one of its target genes (*Lzts1*) suggested that indeed target gene expression is induced in cells of advanced ontogeny stages which had been exposed to *NeuroD1* earlier in development (Figure 42G): In early cortical developmental, *Lzts1* was expressed in the same germinal layer as *NeuroD1*. With progressive ontogeny, *Lzts1* expression was maintained in the cortical plate where cells reside that originated from the *NeuroD1*-expressing cortical layers during earlier development.
The neuronal cells residing in the CP originate from the VZ/SVZ layers and they radially migrate to this destination from the VZ/SVZ during embryonic neurogenesis. Interestingly, although NeuroD1 has been implicated in the context of neuronal migration [65], the underlying molecular mechanism is largely unknown. In order to investigate this further, the following question addressed was whether classical migration genes involved in epithelial–mesenchymal transition (EMT) were induced following NeuroD1 expression. To this end, a recently identified gene set described as upregulated during EMT [66] was analysed for the overlap with genes upregulated by ectopic NeuroD1 induction in A2lox.NeuroD1 cells (URG, see Figure 37B). This analysis showed that URG include a large number of genes that were upregulated during EMT (n = 878, ~40%) and encompassed significant number of hallmark genes like *Cdh2*, *Ncam1*, *Snai1*, *Snai2* or *Twist1*, which are known to promote EMT [66-69] (Figure 42H).
It was intriguing to investigate whether these EMT-linked URG contained also direct targets of NeuroD1. Interestingly, the overlap with URT revealed that several of these genes, including *Ncam1* and *Snai2*,

exhibited ectopic NeuroD1 occupancy at their regulatory elements (n = 87, ~10%) (Figure 42I). These findings suggested that NeuroD1 may induce neuronal migration by activating the transcriptional induction of critical EMT genes essential for mesenchymal, migratory identity.

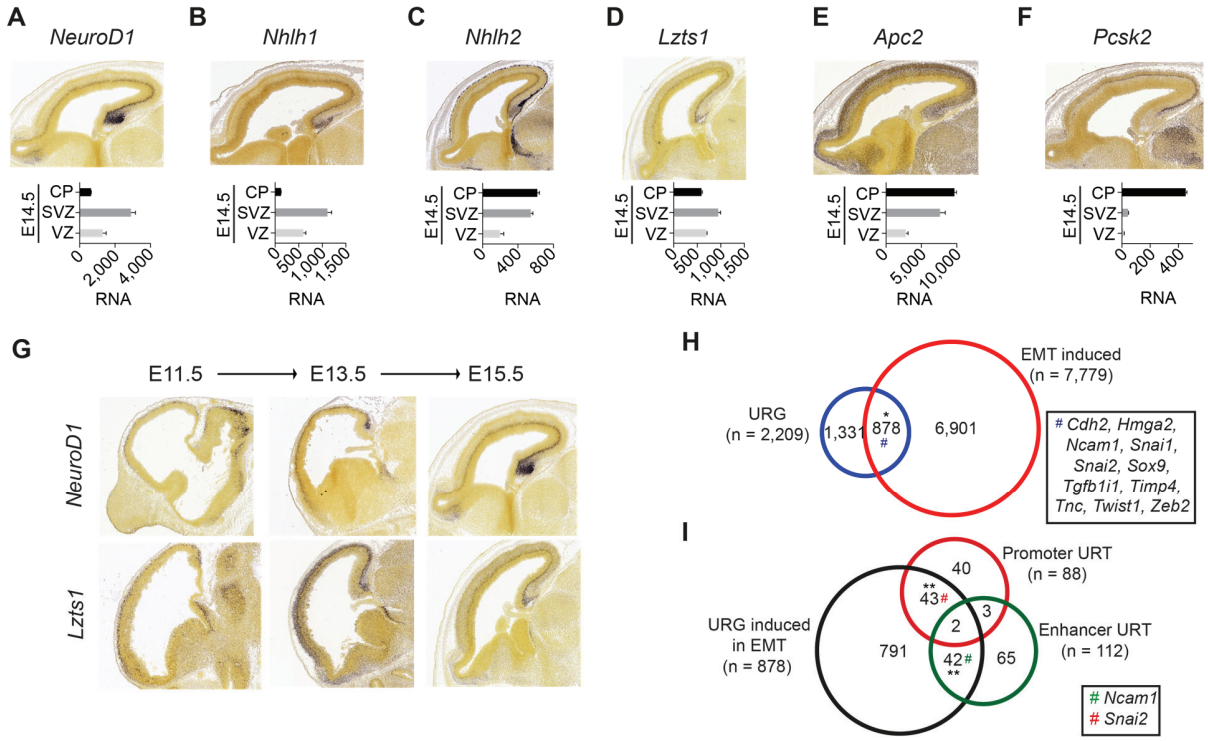


Figure 42: NeuroD1 induces expression of target genes involved in neurogenesis and epithelial-to-mesenchymal transition.
(A - F) (Upper panel) E15.5 cortex *in situ* hybridization images of antisense probes from the Allen Brain Atlas depicting the localization of *NeuroD1* (A) and representative target genes [*Nhlh1* (B), *Nhlh2* (C), *Lzts1* (D), *Apc2* (E), and *Pcsk2* (F)]. (Lower panel) Bar plots showing normalized tag counts (RNA) from RNA-seq analyses for the corresponding genes in E14.5 cortical layers. NeuroD1 was highly expressed in the SVZ whereas target genes were induced in the SVZ/CP. (G) *In situ* hybridization images of antisense probes from the Allen Brain Atlas showing the expression of *NeuroD1* and *Lzts1* during several stages of cortical development (E11.5, E13.5 and E15.5). Cortical layers exhibiting *Lzts1* mRNA originated from developmentally younger cellular layers positive for *NeuroD1* mRNA. (H) Venn diagram depicting the overlap between genes that are upregulated during murine EMT (FC > twofold at any stage, n = 7,779) and URG (n = 2,209, overlap with EMT induced genes: ~40%). The overlap of n = 878 genes contained several hallmark regulators of EMT induced in mesenchymal fate as exemplarily shown as inset. (I) Venn diagram depicting the significant overlap (~10% of URG induced in EMT) between URG that were also induced during murine EMT (n = 878) with promoter (n = 88) and enhancer URT (n = 112). The overlapping set of NeuroD1-bound genes and genes transcriptionally induced during EMT contained hallmark EMT regulators like *Ncam1* and *Snai2* as shown as inset.
Figure information: Error bars reflect standard error of the mean from two biological replicates. Significance was determined by Fischer's test with *P < 0.05 and **P < 0.01; URG, upregulated genes; URT, upregulated targets; EMT, epithelial-to-mesenchymal transition; VZ, ventricular zone; SVZ, subventricular zone; CP, cortical plate. Computational analyses performed by AP, ISH pictures derived from the Allan Brain Atlas (http://brain-map.org) by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.5 Selection of NeuroD1 target regions for further functional studies

3.5.1 NeuroD1 bound targets are cortex-specifically expressed *in vivo*

The global NGS analysis of the NeuroD1 ChIP-seq data revealed a set of direct target regulatory elements which were associated to 195 genes (URT). To study the consequence of NeuroD1 binding at its target

sites, a subset of genes was chosen for further functional investigations. To this end, the identified target genes were further shortlisted based on three criteria: (1) Being previously described in the context of neurogenesis, (2) exhibiting a cortex specific *in vivo* expression profile in E14.5 tissues, as well as (3) independently from the analysed NGS data being validated for NeuroD1 enrichment in a ChIP-qPCR assay. Several URTs were shortlisted based on existing literature (data not shown). In order to measure the expression of pre-selected NeuroD1-regulated neuronal genes, several tissues were dissected from E14.5 murine embryos and subjected to RT-qPCR analysis (Figure 43A - H). Strikingly, three promoter targets (*NeuroD1*, *Hes6* and *Dll3*) (Figure 43A - C) as well as three target genes bound at associated enhancer sites (*Ncam1*, *NeuroD4* and *Nhlh1*) (Figure 43D - F) exhibited a cortex-specific expression profile among the analysed tissues. In comparison to these neuron-specific target genes, the tested control genes (*Tbp* and *Ctcf*) were more uniformly expressed (Figure 43G - H). To validate NeuroD1 occupancy at these selected target sites, a ChIP-qPCR assay for ectopic NeuroD1 was carried out in A2lox.NeuroD1 cells after 24 h of doxycycline induction. Both, the promotor (Figure 43I) and the enhancer target sites (Figure 43J) were occupied by ectopic NeuroD1 whereas control regions for both target classes were not enriched for NeuroD1. For further validation in an additional cellular context, A2lox.NeuroD1 mESCs were differentiated into early neurons and a ChIP-qPCR assay was performed after 24 h of ectopic NeuroD1 induction. Importantly, all of the above identified NeuroD1 target regions were also occupied in neuronal cells (Figure 43K). Strikingly, these data suggested that NeuroD1 is able to bind and regulate its neuronal targets in a non-neuronal chromatin context.

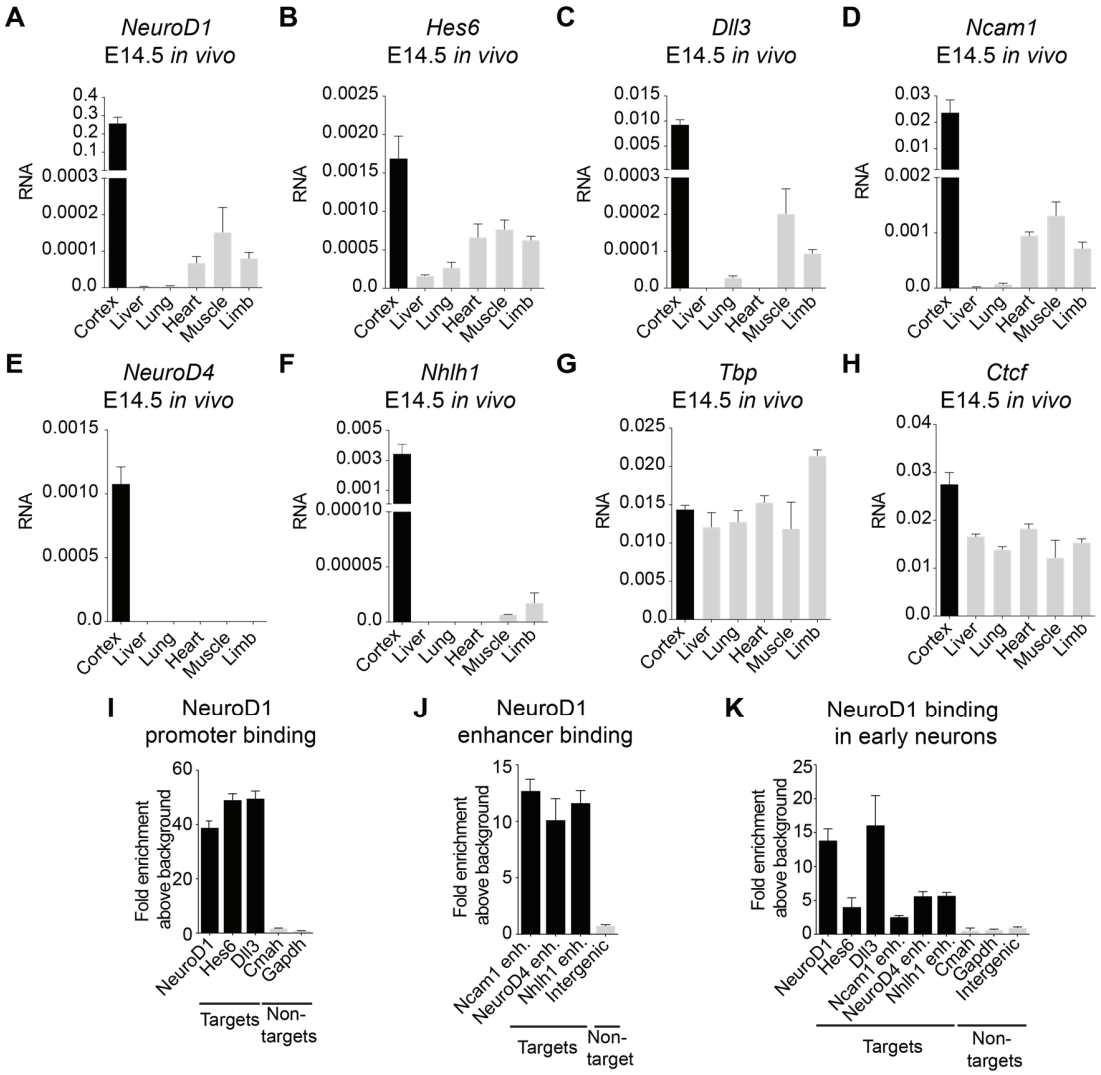


Figure 43: Selected NeuroD1-bound targets are cortex-specifically expressed *in vivo*.
(A - F) RT-qPCR analysis for the expression of selected NeuroD1 target genes in E14.5 tissues. The three tested promoter targets [*NeuroD1* (A), *Hes6* (B) and *Dll3* (C)] as well as the three genes associated with enhancer target sites [*Ncam1* (D), *NeuroD4* (E) and *Nhlh1* (F)] were almost exclusively expressed in the embryonic cortex. (G, H) RT–qPCR analysis for the expression of control genes in E14.5 tissues. The control genes *Tbp* (G) and *Ctcf* (H) were almost uniformly expressed among the tested embryonic tissues. (I) ChIP-qPCR analysis of ectopic NeuroD1 occupancy at selected target (black) and control (grey) promoters in A2lox.NeuroD1 mESCs after 24 h of induction. Ectopic NeuroD1 was specifically enriched at target promoters. (J) ChIP-qPCR analysis of ectopic NeuroD1 occupancy at target enhancers (black) and additional intergenic control region (grey) in A2lox.NeuroD1 cells after 24 h of induction. Ectopic NeuroD1 specifically occupied target enhancers. (K) ChIP-qPCR analysis of NeuroD1 occupancy at target promoters and enhancers (black) as well as control regions (grey) in early neurons derived *in vitro* (TN d1) from A2lox.NeuroD1 mESCs. Ectopic NeuroD1 was induced 24 h prior to the ChIP. Ectopic NeuroD1 was specifically enriched at all target elements.
Figure information: Error bars reflect standard error of the mean from three (A - H) or two (I - K) biological replicates. RNA in figure A - H reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*). The y-axis of ChIP-qPCR results (I - K) shows the relative ChIP enrichment plotted as the ratio of precipitated DNA (bound) to input DNA and further normalized to an intergenic control region (fold enrichment above background). Data derived by JJ. Panels (I - K) are taken from Pataskar & Jung et al., 2016.

3.5.2 NeuroD1 occupies identified target regulatory regions and induces the expression of associated genes also in murine fibroblasts

The findings above showed that NeuroD1 can occupy its neuron specific target elements in embryonic stem cells and induce the expression of their associated genes. However, these pluripotent cells that are able to differentiate into cells of all three germ layers upon an adequate stimulus [70, 71] display overall a more open chromatin structure [3, 72-74] that might allow easier access of ectopically expressed transcription factors to their target regions. Therefore, the next question addressed was to analyse whether ectopic NeuroD1 could also bind to its specific targets in terminally differentiated cells from another lineage such as fibroblasts and could concomitantly induce expression of these neuronal targets. To this end, the murine fibroblast cell line NIH/3T3 was first transfected with the pHAFlagNeuroD1 vector to transiently overexpress HA-FLAG-tagged NeuroD1 for 48 h. A ChIP-qPCR assay using an antibody directed against this FLAG-tag was performed and revealed that ectopically expressed HA-FLAG-NeuroD1 was indeed enriched at previously identified neuronal target regions also in the chromatin context of terminally differentiated fibroblasts (Figure 44A). Intrigued by this finding, the effect of NeuroD1 overexpression on the transcriptional state of such bound target genes was examined next in NIH/3T3 cells. To this end, fibroblast were transfected either with the pCIDRE-NeuroD1 vector, encoding NeuroD1 as well as a red fluorescent protein (RFP) separated by an internal ribosomal entry site (IRES), or a vector without the additional CDS for NeuroD1 (pCIDRE) as a control. The transfected cells were purified based on the expression of their fluorescent marker (RFP) by FACS and together with non-transfected NIH/3T3 cells (WT) subjected to expression profiling by RT-qPCR 48 h post transfection (Figure 44B - I). Strikingly, the majority of tested neuronal genes bound by NeuroD1 at their associated regulatory regions in these fibroblast (see Figure 44A) were also transcriptionally activated upon ectopic NeuroD1 expression, except *NeuroD4* (Figure 44B - G). Interestingly, also the neuronal marker *Tubb3* was transcriptionally induced (Figure 44H), but not the housekeeping gene *Tbp* (Figure 44I).

These observations suggested that NeuroD1 is able to trigger the activation of neuronal developmental genes by directly binding to their regulatory elements irrespective of cell type and the pre-existing chromatin landscape.

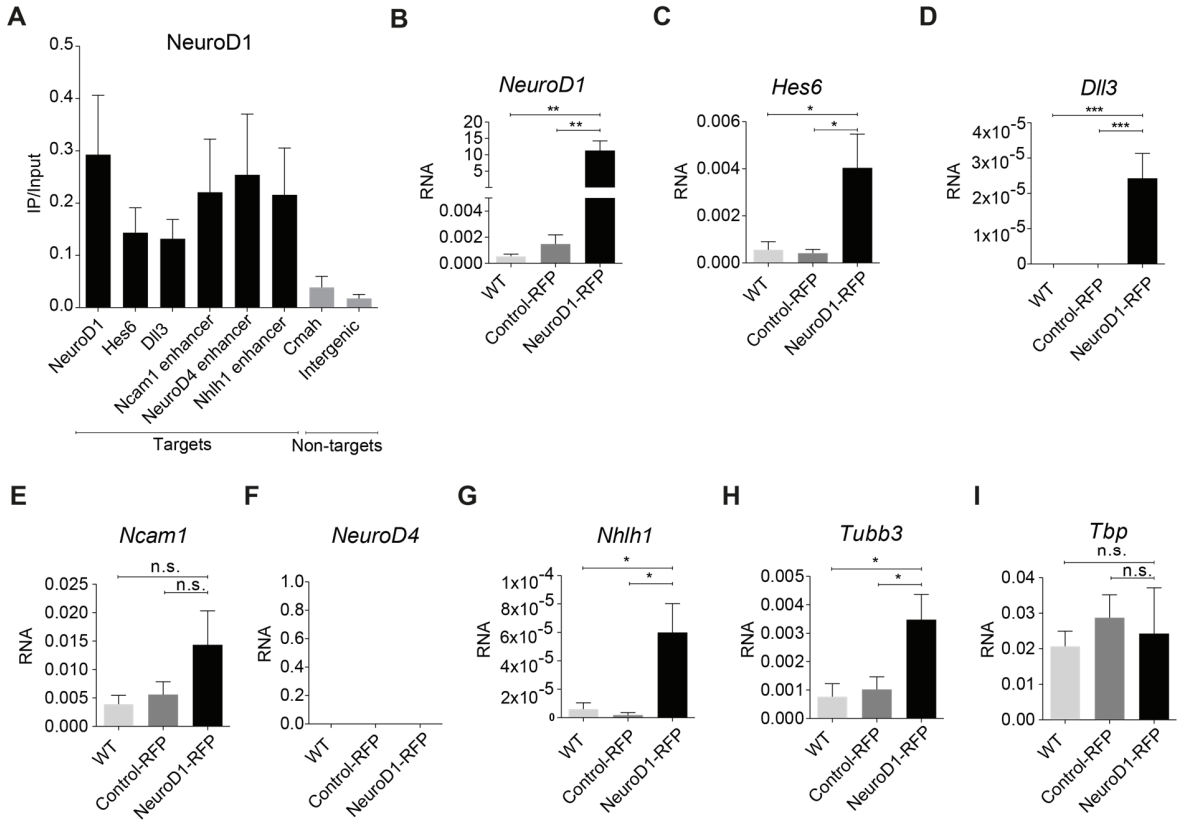


Figure 44: NeuroD1 binds its targets and induces transcription of associated genes in murine fibroblasts.
(A) ChIP-qPCR analysis for the NeuroD1 occupancy at target sites (black) and control regions (grey) in NIH/3T3 fibroblasts after transient transfection of pHAFlagNeuroD1 for 48 h. Target regions were specifically enriched for ectopic HA-FLAG-tagged NeuroD1. The y-axis shows the ChIP DNA normalized to the respective input DNA (IP/Input). (B - I) RT-qPCR analysis for the expression of genes associated to NeuroD1 target regions (B - G), the neuronal marker *Tubb3* (H) and a housekeeping control *Tbp* (I). The analysis was carried out on FAC-sorted NIH/3T3 cells after transfection with pCIDRE-NeuroD1 (NeuroD1-RFP) or pCIDRE (Control-RFP) for 48 h or in non-transfected control cells (wild type, WT). The majority of NeuroD1 associated genes was transcriptionally induced, accompanied by the upregulation of *Tubb3*, whereas *Tbp* transcription was not influenced by the overexpression of NeuroD1. RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*).
Figure information: Error bars indicate standard error of the mean from four biological replicates. Significance was determined by t-test with * P < 0.05, ** P < 0.01 and *** P < 0.001. n.s., not significant. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.6 NeuroD1 functions by reprogramming the chromatin and transcription factor landscapes at target gene promoters

The previous results suggested that NeuroD1 targets and activates neuronal developmental genes. This observation led to the question whether mechanisms exist that are able to repress the transcriptional activation of NeuroD1 target genes in the absence of this bHLH protein. Furthermore, it was intriguing to investigate whether NeuroD1 functions by overriding such silencing programs in order to timely activate the expression of neuronal target genes.

3.6.1 Bayesian modeling reveals a distinct chromatin landscape of promoters associated to induced genes based on their additional occupancy with NeuroD1

The chromatin landscape of a given gene and its regulatory regions serves as the template for TFs and other regulatory proteins to act on in order to influence the transcriptional state [75]. As NeuroD1 was able to bind its target regulatory elements in mESCs as well as terminally differentiated fibroblasts and induced the transcription of associated genes, it is likely to interact with the local chromatin environment to function. A first step towards a deeper functional characterisation of its mode of action aimed at a prediction of the chromatin landscape present in wild type mESCs at NeuroD1 target versus non-target regions. Towards this, a number of NGS datasets for murine mESCs (see 8.5) were comprehensively analysed, encompassing ChIP-seq data of established histone modifications (n = 8) and TFs (n = 50) as well as a chromatin accessibility dataset (FAIRE-seq). These datasets were subjected to Bayesian modeling to probe for features that were generally enriched at NeuroD1 target promoters in pluripotent wild type mESCs without the overexpression of NeuroD1. Moreover, an additional aim was to discriminate between the two classes of promoters among URG, either being NeuroD1-bound or non-bound upon ectopic NeuroD1 induction (data not shown). This analysis was able to discriminate the chromatin and transcription factor landscape of promoters belonging to upregulated genes associated with NeuroD1 target or control genes with high accuracy (data not shown). The distinction was based on several features, for example it was found that components of the basal transcriptional machinery (such as Taf3 and RNA Pol II) were less enrichment at NeuroD1 target promoters in comparison to non-target URG promoters. Interestingly further, the transcription factors Tbx3 (T-box transcription factor 3) and Utf1 (Undifferentiated embryonic cell transcription factor 1), the histone modifications H3K27ac and H3K27me3 as well as chromatin accessibility were identified as high-confidence predictors that could significantly distinguish NeuroD1-target from non-target promoters among URG (Figure 45A and B): In mESCs, Tbx3, Utf1, and the Polycomb repressive mark H3K27me3 were enriched at URT whereas these sites lacked the active histone modification H3K27ac and exhibited less accessible chromatin in comparison to non-bound URG promoters. Interestingly, many of the promoter URT were co-occupied by Utf1 and Tbx3 (data not shown). The enrichment of Tbx3 at these repressed neuronal genes was intriguing to find, as a recent study has linked this protein to the specification of mesoendoderm lineage [76]. Overall, these findings suggested that NeuroD1 target promoters are repressed by distinct factors within a heterochromatin state in the absence of NeuroD1 (as exemplified in Figure 45C).

The H3K27ac mark is a prominent predictor for transcriptionally active promoters. To investigate its enrichment dynamics at URT during neuronal development, an *in vitro* differentiation time course of mESCs was carried out and H3K27ac ChIP-seq was performed at prominent stages encompassing mESCs (ES), neuronal progenitors (NP/CA d8), early neurons (TN d1) as well as terminally differentiated neurons (TN d10). A large fraction of promoter URT gained H3K27ac enrichment during the onset of *in vitro* neurogenesis (Figure 45D and E), a stage where *NeuroD1* was most highly expressed (see Figure 35B). The gain of this activating histone modification at promoter URT sites was in line with the overall observed transcriptional induction of URT during the onset of neurogenesis (see Figure 39E). Extending the H3K27ac analyses to

non-neuronal tissues *in vivo* using publicly available ChIP-seq datasets (see 8.5) revealed that promoters of URT lacked the H3K27ac mark in this context (data not shown). This was in agreement with the observations that *NeuroD1* and its targets were transcribed in a neurogenesis related context (see Figure 35B and Figure 39G and H).

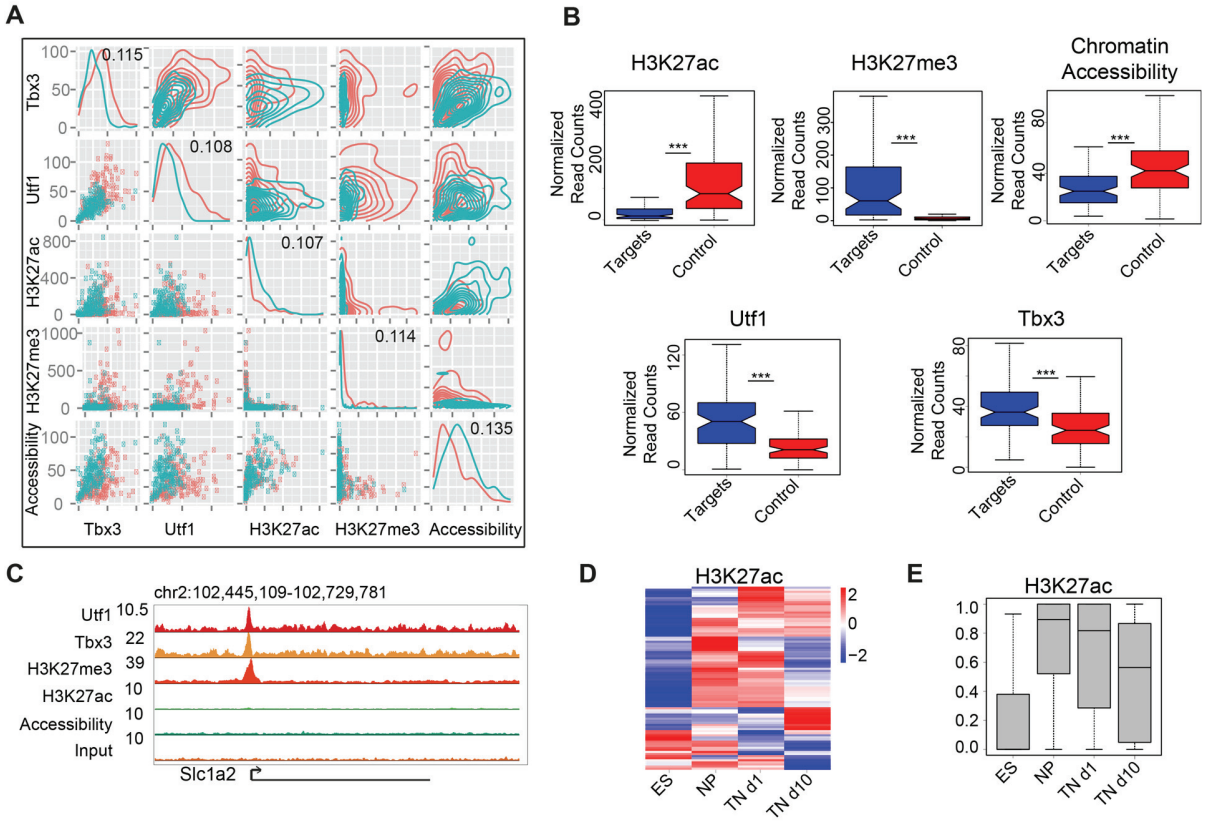


Figure 45: The chromatin and transcription factor landscape can distinguish URT promoter regions from non-target URG promoters.

(A) Distribution plot depicting how ChIP-seq signal strength of the Bayesian model predicted features Tbx3, Utf1, H3K27ac, H3K27me3 as well as chromatin accessibility (FAIRE-seq) correlate at NeuroD1-bound and non-bound promoters of URG in wild type mESCs. NeuroD1 bound promoters of URT are depicted in red, and non-bound promoters of URG are depicted in cyan. The promoters of URT were more enriched for Tbx3, Utf1, and H3K27me3 while lacking H3K27ac and chromatin accessibility in comparison to control promoters of URG. Top-right and diagonal charts group show one and two-dimensional density plots, respectively. The ChIP-seq enrichment is depicted by scatter plots in the bottom left charts. The information gain coefficients for the corresponding features are shown in the diagonal. (B) Boxplots showing ChIP-seq enrichment of H3K27ac, H3K27me3, chromatin accessibility, Utf1, and Tbx3 as normalized read counts at promoters of URT (blue) and non-bound promoters of URG as control (red). Significance was determined by t-test with *** $P < 0.001$. (C) UCSC genome browser screenshot depicting the in (A) identified distinctive chromatin and TF landscape ChIP-/FAIRE-seq features at a representative NeuroD1 target promoter associated to URT. Tbx3, Utf1, and H3K27me3 were co-localized at the promoter whereas H3K27ac was not enriched and chromatin was not accessible. The wiggle files represent on the y-axis a "0" value baseline. The highlighted gene is indicated as a black arrow representing the direction of transcription and therefore the location of the associated promoter to the TSS. (D) Heatmap depicting the dynamics of H3K27ac ChIP-seq enrichment at NeuroD1-bound promoters of URT during an *in vitro* neuronal differentiation time course. These promoters strongly gained H3K27ac enrichment during the onset of neurogenesis (NP/TN d1). Each row in the heatmap represents one promoter where the H3K27ac ChIP-seq enrichment is scaled from blue (low enrichment) to red (high enrichment). The enrichment was calculated from two biological replicates per developmental stage. (E) The same information as in (D) but represented as a boxplot. ChIP-seq enrichment values are scaled between 0 and 1 on the y-axis.

Figure information: ES, embryonic stem cells; NP, neuronal progenitors (CA d8); TN d1, terminally differentiated neurons after one day in culture; and TN d10, terminally differentiated neurons after ten days in culture. Computational analyses performed by AP and PS. All panels are taken from Pataskar & Jung et al., 2016.

3.6.2 NeuroD1 induction leads to remodeling of the heterochromatin landscape at target promoters towards euchromatin

As the Bayesian model predicted several distinctive features for bound and non-bound promoters of upregulated genes (see Figure 45A and B), it raised the question whether these features might be remodeled upon ectopic NeuroD1 expression in mESCs. For this purpose, ectopic NeuroD1 was induced in A2lox.NeuroD1 mESCs for 48 h with doxycycline. Afterwards, ChIP-qPCR assays for Tbx3, H3K27me3, and H3K27ac in addition to FAIRE-qPCR as a measure for chromatin accessibility were carried out in induced (+Dox) as well as in non-induced (-Dox) cells. Strikingly, the tested target promoters gained the active histone mark H3K27ac upon ectopic NeuroD1 induction (Figure 46A). This was accompanied by a loss of the repressive mark H3K27me3 (Figure 46B) and a gain in chromatin accessibility (Figure 46C). This remodeling of the promoter chromatin landscape is in line with the observed increased expression of the associated genes (see Figure 36B). Surprisingly, also the enrichment of Tbx3 decreased at the target promoters upon ectopic NeuroD1 induction (Figure 46C). These analyses showed that heterochromatic gene promoters which become occupied by NeuroD1 (see Figure 43I and K) lost Tbx3 and H3K27me3 occupancy whereas they displayed increased enrichment in the active histone mark H3K27ac and chromatin accessibility after two days. These findings suggested that the potential NeuroD1-mediated antagonism of neuronal lineage repressors such as Tbx3 might provide additional competence for the activation of neuronal genes during neurogenesis.

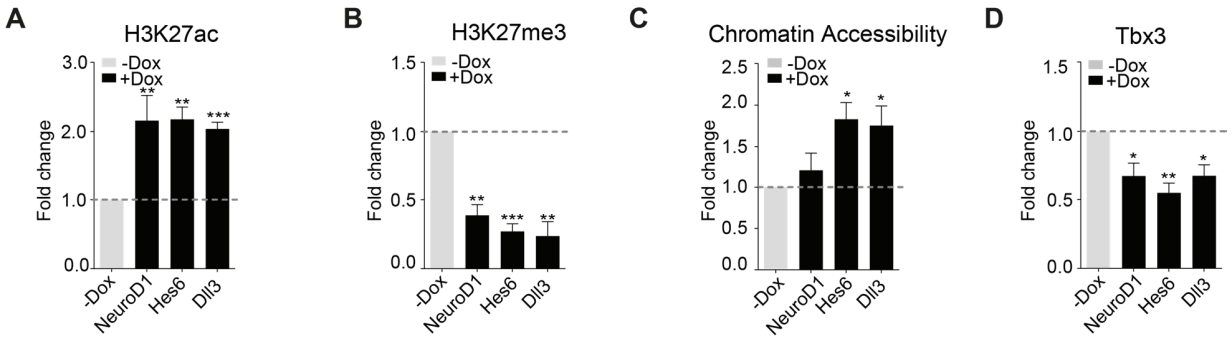


Figure 46: NeuroD1 binding to target promoters remodels the transcription factor landscape and causes a transition from heterochromatin towards euchromatin within 48 h of ectopic NeuroD1 induction.
(A - D) ChIP-qPCR analyses for H3K27ac enrichment (A), H3K27me3 enrichment (B), chromatin accessibility as determined by FAIRE-qPCR (C), and Tbx3 enrichment (D) at representative promoters of URT. A2lox.NeuroD1 mESCs were either induced for 48 h to express ectopic NeuroD1 (+Dox) or used non-induced (-Dox) as a control. The fold change of ChIP-qPCR enrichment (A, B, D) and FAIRE-qPCR signal (C) from +Dox to -Dox of the average enrichment of bound DNA normalized to the respective genomic DNA input is plotted on the y-axis. The tested promoters of URT lost Tbx3 and H3K27me3 enrichment whereas they displayed an increase in the active histone mark H3K27ac and chromatin accessibility after being exposed to ectopic NeuroD1 for two days.
Figure information: Error bars reflect standard error of the mean from four biological replicates. Significance was determined by t-test with * P < 0.05, ** P < 0.01 and *** P < 0.001. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.6.3 Transcriptional induction of NeuroD1 target genes involves consecutive remodeling of the promoter chromatin landscape

The previous results were prompting for a more detailed, time-resolved chromatin analysis during the applied experimental window of two days as the data was derived as an endpoint analysis and did not reveal insights into the timely kinetics underlying epigenetic reprogramming of target promoters. The exact order of events upon NeuroD1 overexpression at target promoters that led to the activation of the associated genes remained elusive. To this end, ChIP-qPCR as well as RT-qPCR time course assays were carried out in A2lox.NeuroD1 cells following the overexpression of ectopic *NeuroD1*. Based on the observed induction kinetics for the ectopic NeuroD1 protein (see Figure 34B), the analyses were carried out 6 h, 12 h, 18 h, 24 h and 48 h after treating A2lox.NeuroD1 mESCs with doxycycline. Untreated cells were taken as control (-Dox). The ChIP assays were performed for ectopic NeuroD1, H3K27me3, H3K27ac, and RNA polymerase II (Pol II) and expression profiling was carried out for the genes associated to selected target promoters. These analyses revealed that both the increase of H3K27ac occupancy as well as the decrease of H3K27me3 enrichment occurred after the binding of ectopic NeuroD1 at these profiled promoters (Figure 47A - C). Interestingly further, this chromatin dynamics was accompanied by a time-dependent recruitment of Pol II, thereby leading to transcriptional induction of the associated genes (Figure 47A - C). Taken together, these findings suggested that ectopic NeuroD1 is able to occupy its target promoter sites despite their heterochromatic state and triggered their remodeling towards euchromatin, which consecutively ushered in gene expression.

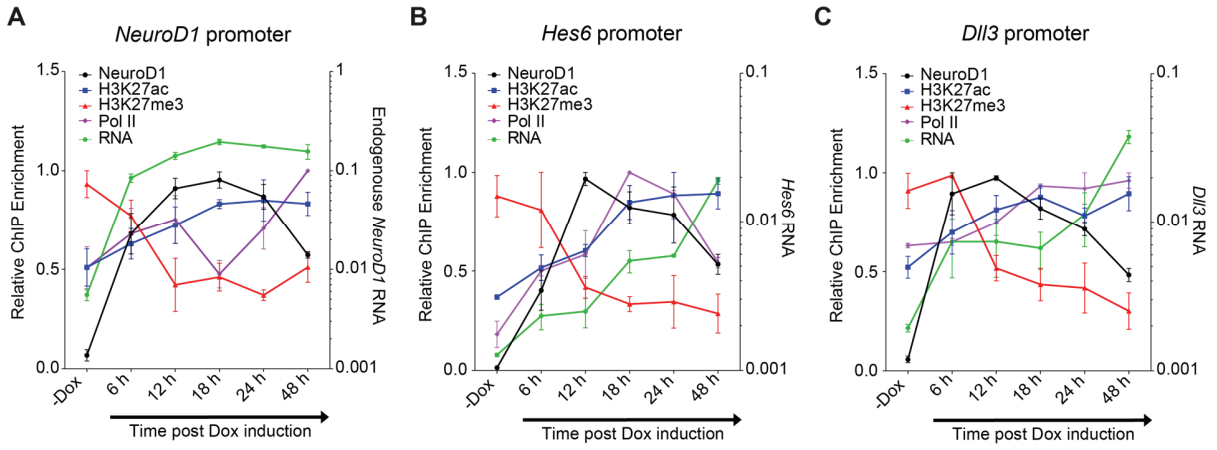


Figure 47: Remodeling of chromatin landscape at NeuroD1 target promoters is triggered by NeuroD1 occupancy and leads to a time-dependent recruitment of RNA-Polymerase II and in transcription of the associated genes.
(A-C) ChIP-qPCR and RNA expression time course analyses at 6 h, 12 h, 18 h, 24 h, and 48 h after ectopic NeuroD1 induction as well as in non-induced A2lox.NeuroD1 cells (-Dox). ChIP-qPCR assays were carried out for enrichment dynamics of ectopic NeuroD1 (black), H3K27ac (blue), H3K27me3 (red), and RNA polymerase II (Pol II, purple) at three representative promoters of URTs [*NeuroD1* in (A), *Hes6* in (B), and *Dll3* in (C)] plotted on the left y-axis. The expression dynamics of associated genes is shown in green plotted on the right y-axis.
Figure information: The average enrichment of bound DNA normalized to the respective genomic DNA input in the ChIP-qPCRs was further normalized to the highest enrichment for each time course replicate and plotted on the left y-axis as relative ChIP enrichment. ChIP-qPCR assays were carried out in four biological replicates. RNA on the logarithmic right y-axis reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*) from three biological replicates. Error bars reflect the standard error of the mean. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.7 NeuroD1 targets to intergenic regions and activates enhancers

The analysed NeuroD1 ChIP-seq data aside from predominant promoter binding also exhibited a large number of intergenic regions which were occupied by NeuroD1 (see Figure 39B - D) and associated to neuronal genes expressed for example during neurogenesis *in vivo* (see Figure 43D - F). The next question was whether relevant target enhancer regions associated with URG (see 3.3.1) exhibited similarly to promoter targets a distinct chromatin and transcription factor landscape in comparison to non-target regions associated with URG. In order to address this question, similar to section 3.6, a Bayesian modeling approach was carried out on the obtained large-scale ChIP-seq data for murine mESCs (see 8.5).

3.7.1 Bayesian modelling predicts a unique chromatin and TF landscape for NeuroD1-bound enhancer elements

The Bayesian model identified three distinct features that enabled a discrimination between enhancers associated with URG, which are either NeuroD1-bound or non-bound upon ectopic NeuroD1 expression (Figure 48A - B): One distinctive feature was the transcription factor Tbx3, which was already identified at promoter targets (see 3.6.1), supporting a general antagonism between Tbx3 and NeuroD1 at target sites. Interestingly, a second differentiator specifically enriched at NeuroD1 target enhancers was the Methyl-CpG-binding domain protein 3 (Mbd3), a subunit of the NuRD (Nucleosome Remodeling Deacetylase) complex, which has been described to facilitate nucleosome remodeling and to deacetylate histones [77-79]. Furthermore, the lysine 4 residue of histone H3 was found to be significantly higher monomethylated (H3K4me1) at these NeuroD1 enhancer targets in comparison to control regions associated with URG. Additionally, a number of enhancer target sites exhibited a co-occupancy of these three discriminative features (exemplified in Figure 48C). These data suggested that NeuroD1 target enhancers might be primed in mESCs [63] for later activation during ontogeny.

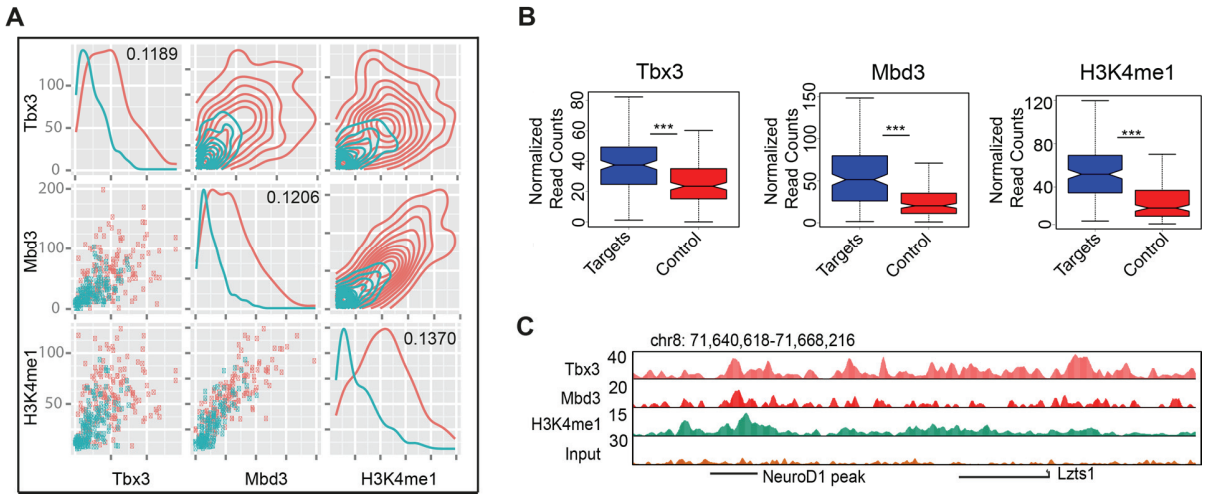


Figure 48: NeuroD1 target enhancers exhibit a different chromatin and transcription factor landscape as non-target intergenic regions associated with URG.
(A) Distribution plots showing how ChIP-seq signal strength for Tbx3, Mbd3 as well as H3K4me1 (Bayesian model predicted features) correlate at NeuroD1-bound enhancers in comparison to non-target enhancers associated to URG. NeuroD1-bound enhancers associated to URG (Enhancer URT) are depicted in red and non-bound enhancers associated with URG are depicted in cyan. Enhancer URT were more enriched for Tbx3, Mbd3 and H3K4me1 in comparison to the control regions. Top-right and diagonal charts group show one and two-dimensional density plots, respectively. The ChIP-seq enrichment is depicted by scatter plots in bottom left charts. The information gain coefficients for the corresponding features are shown in the diagonal. (B) Boxplots depicting ChIP-seq enrichment of Tbx3, Mbd3, and H3K4me1 as normalized read counts at enhancer URT (blue) and non-bound enhancers associated with URG as control (red). Significance was determined by t-test with *** P < 0.001. (C) UCSC genome browser screenshot depicting the in (A) identified distinctive chromatin and TF landscape ChIP-seq features at a representative NeuroD1 target enhancer (NeuroD1 peak) associated to *Lzts1*. Tbx3, Mbd3 and H3K4me1 were co-localized at the enhancer target. The wiggle files represent a "0" value baseline on the y-axis. The highlighted gene is indicated as a black arrow representing the direction of transcription and therefore the location of the associated promoter with the TSS.
Figure information: Computational analyses performed by AP and PS. All panels are taken from Pataskar & Jung et al., 2016.

The histone mark H3K27ac is a prominent predictor of active enhancers [63]. In order to investigate whether the identified enhancer target sites were exhibiting a neurogenesis-related H3K27ac dynamics, the previously generated H3K27ac ChIP-seq time course data of neuronal differentiation *in vitro* (see 3.6.1) was probed for the presence of this histone modification at enhancer URT sites (Figure 49A and B). Indeed, the majority of these distal regulatory elements gained H3K27ac during the onset of neurogenesis. In line with these endogenous chromatin changes at NeuroD1-bound enhancer target sites, also the expression of associated URG increased during neurogenesis both, *in vitro* and *in vivo* (Figure 49C - E). The onset of transcriptional induction was prominently paralleling the dynamics of endogenous *NeuroD1* expression as previously shown (see Figure 35). Furthermore, the analysis of publicly available *in vivo* H3K27ac ChIP-seq data showed that these distal regulatory regions as well as exonic and intronic enhancers associated with URTs were depleted of H3K27ac enrichment in non-neuronal tissues (data not shown). Taken together, these findings suggested that NeuroD1 target enhancers are kept in an inactive state by mechanisms employing Mbd3/NuRD occupancy and a lack of active histone modifications like H3K27ac in non-neuronal developmental stages where NeuroD1 is not present.

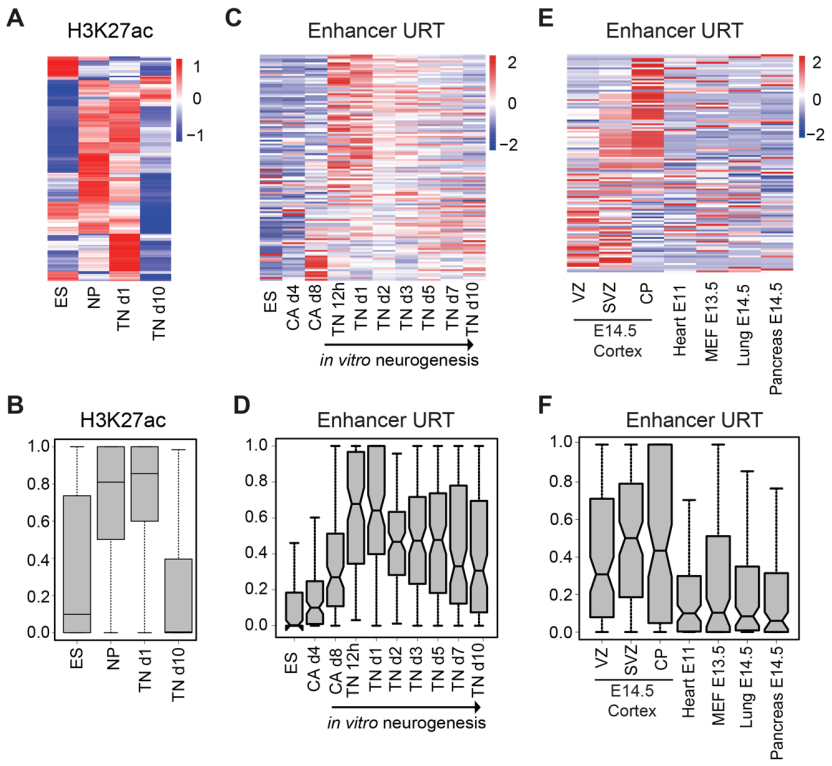


Figure 49: Enhancer URT are activated by a gain of H3K27ac and their associated genes are transcriptionally induced during neuronal differentiation.
(A) Heatmap depicting the dynamics of H3K27ac ChIP-seq enrichment at NeuroD1 bound enhancers associated to URG (enhancer URT sites) during an *in vitro* neuronal differentiation time course of mESCs. These enhancer URT sites gained H3K27ac enrichment during the onset of neurogenesis (NP/TN d1). (B) The same information as in (A) but represented as a boxplot. ChIP-seq enrichment values are scaled between 0 and 1 on the y-axis. (C) Heatmap depicting the expression of enhancer URT in an *in vitro* neuronal RNA-seq time course. Most URT bound by NeuroD1 at their associated enhancer were endogenously upregulated during neurogenesis. (D) Same information as in (C) but represented as a boxplot. (E) Heatmap depicting the expression of enhancer URT in various embryonic tissues. Most URT bound by NeuroD1 at their associated enhancer were endogenously upregulated exclusively during *in vivo* neurogenesis from VZ to SVZ. (F) The same information as in (E) but represented as a boxplot.
Figure information: Each row in the heatmaps represents one enhancer site where the H3K27ac ChIP-seq enrichment (A) or the RNA-seq expression value of the associated gene (C and E) is scaled from blue (low enrichment/expression) to red (high enrichment/expression). The ChIP-seq enrichment was calculated from two biological replicates, the RNA-seq expression values from biological triplicates. Boxplots depicting RNA-seq data (D and F) contain expression values scaled between 0 and 1 on the y-axis. VZ, ventricular zone; SVZ, subventricular zone; CP, cortical plate; ES, embryonic stem cells; NP, neuronal progenitors (CA d8); TN d1, terminally differentiated neurons after one day in culture; TN d10, terminally differentiated neurons after ten days in culture. Computational analyses performed by AP. *In vitro* expression data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.7.2 NeuroD1 induction leads to chromatin remodeling at target enhancer sites

So far, the observations showed that NeuroD1 targeted enhancers were only endogenously active (H3K27ac enriched) at specific stages during neurogenesis (see 3.7.1). Their associated genes were induced upon ectopic NeuroD1 expression in mESCs (see Figure 39D). These findings raised the question whether the Bayesian model predicted repressive features of enhancer URT sites might be altered upon NeuroD1 induction. Furthermore, it remained elusive if these regulatory regions might gain H3K27ac enrichment in parallel to induce the neuronal program, accompanying the remodeling of the promoter

landscape (see 3.6). To this end, ectopic NeuroD1 was induced in A2lox.NeuroD1 mESCs for 48 h with doxycycline. Afterwards, ChIP-qPCR assays were performed for Mbd3, H3K4me1, and H3K27ac in addition to FAIRE-qPCR as a measure for chromatin accessibility in induced (+Dox) as well as in non-induced (-Dox) cells. Strikingly, the tested enhancer sites indeed gained enrichment for H3K27ac, mostly accompanied by H3K4me1 occupancy (Figure 50A and B). Furthermore, the accessibility of chromatin as measured by FAIRE-qPCR increased significantly at these enhancer URT sites (Figure 50C), which was complemented by a reduction of Mbd3 occupancy (Figure 50D). These analyses showed that (primed) distal regulatory regions which become occupied by NeuroD1 (see Figure 43J and K) are remodeled into active enhancers.

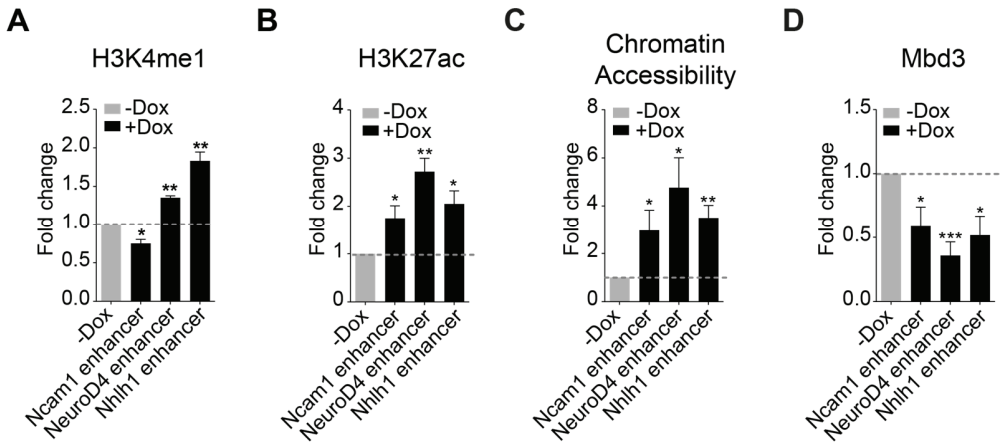


Figure 50: NeuroD1 target distal regulatory elements are remodeled into active enhancers within 48 h of ectopic NeuroD1 induction in mESCs.
(A - D) ChIP-qPCR analyses for H3K4me1 enrichment (A), H3K27ac enrichment (B), chromatin accessibility as determined by FAIRE-qPCR (C), and Mbd3 enrichment (D) at representative enhancer URT sites. A2lox.NeuroD1 cells were either induced for 48 h to express ectopic NeuroD1 (+Dox) or used non-induced (-Dox) as a control. The fold change of ChIP-qPCR enrichment (A, B, D) and FAIRE-qPCR signal (C) from +Dox to -Dox of the average enrichment of bound DNA normalized to the respective genomic DNA input is plotted on the y-axis. The tested enhancer URT sites gained H3K4me1 and H3K27ac enrichment as well as chromatin accessibility whereas they displayed a decrease in Mbd3 occupancy after being exposed to ectopic NeuroD1 for two days.
Figure information: Error bars reflect standard error of the mean from four biological replicates. Significance was determined by t-test with * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.7.3 NeuroD1 binding at distal regulatory regions activates enhancers that induce the expression of associated genes

The Aforementioned results prompted for a more detailed chromatin kinetics analysis during the applied experimental window of two days to determine the dynamics of enhancer activation by ectopic NeuroD1 as already performed for NeuroD1 target promoters (see 3.6.3). For this purpose, ChIP-qPCR as well as RT-qPCR time course assays were carried out in A2lox.NeuroD1 cells following the overexpression of ectopic NeuroD1 at the already established time points. The ChIP assays were performed for ectopic NeuroD1 and H3K27ac, the expression profiling was carried out for genes associated to selected target enhancers (enhancer URT sites) (Figure 51A - C). Strikingly, these analyses revealed that ectopic NeuroD1 occupied its target distal regulatory elements within 6 h after its induction. This was accompanied by a consecutive

increase of H3K27ac enrichment during later time points. These enhancer chromatin dynamics led to the transcriptional activation of associated target genes (Figure 51A - C). However, the transcriptional induction was initiated at later time points in comparison to the induction of promoter URTs (see Figure 47A - C). These observations suggested that NeuroD1 targeting to distal regulatory elements initiates their remodeling towards active enhancers, which consecutively leads to transcriptional induction of associated neuronal genes.

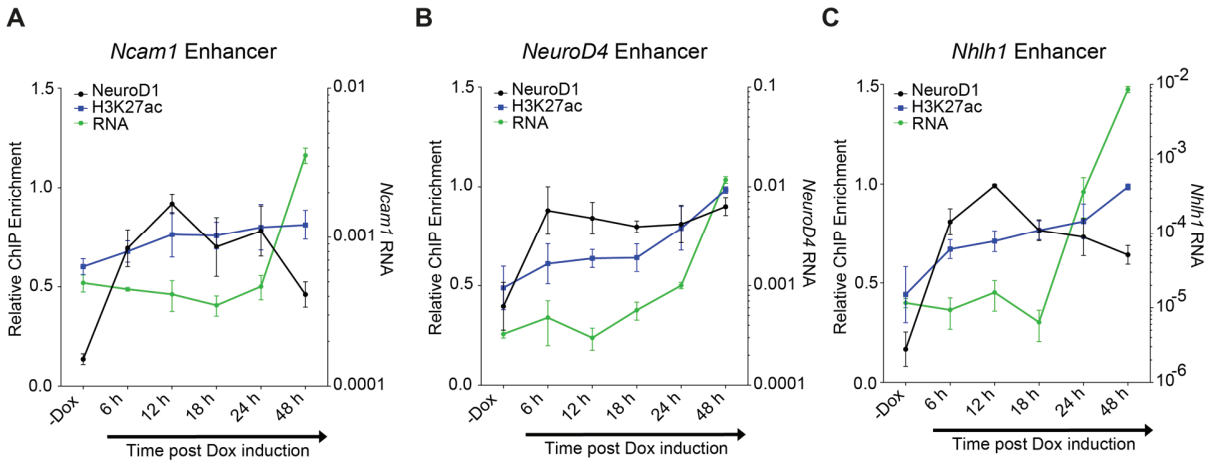


Figure 51: NeuroD1 binding to target enhancers leads to a time dependent increase of H3K27ac and to the induction of associated genes.
(A - C) ChIP-qPCR and RNA expression time course analyses encompassing the time points 6 h, 12 h, 18 h, 24 h, and 48 h after ectopic NeuroD1 induction as well as in non-induced A2lox.NeuroD1 mESCs (-Dox). The assays were carried out for the ChIP-qPCR enrichment dynamics of ectopic NeuroD1 (black) and H3K27ac (blue) at three representative NeuroD1 enhancer sites associated to URG [*Ncam1* in (A), *NeuroD4* in (B), and *Nhlh1* in (C)] plotted to the left y-axis as well as the expression of their associated genes (green) plotted on the right y-axis.
Figure information: The average enrichment of bound DNA normalized to the respective genomic DNA input in the ChIP-qPCRs was further normalized to the highest enrichment for each time course replicate and plotted on the left y-axis as relative ChIP enrichment. The ChIP-qPCR assays were carried out in four biological replicates. RNA on the logarithmic right y-axis reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*) from three biological replicates. Error bars reflect the standard error of the mean. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.8 NeuroD1 binds the identified targets and remodels their chromatin landscape in the embryonic cortex

The previous analyses revealed that ectopic NeuroD1 occupied its target regulatory elements and induced the transcriptional upregulation of associated genes *in vitro* by initiating remodeling events of their chromatin and TF environment. However, it remained elusive whether NeuroD1 is able to occupy the *in vitro* identified target regions also *in vivo*. Furthermore, it was intriguing to investigate if NeuroD1 also functions during cortical development *in vivo* by initiating transcriptional induction of its target genes via modulation of the epigenetic state of regulatory elements. During murine forebrain development, apical progenitors (APs) residing in the VZ express the neural progenitor marker Pax6 and are able to give either rise to neurons as well as neural/neuronal progenitors [43]. This cell population played therefore a crucial role to address the *in vivo* function of NeuroD1.

3.8.1 NeuroD1 binds the identified targets in embryonic cortex and induces a euchromatin state

In order to investigate whether NeuroD1 occupies the identified target elements also *in vivo*, a ChIP-assay was performed with a NeuroD1-specific antibody on cells from freshly dissected murine embryonic cortices at mid-neurogenesis (E14.5). The ChIP material was consecutively probed by qPCR for NeuroD1 enrichment at the previously selected target regions (both promoter and enhancer sites, see 3.5) as well as additional potential *in vivo* targets (see e.g. Figure 42C - F). The analysis revealed that these *in vitro* identified targets of ectopic NeuroD1 were indeed occupied by the endogenous NeuroD1 protein in the embryonic cortex whereas control regions did not exhibit enrichment (Figure 52A). These results suggested that the chosen approach reliably identified authentic NeuroD1 target regions with *in vivo* relevance.

To examine the impact of NeuroD1-binding on the epigenetic landscape and the transcriptional state of its targets *in vivo*, NeuroD1 was ectopically induced directly in the apical progenitor population of the developing embryonic cortex. To this end, the *in utero* electroporation (IUE) methodology was employed as described previously [80]. This assay enabled to specifically manipulate the radial glial cell population in the VZ at mid-neurogenesis by directly electroporating a vector into the cellular population attached to the ventricular surface [81]. Either the pCIDRE-NeuroD1 (NeuroD1-RFP) or the pCIDRE vector without the additional CDS for NeuroD1 (Control-RFP) were electroporated at E13.5 into the apical progenitor population of the developing embryonic cortex (Figure 52B). Two days later, at E15.5, the embryos were sacrificed and the manipulated cell populations were purified directly from the embryonic brain. For this purpose, FACS methodology was employed based on the fluorescence intensity of the RFP marker encoded by the vector backbone (Figure 52B). The isolated populations were subjected to an analysis of chromatin changes by profiling the enrichment of H3K27ac employing ChIP-qPCR assay. Strikingly, the overexpression of NeuroD1 during 48 h of *in vivo* cortical development led to an increase of H3K27ac enrichment at both classes of previously studied NeuroD1 regulatory elements, the promoter as well as enhancer targets, but not at non-target control regions (Figure 52C). This findings suggested that the associated genes might concomitantly be transcriptionally induced in the population exposed to ectopic NeuroD1 *in vivo*. The gene expression profiling of transcriptional changes by RT-qPCR revealed that the selected target genes were indeed more expressed in the cell population exposed to ectopic NeuroD1 for 48 h *in vivo* in comparison to the control population (Figure 52D). These observations strongly supported the *in vivo* relevance of the previously *in vitro* identified NeuroD1 targets that are likely essential to drive NeuroD1-dependent cell fate changes towards neuronal state.

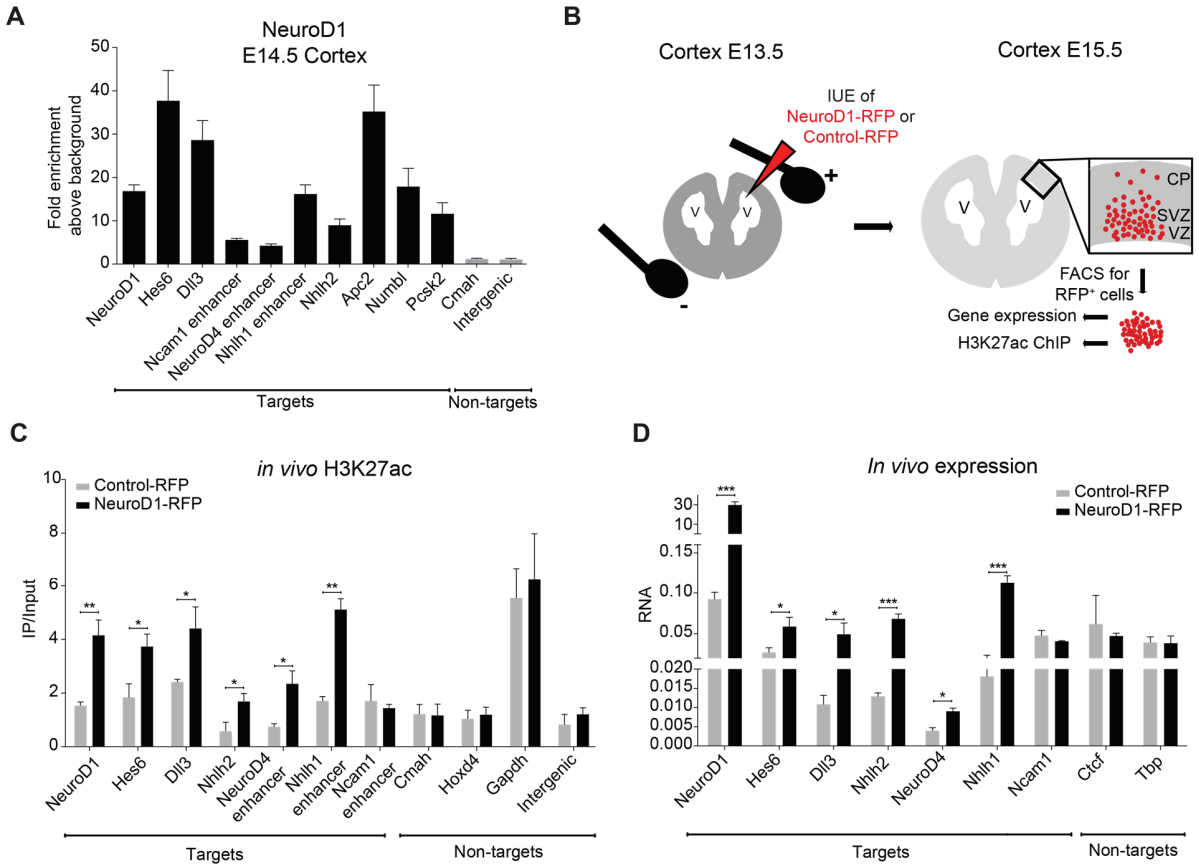


Figure 52: NeuroD1 occupies identified targets in the murine embryonic cortex and induces their euchromatin state with expression of associated genes.

(A) ChIP-qPCR analysis showing enrichment of endogenous NeuroD1 at target (black) and control sites (grey) in the embryonic cortex at E14.5. NeuroD1 specifically occupied the previously *in vitro* identified target regulatory elements also *in vivo*. The y-axis depicts the relative ChIP enrichment plotted as the ratio of precipitated DNA (bound) to input DNA and further normalized to an intergenic control region (fold enrichment above background). (B) Scheme showing the *in utero* electroporation (IUE) approach for NeuroD1-RFP (pCIDRE-NeuroD1) or Control-RFP (pCIDRE) plasmids at E13.5 (left panel). The vector was injected into the ventricle of the embryonic cortex. The negatively charged plasmid DNA was electroporated into apical progenitors in the dorsal ventricular zone due to directing the applied electric current to the region of interest by positioning the cathode (+) towards the pial surface. Scheme depicting the isolation procedure of RFP-positive cortical cells by FACS at E15.5 and indicated downstream applications (right panel). (C) ChIP-qPCR analysis for H3K27ac enrichment at NeuroD1 target *cis*-regulatory elements as well as control regions in E15.5 FAC-sorted cortical cells 48 h post IUE. NeuroD1 target regions specifically gained H3K27ac enrichment in the population exposed to ectopic NeuroD1 (NeuroD1-RFP, black) in comparison to the control population (Control-RFP, grey). The y-axis depicts the relative H3K27ac enrichment plotted as the ratio of precipitated DNA (bound) to input DNA (IP/Input). (D) RT-qPCR analysis for the expression of NeuroD1 target and non-target genes in E15.5 FAC-sorted cortical cells 48 h post IUE. Genes associated to NeuroD1 bound regulatory regions (targets) specifically exhibited an increased expression in the cell population exposed to ectopic NeuroD1 (NeuroD1-RFP, black) in comparison to the control population (Control-RFP, grey). RNA is plotted as the relative gene expression normalized to a housekeeping gene (*Rpl19*) on the y-axis.

Figure information: Error bars reflect standard error of the mean from three (A, D) or four (C) biological replicates. Significance was determined by t-test with * P < 0.05, ** P < 0.01 and *** P < 0.001. CP, cortical plate; SVZ, subventricular zone; V, ventricle; VZ, ventricular zone; IUE, *in utero* electroporation. IUE and FAC-sorting performed by FN, data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.8.2 NeuroD1 overexpression promotes the neurogenic differentiation of neural progenitors *in vivo*

The chromatin and expression analyses *in vivo* suggested that NeuroD1 is able to induce its neuronal targets upon overexpression in neural progenitors of the VZ. However, it remained elusive whether these manipulated progenitor cells consecutively differentiated towards neurons. To address this question, an immunohistochemical analysis was perfomred on coronal cryosections of E15.5 embryonic brains which had been electroporated at E13.5 with the pCIDRE-NeuroD1 or pCIDRE vector as previously described (see Figure 52B). An immunofluorescence staining was carried out for the neural progenitor marker Pax6 and the spatial distribution of the progeny of electroporated cells along the cortical layers was detected by the RFP fluorescence signal encoded by the vector backbone. Interestingly, upon overexpression of NeuroD1 for 48 h *in vivo*, the VZ (accommodating proliferating progenitors) exhibited a decrease in the proportion of RFP-expressing cells in comparison to electroporated control cortices (Figure 53A - C). In parallel, the proportion of RFP-positive cells residing in the intermediate zone (IZ, a layer which consists mainly of newborn migrating neurons) increased upon overexpression of NeuroD1 in comparison to control brains (Figure 53A - C). This observation was supported by the almost complete loss of NeuroD1-overexpressing cells which exhibited a co-staining for Pax6 (Figure 53D). Taken together, this finding suggested that the pool of apical progenitors overexpressing NeuroD1, which is normally not highly expressed in these cells, most likely prematurely differentiated into neurons. These results supported the previous *in vitro* findings and established that NeuroD1 drives neuronal cell fate acquisition *in vivo*.

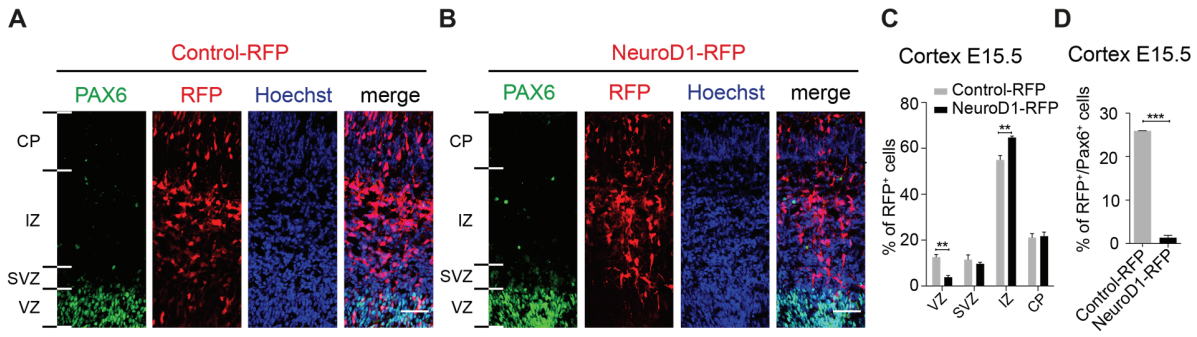


Figure 53: NeuroD1 overexpression promotes the neurogenic differentiation of Pax6-positive progenitors *in vivo*.
(A - B) Immunohistochemical analysis for Pax6 (neural progenitor marker), RFP (electroporated cells), and Hoechst (nucleus) in a representative cortical cryosection of a E15.5 mouse brain 48 h post IUE with pCIDRE (Control-RFP) vector (A) or with pCIDRE-NeuroD1 (NeuroD1-RFP) vector (B). Almost no RFP-positive cells were found in the VZ upon overexpression of NeuroD1. Shown is a representative example out of two biological replicates. Scale bar, 50 µm. (C) Quantifications of the spatial distribution electroporated cells at E15.5 among the four cortical layers 48 h after IUE of pCIDRE or pCIDRE-NeuroD1. NeuroD1 overexpression led to a decrease of RFP-positive cells residing in the VZ and to an increase in the population of RFP-positive cells in the IZ. (D) Quantification of RFP and Pax6 double-positive cells at E15.5 48 h after IUE of pCIDRE or pCIDRE-NeuroD1 vectors showed a loss of double-positive cells upon overexpression of NeuroD1.
Figure information: Error bars reflect standard error of the mean from two biological replicates. Significance was determined by t-test with ** P < 0.01 and *** P < 0.001. CP, cortical plate; IZ, intermediate zone; SVZ, subventricular zone; VZ, ventricular zone. IUE performed by FN, data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.9 Transient action of NeuroD1 is sufficient to confer stable epigenetic changes at its target regulatory elements

The *in vitro* experiments with A2lox.NeuroD1 mESCs had been carried out with cell culture media containing LIF (leukemia inhibitory factor), a cytokine which supports the pluripotent nature of embryonic stem cells [82-86]. The next aim was to review the previous *in vitro* findings in a cellular environment *per se* promoting neuronal differentiation.

3.9.1 Principle of the induced terminal neuron differentiation protocol

In order to remove pluripotency signals and culture A2lox.NeuroD1 cells in a neurogenic environment, the iTN (induced terminal differentiated neurons) differentiation protocol was developed (Figure 54). In contrast to the established neuronal differentiation protocol by Bibel and colleagues [16, 17], the iTN system omitted the stage of cellular aggregate formation. In contrast, A2lox.NeuroD1 cells were directly plated on culture dishes coated with polyornithine and laminin to provide an enhanced matrix for neuronal differentiation and growth of neuronal processes. To facilitate neuronal differentiation, ectopic NeuroD1 was induced at the time of plating by the addition of doxycycline to the N2 medium, which did not contain LIF and serum. After two days, the medium was changed to complete medium which enhances neuronal differentiation. The doxycycline treatment was continued up to seven days and allowed the culture of A2lox.NeuroD1 cells for a longer time period than before. As a control condition, A2lox.NeuroD1 mESCs without doxycycline treatment were taken along. The color-coded arrows in Figure 54 indicate the performed assays at certain developmental stages.

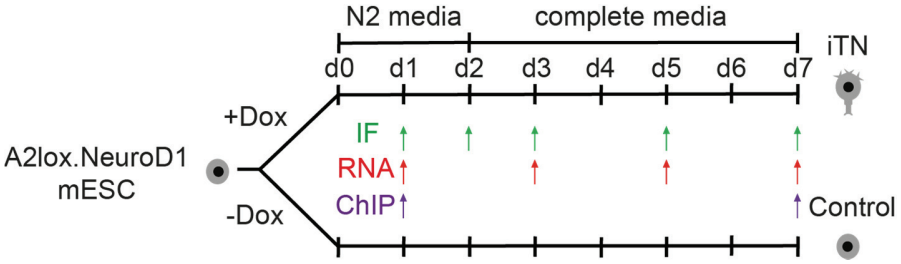


Figure 54: Principle of induced terminal neuron formation.
Scheme depicting the protocol for differentiation of A2lox.NeuroD1 mESCs into induced terminal neurons (iTN). The iTNs were induced by expression of ectopic NeuroD1 (+Dox) whereas the non-induced cells (-Dox) served as control. The arrows indicate time points at which experiments were performed [Immunocytochemistry (IF), RNA expression analysis by RT-qPCR (RNA), and chromatin immunoprecipitation (ChIP)]. ChIP-qPCR for ectopic NeuroD1 was performed at day one (d1), and all ChIP-qPCRs for changes in transcription factor and chromatin landscapes as well as FAIRE-qPCRs for chromatin accessibility were performed at day seven. System and model generated by JJ. Figure taken from Pataskar & Jung et al., 2016.

3.9.2 Validation of neuron formation in the iTN system

As a first measure, the expression level of *NeuroD1* was probed during iTN differentiation of A2lox.NeuroD1 mESCs (Figure 55). The expression of ectopic *NeuroD1* was only detectable in the induced condition

(+Dox), in which expression levels peaked at day three to five of differentiation. The RNA levels of endogenous *NeuroD1* mimicked this pattern upon induction of ectopic *NeuroD1* until day five, whereas the minimal induction in the control (-Dox) condition decreased over time. Overall, the expression pattern of *NeuroD1* during iTN differentiation closely resembled the previous observations in mESCs (see Figure 34A).

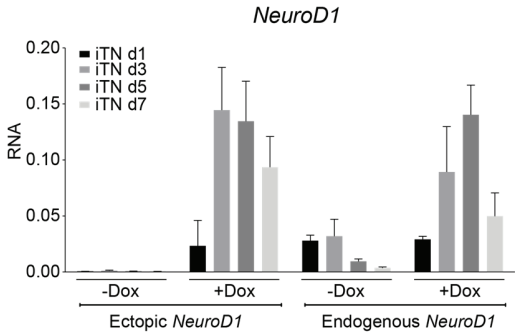


Figure 55: Expression dynamics of *NeuroD1* during iTN formation.
RT-qPCR analysis for the expression dynamics of ectopic as well as endogenous *NeuroD1* during iTN differentiation of A2lox.NeuroD1 mESCs for seven days. Cells were either continuously induced with doxycycline (+Dox) to express ectopic *NeuroD1* or left untreated (-Dox) as control. Both, ectopic as well as endogenous *NeuroD1* expression levels increased upon differentiation of A2lox.NeuroD1 mESCs into induced terminal neurons. Ectopic *NeuroD1* was not detectable in the control (-Dox) condition, whereas the expression of endogenous *NeuroD1* was comparable at day one in both conditions, but dropped afterwards in control cells.
Figure information: RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*). Error bars indicate standard error of the mean from two biological replicates. Data derived by JJ. Figure taken from Pataskar & Jung et al., 2016.

After the confirmation of *NeuroD1* induction during iTN formation, the next question addressed was whether the increased levels of this bHLH factor also lead to a neuronal like morphology and expression of neuronal markers such as a TUJ1 in iTN cells as observed for the induction of NeuroD1 under culture conditions for embryonic stem cells (see Figure 36A). Therefore, an immunocytochemistry time course analysis was carried out during seven days of iTN formation in both, the induced (+Dox) and control (-Dox) condition (Figure 56A - B). The A2lox.NeuroD1 cells expressing ectopic *NeuroD1* (+Dox) exhibited a gradual increase in the number of TUJ1-positive cells during seven days of iTN culture in a neurogenic environment (Figure 56B). However, although a small number of TUJ1-positive cells was detected in non-induced cells (-Dox), their appearance was limited to the last investigated time point (Figure 56A). The occurrence of a neuron-like morphology was in line with the detection of the neuronal marker TUJ1 in all iTN cells and suggested the manifestation of a neuronal developmental program upon NeuroD1 induction also in these culture conditions.

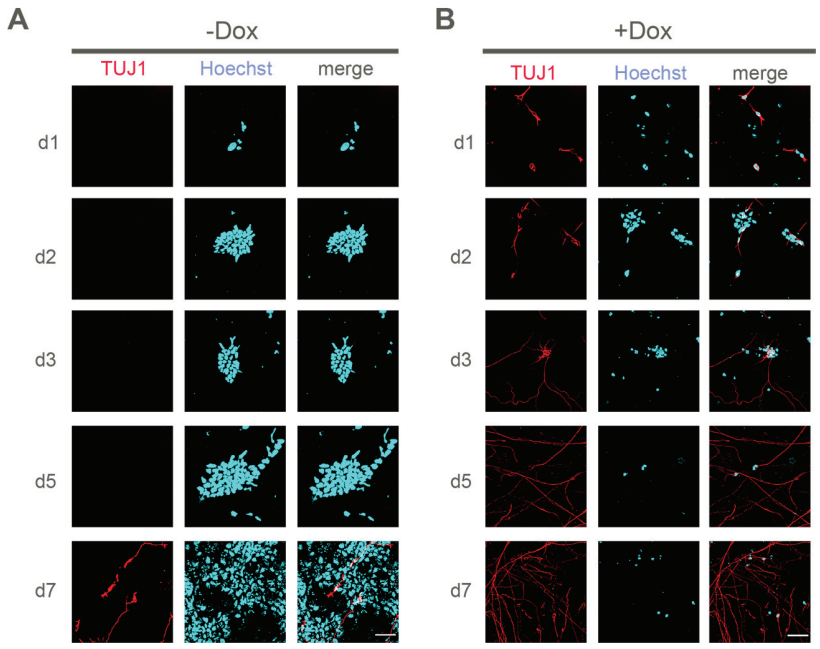


Figure 56: iTNs exhibit a neuronal morphology and express TUJ1.
(A , B) Immunocytochemistry time course analysis for TUJ1 (neuronal marker) and Hoechst (nucleus) during iTN differentiation of A2lox.NeuroD1 cells for seven days. Cells where either continuously induced with doxycycline (+Dox, B) to express ectopic *NeuroD1* or left untreated (-Dox, A) as control. The formation of TUJ1-positive cells was visible after 24 h of doxycycline induction and characterised by a gain of increasingly complex TUJ1-positive processes within seven days. The leower number of cells during iTN formation (+Dox) was most likely a result of cells acquiring a post-mitotic state as part of neurogenesis whereas control cells (-Dox, A) that lacked ectopic *NeuroD1* and therefore an induction of neurogenesis continued to proliferate. However, virtually all cells were TUJ1-positive after five days of iTN differentiation.
Figure information: Scale bar, 50 µm. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

In order to investigate the expression of the previously measured hallmark pluripotency and neuronal marker genes (see Figure 36B - C), a RT-qPCR time course analysis was performed during iTN formation along with the control condition (Figure 57). Interestingly, all tested hallmark pluripotency genes, except *Sox2*, were downregulated following ectopic *NeuroD1* induction as well as in the control condition within seven days (Figure 57A). However, the reduction of expression levels was achieved earlier in the doxycycline treated cells in comparison to the non-induced condition, especially prominent in the RNA profile of *Nanog* at day one. Interestingly further, *Sox2* expression became similarly induced in both conditions, which was not observed for the previous differentiation of mESCs in LIF containing medium. Importantly, this effect was accompanied by a strong and constant increase of the RNA levels measured for the neuronal markers *MapT*, *Syp* and *vGlut2* only in induced cells (Figure 57B). Interestingly, although the expression of *Tubb3* complemented this expression profile upon ectopic *NeuroD1* induction, its RNA levels were briefly upregulated during the first three days in non-induced cells and thereafter decreases continuously.
Taken together, these results suggested that a rapid and efficient neuron formation is accomplished during iTN differentiation. Although the culture medium with reduced pluripotency stimuli in combination with a neurogenic growth environment led to a minimal induction of neurogenesis, only a few cells exhibited an iTN characteristic morphology after seven days. These data showed that NeuroD1 can generate neurons

under these applied culture conditions and therefore these cells are referred to as “induced terminally differentiated neurons” (iTNs) from here onwards.

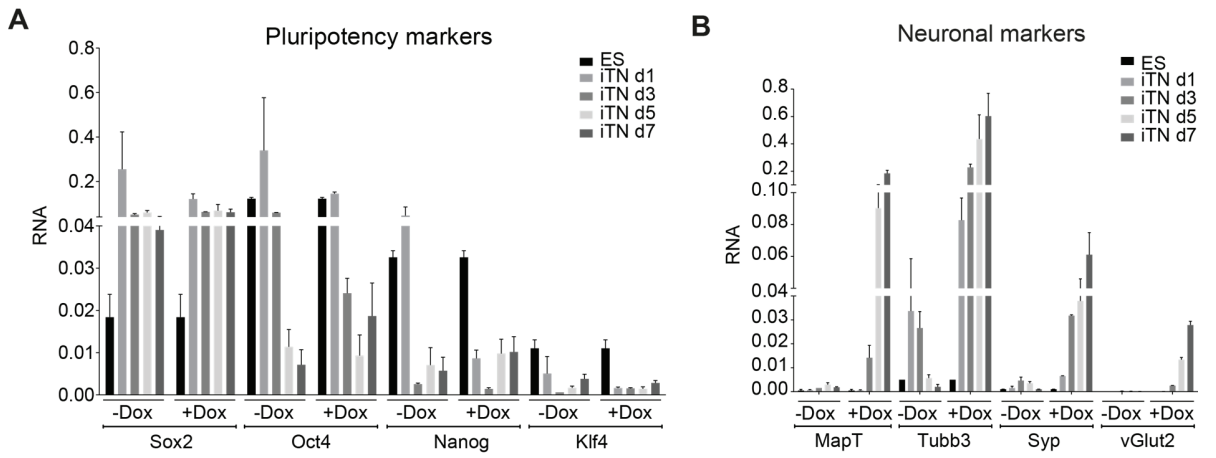


Figure 57: The formation of iTN cells accompanies an induction of neuronal and repression of pluripotency genes.
(A, B) RT-qPCR analysis for the expression of hallmark genes linked to pluripotency (A) in addition to neuronal marker genes (B) during iTN differentiation of A2lox.NeuroD1 mESCs for seven days. Cells were either continuously induced with doxycycline (+Dox) to express ectopic *NeuroD1* or left untreated (-Dox) as control. Additionally, 159.2 mESCs (ES) were probed as control. The pluripotency genes, except *Sox2*, displayed decreased expression whereas neuronal markers were increased in expression within seven days.
Figure information: Error bars indicate standard error of the mean from two biological replicates. RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*). Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.9.3 NeuroD1 remodels the chromatin landscape at its target sites during iTN formation to induce the expression of associated neuronal genes

The formation of iTNs suggested that NeuroD1 might function similarly as observed in mESCs. The first question addressed was whether ectopic NeuroD1 targets the same previously identified regulatory elements also during iTN differentiation. To this end, a ChIP assay for ectopic NeuroD1 was carried out during iTN formation and the six selected target regions (see 3.5) were probed for NeuroD1 occupancy by qPCR analysis. Strikingly, all tested promoter and enhancer sites were enriched for ectopic NeuroD1 within 24 h of iTN formation (Figure 58A). To test if this NeuroD1 occupancy was accompanied by the acquisition of an active chromatin signature at these sites, ChIP-qPCR assays for H3K27ac and H3K27me3 were carried out after seven days of iTN differentiation. Indeed, the tested promoter targets lost the repressive H3K27me3 chromatin modification and likewise gained in parallel to the enhancer target regions the active H3K27ac mark (Figure 58B - C). The activation of these regulatory elements was further supported by an increase in their chromatin accessibility levels at day seven as measured by FAIRE-qPCR (Figure 58D). To further investigate if such epigenetic remodeling of NeuroD1 target regulatory elements results in transcriptional activation of their associated genes, a RT-qPCR assay for these genes was carried out after seven days of iTN formation. Indeed, all the tested URTs exhibited a significant transcriptional activation in iTN cells (Figure 58E). Taken together, these findings further supported the conclusion that NeuroD1 can

bind its genomic target regions independent of chromatin context and mediate their epigenetic remodeling to confer transcriptional competence to associated neuronal target genes.

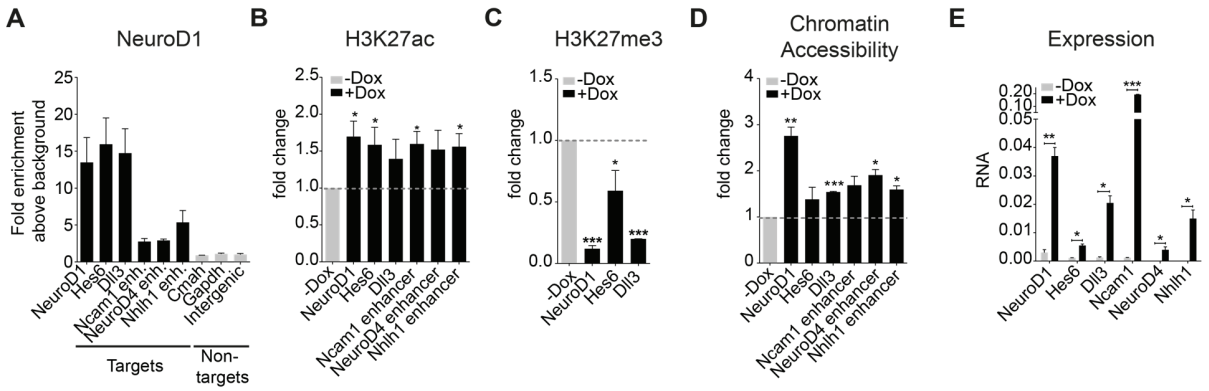


Figure 58: NeuroD1 binding during iTN formation leads to euchromatin formation at target elements promoting expression of associated neuronal genes.
(A) ChIP-qPCR analysis for ectopic NeuroD1 enrichment at identified target promoters and enhancers 24 h after induction of iTN differentiation. Ectopic NeuroD1 occupied its target *cis*-regulatory elements also during iTN formation. (B) ChIP-qPCR analysis for the fold change of H3K27ac enrichment at NeuroD1 target regulatory elements after seven days of iTN formation between induced (+Dox) and non-induced control cells (-Dox). All target promoter and enhancer elements gained enrichment of the active H3K27ac mark upon iTN differentiation. (C) ChIP-qPCR analysis for the fold change of H3K27me3 enrichment at NeuroD1 target regulatory elements after seven days of iTN formation between induced (+Dox) and non-induced control cells (-Dox). All target promoter elements lost enrichment of the repressive H3K27me3 mark upon iTN differentiation. (D) FAIRE-qPCR analysis for the fold change of chromatin accessibility at NeuroD1 target regulatory elements after seven days of iTN formation between induced (+Dox) and non-induced control cells (-Dox). NeuroD1 target regulatory regions displayed an increase in accessible chromatin upon iTN differentiation. (E) RT–qPCR analysis for the expression of genes associated with the tested target regulatory elements after seven days of iTN formation. All representative URTs were transcriptionally induced during iTN formation.
Figure information: A2lox.NeuroD1 mESCs were either continuously induced with doxycycline (+Dox) to express ectopic *NeuroD1* or left untreated (-Dox) as control. Error bars reflect standard error of the mean from three (A - D) or two (E) biological replicates. The y-axis in (A) shows the relative ChIP enrichment plotted as the ratio of precipitated DNA (bound) to input DNA and further normalized to an intergenic control region (background). The fold change of ChIP-qPCR enrichment (B, C) and FAIRE-qPCR signal (D) between tested conditions is plotted as the fold change of the average enrichment of bound DNA normalized to the respective genomic DNA input on the y-axis of remaining ChIP/FAIRE-qPCR plots. RNA in (E) reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*). Significance was determined by t-test with * P < 0.05, ** P < 0.01, *** P < 0.001. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.9.4 The expression level of neuronal marker genes is triggered by the induction levels of ectopic *NeuroD1*

Previous studies showed, that the expression of several bHLH factors (e.g. *Hes1* or *Ascl1*) is oscillating in multipotent neural progenitors and that upon cell fate determination only the expression of the fate determining bHLH factor(s) is accumulating [87-89]. These findings raised the question whether NeuroD1-induced neuronal gene induction was dependent on the expression levels of this bHLH factor itself. To this end, induction kinetics of ectopic *NeuroD1* were performed using three different concentrations of doxycycline (50 ng/ml, 500 ng/ml and 1,000 ng/ml) during an iTN differentiation time course of A2lox.NeuroD1 mESCs. The consecutive RNA expression analysis employing RT-qPCR revealed, that the lowest concentration of doxycycline led to a very minute expression of ectopic *NeuroD1* whereas the higher doses similarly induced its RNA levels comparably to the previously observed levels (Figure 59A).

Interestingly, the tested neuronal marker genes (*Tubb3*, *MapT*, *vGlut2* and *Syp*) were also only minimally induced with 50 ng/ml doxycycline (Figure 59B - E). Both treatments with 500 ng/ml and 1,000 ng/ml doxycycline resulted in a transcriptional increase of tested neuronal marker genes, suggesting a successful iTN formation in line with the previous results. Taken together, these observations suggested that the degree of transcriptional induction of neuronal marker genes correlates with the expression levels of ectopic *NeuroD1*.

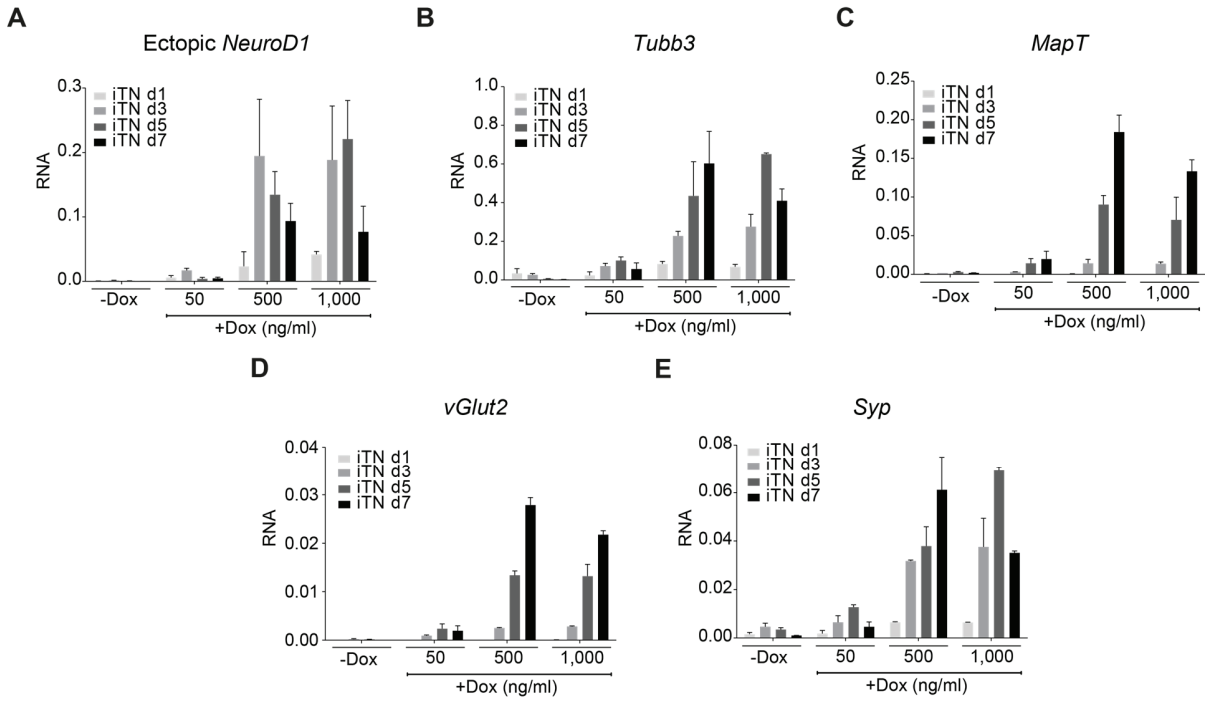


Figure 59: The induction of neuronal marker genes during iTN formation depends on the doxycycline concentration and hence ectopic *NeuroD1* expression.
(A - E) RT-qPCR analyses for the dose-dependent induction of ectopic *NeuroD1* (A) and neuronal marker genes [*Tubb3* (B), *MapT* (C), *vGlut2* (D), and *Syp* (E)] during iTN differentiation of A2lox.NeuroD1 mESCs for seven days. Cells were either continuously induced with three different concentrations of doxycycline (+Dox: 50, 500 and 1,000 ng/ml) to express ectopic *NeuroD1* or left untreated (-Dox) as control. Over time, the lowest concentration of 50 ng/ml doxycycline did not advance into a strong increase of *ectopic* NeuroD1 and neuronal marker genes in comparison to control. Both, the 500 ng/ml and 1,000 ng/ml doxycycline concentration efficiently induced ectopic *NeuroD1* in comparison to the control condition and led to increased expression of neuronal marker genes over time.
Figure information: Error bars indicate the standard error of the mean from two biological replicates. RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*). Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.9.5 Transient induction of ectopic *NeuroD1* induces the complete neuronal fate transition program during iTN formation

The finding that the target gene activation seemed to depend on the expression level of *NeuroD1* was furthermore intriguing in the light of the previous observations: (1) *NeuroD1* was only transiently induced during the onset of neurogenesis *in vitro* (see Figure 35A) as well as *in vivo* (see Figure 35B). (2) Many NeuroD1 target neuronal genes maintained an active transcriptional state at later developmental stages, again both during *in vitro* (see Figure 39E and F) and *in vivo* neurogenesis (see Figure 39G and H).

Therefore, the question arose whether transient availability of ectopic NeuroD1 would still lead to the induction of neuronal program *in vitro*. To address this question, the transient availability of NeuroD1 was mimicked during iTN differentiation by providing doxycycline only for two days instead of a continuous supply in the culture medium. The iTN system served here as a promising model as it had already demonstrated to display a high sensitivity to doxycycline as the transcriptional responses of target genes differed depending on the applied concentration. The extended experimental time window up to seven days allowed to probe for potential long term effects of transient NeuroD1 expression. The expression level investigation of ectopic *NeuroD1* during the transient induction condition (+/-Dox) revealed that indeed ectopic *NeuroD1* was induced during the first two days of iTN formation but was not detected anymore upon the removal of doxycycline after these initial 48 h (Figure 60).

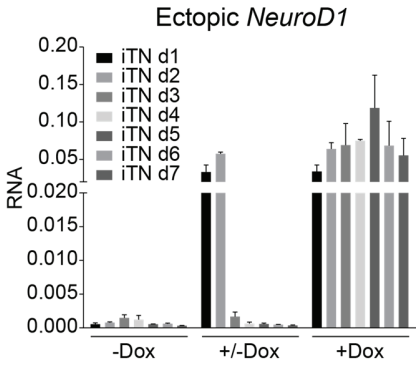


Figure 60: Ectopic *NeuroD1* expression is not detectable anymore after transient induction for 48 h during iTN formation.
RT-qPCR analysis time course for the expression dynamics of ectopic *NeuroD1* during iTN differentiation of A2lox.NeuroD1 cells for seven days. Cells were either continuously induced with doxycycline (+Dox) to express ectopic *NeuroD1*, left untreated (-Dox) as control or doxycycline was removed from the culture medium after 48 h (+/- Dox). RNA levels of ectopic *NeuroD1* were absent within 24 h after the removal of doxycycline.
Figure information: RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*). Error bars indicate the standard error of the mean of two biological replicates. Data derived by JJ. Figure taken from Pataskar & Jung et al., 2016.

An immunocytochemistry analysis for TUJ1 revealed that although ectopic *NeuroD1* was only transiently present during the onset of iTN differentiation as shown above (+/-Dox), TUJ1-positive iTN cells still emerged within seven days of iTN formation (Figure 61A). This observation suggested that the transient expression of NeuroD1 is sufficient to induced neurogenesis *in vitro*.

It was intriguing to investigate whether the previously observed NeuroD1-induced changes in the epigenetic landscape of regulatory elements (see Figure 58B - D) could be maintained after removal of *NeuroD1*. Therefore, ChIP-qPCR assays for H3K27ac and H3K27me3 were carried out in addition to FAIRE-qPCR measuring chromatin accessibility after differentiating A2lox.NeuroD1 mESCs while ectopic *NeuroD1* was only induced during the first two days of iTN formation (+/-Dox). Interestingly, the previously observed NeuroD1-induced increase of H3K27ac levels at target regulatory regions was only maintained at few target sites (Figure 61B). However, the loss of the repressive H3K27me3 histone modification was efficiently preserved at all target sites despite the only transient availability of ectopic *NeuroD1* (Figure 61C). In

addition, the NeuroD1-induced alteration of the chromatin landscape towards euchromatin was supported by the observed tendency of increased chromatin accessibility at target elements, although only few were significantly changing (Figure 61D).

The next step was to investigate whether the transcription factor landscape during iTN differentiation was similarly reprogrammed as observed for continuous ectopic *NeuroD1* expression in mESCs. Therefore, ChIP assays with antibodies directed against Mbd3 and Tbx3 were carried out after seven days of iTN differentiation induced by transient (+/-Dox) as well as constitutive (+Dox) ectopic NeuroD1 expression. Interestingly, the ChIP-qPCR analyses of NeuroD1 target regulatory elements showed that the occupancy of Tbx3 (Figure 61E) and Mbd3 (Figure 61F) decreased during iTN formation (+Dox) in comparison to control (-Dox). Strikingly, even the transient expression of ectopic *NeuroD1* during the onset of iTN differentiation (+/-Dox) led to a loss of both TFs from the probed target sites (Figure 61E - F). In line with the observed maintenance of NeuroD1-induced changes in the chromatin and transcription factor landscape at its target regulatory elements, the associated URTs maintained the previously observed transcriptional induction despite the removal of ectopic *NeuroD1* after the initial two days of iTN differentiation (Figure 61G).

Taken together, the findings suggested that the transient function of NeuroD1 during the onset of neurogenesis might lead to a persistent change of the chromatin and transcription factor landscape which sustains concomitantly the transcriptional induction of associated neuronal genes.

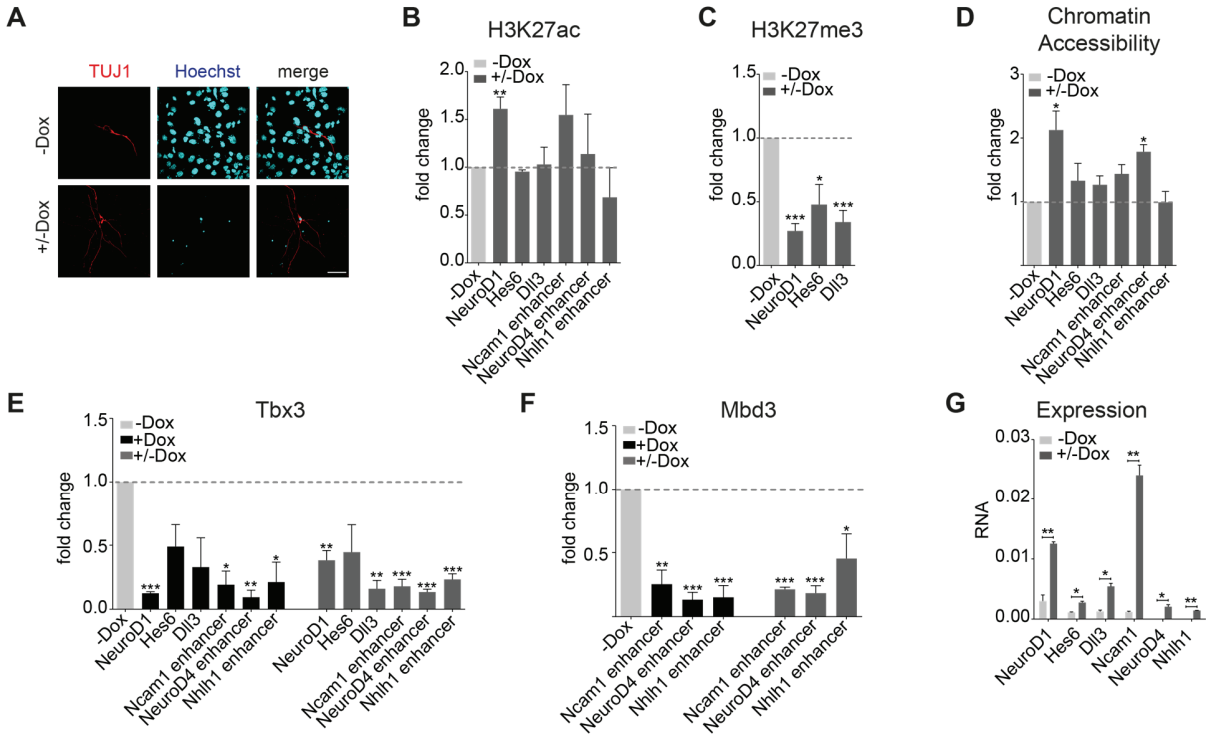


Figure 61: The effect of transient NeuroD1 action persists through epigenetic mechanisms.
(A) Immunocytochemistry analysis for TUJ1-positive cells after seven days of differentiating A2lox.NeuroD1 mESCs into iTNs. Cells were deprived of doxycycline after 48 h (+/-Dox) or left untreated (-Dox) as control. The formation of TUJ1-positive cells exhibiting a neuron-like morphology was visible. Nuclei were visualized with Hoechst. Scale bar, 50 µm. (B) ChIP-qPCR analysis for the fold change of H3K27ac enrichment at NeuroD1 target regulatory elements after seven days of iTN formation between cells exposed to transient ectopic *NeuroD1* induction (+/-Dox) and non-induced control cells (-Dox). Not all target regulatory elements gained enrichment of the active H3K27ac mark. (C) ChIP-qPCR analysis for the fold change of H3K27me3 enrichment at NeuroD1 target regulatory elements after seven days of iTN formation between cells exposed to transient ectopic *NeuroD1* induction (+/-Dox) and non-induced control cells (-Dox). All target promoter elements lost enrichment of the repressive H3K27me3 mark upon iTN differentiation. (D) FAIRE-qPCR analysis for the fold change of chromatin accessibility at NeuroD1 target regulatory elements after seven days of iTN formation between cells exposed to transient ectopic *NeuroD1* induction (+/-Dox) and non-induced control cells (-Dox). NeuroD1 target regulatory regions displayed a tendency for increased accessible chromatin upon iTN differentiation. (E, F) ChIP-qPCR analysis for the fold change of Tbx3 (E) or Mbd3 (F) enrichment at target genomic elements after seven days of iTN formation comparing continuously induced (+Dox) as well as transiently induced (+/-Dox) iTNs with control cells (-Dox). Both factors were replaced form NeuroD1 target *cis*-regulatory regions upon iTN differentiation. (G) RT-qPCR analysis for the expression of representative URT at day seven of iTN differentiation comparing cells exposed to transient ectopic *NeuroD1* induction (+/-Dox) and non-induced control cells (-Dox). All URT associated to NeuroD1 target elements were transcriptionally induced.
Figure information: Error bars reflect standard error of the mean from three (B - F) or two (G) biological replicates. The y-axis of ChIP-qPCR results in (D) shows the relative ChIP enrichment plotted as the ratio of precipitated DNA (bound) to input DNA and further normalized to an intergenic control region (background). The fold change of ChIP-qPCR enrichment and FAIRE-qPCR signal between tested conditions (C - F) is plotted as the fold change of the average enrichment of bound DNA normalized to the respective genomic DNA input on the y-axis. RNA fold change (G) reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*) plotted as the fold change between tested conditions. Significance was determined by t-test with * P < 0.05, ** P < 0.01 and *** P < 0.001. Data derived by JJ. All panels are taken from Pataskar & Jung et al., 2016.

3.10 Working model

Taken together, the findings of this chapter provided novel insights into the function of NeuroD1 mediating the onset of neuronal cell fate acquisition. These data encompassed insights into the global genomic target regions of NeuroD1, including both, promoters as well as enhancers of neuronal genes, which were bound sequence specifically. Interestingly, several of these NeuroD1 targets included transcriptional regulators. NeuroD1 bound its target regulatory elements despite them being in a heterochromatin state. This occupancy initiated a set of events that included the reprogramming of heterochromatin to euchromatin, which led to the transcriptional activation of associated genes. Strikingly, the transient action of NeuroD1 was sufficient to induce this alteration of the chromatin and transcription factor landscape underlying neuronal development. The proposed model (Figure 62) depicts the events which have been investigated at NeuroD1 target promoters in this thesis. However, the obtained observations also supported a similar activation of enhancer elements and concomitant their associated genes due to the action of NeuroD1.

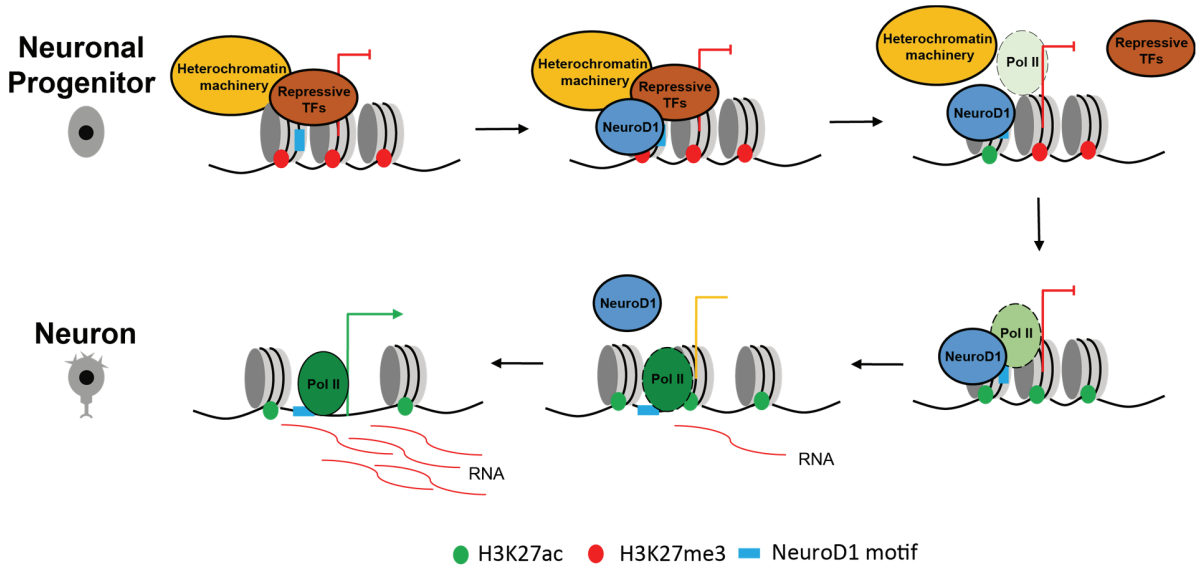


Figure 62: NeuroD1 reprograms chromatin and transcription factor landscapes to induce neuronal program.
NeuroD1 target promoters were embedded in heterochromatic environment. Upon induction, NeuroD1 binds to its target sites by recognizing its sequence motif which was followed by replacement of heterochromatin machinery and repressive transcription factors. The loss of inactive (H3K27me3) and gain of active (H3K27ac) histone marks was accompanied by increased chromatin accessibility, leading to recruitment of RNA polymerase II and gene expression. Similar dynamics were observed at NeuroD1 target enhancers except that H3K27me3 was absent at these sites prior to NeuroD1 binding. Consequently, NeuroD1 induced a gene regulatory program that specified neuronal fate. Despite a transient NeuroD1 action during neuronal differentiation, the transcriptional induction of target neuronal genes persisted via an epigenetic memory. Model developed by JJ. Figure taken from Pataskar & Jung et al., 2016.

4 Results Part B - Zfp354c

The KRAB zinc-finger protein Zfp354c is essential to repress retrotransposons during embryonic neurogenesis

The following section encompasses results which have been derived by a collaborating team: *In utero* electroporations and FAC-sorting of cortical cells were performed by Florian Noack (FN, Lab of Federico Calegari, Dresden) and Vilma Rraklli (VR, Lab of Johan Holmberg, Stockholm). *In vitro* as well as *in vivo* phenotypic characterisation and H3K9me3 ChIPs were performed by Angela Garding (AG). Computational analyses were carried out by Abhijeet Pataskar (AP). All other experiments were performed by myself (JJ). Each figure legend states which data was not derived by me.

The process of neuronal differentiation is known to be regulated by multiple cell-intrinsic and cell-extrinsic mechanisms in a spatio-temporal manner [90], wherein the role of transcription factors and epigenetic regulators is increasingly appreciated [91-102]. Zinc-finger containing proteins (ZFPs) represent one of the largest families of such regulatory proteins in mammals and are involved in a plethora of biological processes [103-107]. However, despite few well established examples a comprehensive picture of the involvement of zinc-finger proteins underlying the control of gene regulation and genomic integrity remains elusive.

4.1 Identification of Zfp354c as a neuronal KRAB-containing zinc-finger protein

4.1.1 Expression profiling of murine zinc-finger proteins in E14.5 tissues

Although cell fate changes during embryonic development are well characterised, the gene regulatory circuitry employed during phenotypic remodeling and commitment towards the neuronal lineage is poorly understood. Zinc-finger proteins are important players in a plethora of biological processes, however, their role during embryonic neurogenesis remains largely undefined. Given the diversity of ZFPs, the first question addressed was which genes in the mouse genome are coding for proteins harboring at least one zinc-finger domain. To this end, computational approaches were applied to identify all genes coding for zinc-fingers in the murine genome (data not shown). In order to get further insights into the expression pattern of the identified 1,674 ZFPs (“ZincOme”), several publicly available RNA-seq datasets of different embryonic murine tissues (E14.5) of several lineages were analysed. This investigation revealed several clusters of ZFPs which were highly expressed in a tissue-specific manner (Figure 63A). To select for ZFPs potentially important for embryonic neurogenesis, the cluster of genes specifically highly expressed in E14.5 brain (n = 445) was further investigated. The expression profile of these genes in defined layers of the E14.5 cortex (VZ, SVZ and CP) was compared along with tissues from the remaining lineages (Figure 63B). Interestingly, the majority of these brain-specific ZFPs (“NeuroZincOme”, see 8.7) was upregulated in the

CP in comparison to the germinal cortical layers VZ and SVZ, suggesting that these genes might play a role after the onset of neuronal differentiation.

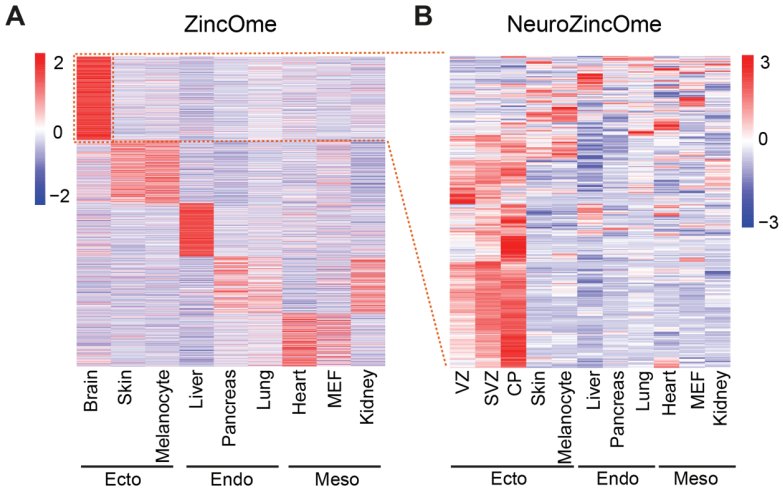


Figure 63: Identification of brain-specific zinc-finger proteins.
(A) Heatmap depicting the expression of 1,674 murine genes encompassing at least one zinc-finger domain (ZincOme) in different E14.5 tissues employing RNA-seq datasets. One cluster showed a selectively high expression in the embryonic brain (dotted orange line). (B) Heatmap depicting expression of 445 brain-specific zinc-finger genes (NeuroZincOme) from (A) in different E14.5 tissues including three cortical layers (VZ/SVZ/CP) of the same developmental age. Most ZFPs were highly expressed in the cortical plate and exhibited a low expression level in tissues of other lineages.
Figure information: Each row in the heatmaps represents a gene whose RNA-seq expression value is scaled from blue (low expression) to red (high expression). Biological replicates have been used for analysis. VZ, ventricular zone; SVZ, subventricular zone; CP, cortical plate; Ecto, ectoderm; Endo, endoderm; Meso, mesoderm. Computational analyses performed by AP.

4.1.2 Selection of *Zfp354c* as a neuronal KRAB-containing zinc-finger protein for further characterisation

One of the largest class of poly zinc-finger containing proteins harbors a Krueppel-associated box (KRAB), which is found in their amino-terminal region. Interestingly, the combination of this domain with C-terminal zinc-fingers is exclusively found in tetrapods [104, 107, 108]. These KRAB-ZFPs seem to have evolved to silence newly emerging endogenous retroviruses (ERVs) that colonize the host genome [109-111]. As the majority of studies on the regulation of embryonic neurogenesis are focused on the coding genome, the KRAB-ZFPs and their potential regulatory influence on the non-coding genome during embryonic development represents an exciting field of study. As KRAB-ZFPs are largely unexplored in the context of embryonic neurogenesis, the NeuroZincOme genes were probed for the presence of KRAB-ZFPs to unravel potentially novel regulators in the context of neuronal development.
Zfp354c was identified as a candidate for further studies. Despite its membership in the family of KRAB-containing zinc-finger proteins, very little was known about this protein: It had been suggested to act as a negative regulator of osteogenesis [112] and to be regulated by bone morphogenetic protein (BMP) signaling [113] as well as retinoic acid (RA) signaling [114]. As BMP- [115], and RA signaling [116] are

known to be involved in neurogenic processes, this data provided a first hint for a possible regulation of *Zfp354c* during neurogenesis. Its potential role during neurogenesis was further supported by a study which specifically detected this KRAB-ZFP in human fetal brain [117, 118]. Additionally, *Zfp354c* was among the NeuroZincOme candidates continuously transcriptionally induced from VZ to SVZ and CP and almost not expressed in other tissues shown in Figure 63B.

The 64 kDA Zfp354c protein contains 11 C2H2 zinc-fingers that are all encoded by a single exon at the C-terminus and an N-terminal KRAB domain (Figure 64A). The expression pattern of *Zfp354c* was confirmed to be specifically high in the embryonic cortex by a RT-qPCR analysis of embryonic tissues (Figure 64B). In order to probe for an indication of regulatory function in the context of DNA, the next question was whether this KRAB-ZFP is localized to the cell nucleus. As the commercially available antibodies for this largely unexplored protein where very limited and poorly characterised, Zfp354c was fused with an N-terminal HA-tag and FLAG-tag. In order to generate the pHAFlagZfp354c vector, the coding sequence (CDS) for Zfp354c was retrieved from NCBI (NM_013922) and custom synthesized at GeneOracle. Next, the CDS was cloned into the pHAFlagAgo1 destination vector (data not shown), leading to the N-terminal fusion of a HA-FLAG-tag with the Zfp354c CDS. The pHAFlagZfp354c vector was transfected into Neuro2a cells [119-121] which served as a model system to study the functional properties of this largely uncharacterised protein. An immunocytochemistry was performed for the HA- as well as the FLAG-tag to avoid possible false interpretations by using only a single antibody (Figure 64C). Strikingly, Zfp354c exhibited a predominant nuclear localization in Neuro2a cells after ectopic expression, suggesting that this protein might act on chromatin.

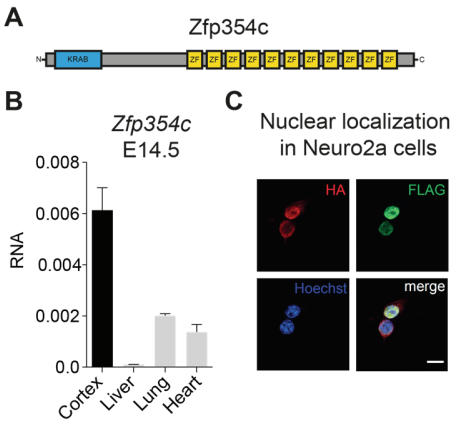


Figure 64: *Zfp354c* is a nuclear KRAB-containing zinc-finger protein highly expressed in the embryonic brain.
(A) Schematic representation of the Zfp354c protein. The blue box represents the KRAB domain and the yellow boxes indicate the 11 C2H2 zinc-fingers. (B) RT-qPCR analysis for the expression of *Zfp354c* in E14.5 tissues. *Zfp354c* showed the highest expression in the embryonic cortex (black) in comparison to other tested tissues (grey). Error bars reflect standard error of the mean from three biological replicates. RNA reflects the relative gene expression normalized to a housekeeping gene (*Rpl19*) plotted on the y-axis. (C) Immunocytochemistry in Neuro2a cells for the HA-tag and FLAG-tag of HA-FLAG-Zfp354c and Hoechst (nucleus) 48 h after transfection with the pHAFlagZfp354c vector. The HA- and FLAG-tagged Zfp354c protein localized to the nucleus. Scale bar: 10 µm. Data derived by JJ.

4.2 Zfp354c interacts via its KRAB domain with KAP1 and requires the zinc-finger cluster for nuclear localization

The nuclear localization of Zfp354c was furthermore interesting as KRAB-containing zinc-finger proteins have been described to facilitate target site specific heterochromatin formation via interaction with a protein complex containing the nuclear corepressor KAP1/Trim28 (KRAB-associated protein 1/ Tripartite motif-containing 28) [108, 111, 122]. It was tempting to address whether Zfp354c is a novel interaction partner of KAP1. To this end, SILAC mass spectrometry (SILAC-MS) was performed in Neuro2a cells transfected with the pHAFlagZfp354c vector. Strikingly, the quantitative SILAC-MS approach identified KAP1 as an interaction partner of Zfp354c (Figure 65A). This analysis revealed two other interaction partners, Importin-7 (Ipo7) and Importin subunit beta-1 (Kpnb1), which were shown to function in the context of nuclear protein import [123-126] and might be essential to shuttle Zfp354c into the nucleus.

Having observed the interaction with the nuclear corepressor component KAP1, it was interesting to unravel which domain of Zfp354c was responsible for this interaction. By employing site-directed mutagenesis, two mutant Zfp354c proteins were generated in addition to wild type (WT) Zfp354c (Figure 65B): In the first mutant, the KRAB domain was removed (KRAB mutant) whereas in the second mutant the poly zinc-finger cluster (ZFC) was deleted (ZFC mutant). The HA-FLAG-tagged WT Zfp354c as well as the KRAB and ZFC mutants were transfected into Neuro2a cells. Next, co-immunoprecipitation assays against the FLAG-tagged proteins were carried out and subjected to western blot analysis for KAP1 and the HA-tag (Figure 65C). This approach validated the interaction of Zfp354c with KAP1 as determined by SILAC-MS. Furthermore, such interaction was independent of the zinc-finger cluster as the ZFC mutant could also pull down KAP1 from the Neuro2a extract. However, the co-immunoprecipitation of KAP1 in whole cell extracts was not observed for the KRAB mutant (Figure 65C), suggesting that indeed the KRAB domain of Zfp354c mediates the interaction with KAP1. The next question addressed was which domain of Zfp354c was essential for its nuclear localization. To this end, the previously used three HA-FLAG constructs where transfected into Neuro2a cells and an immunocytochemistry assay against their FLAG-tag was carried out. Interestingly, the deletion of the whole ZFC prevented the nuclear localization of Zfp354c whereas the removal of the KRAB domain did not interfere with Zfp354c's subcellular localization pattern (Figure 65D).

Taken together, these results indicated that Zfp354c is a KRAB-ZFP which requires its zinc-fingers for nuclear localization and interacts with KAP1 mediated by its KRAB domain. This protein therefore is likely to act as a mediator of KAP1-facilitated heterochromatin formation.

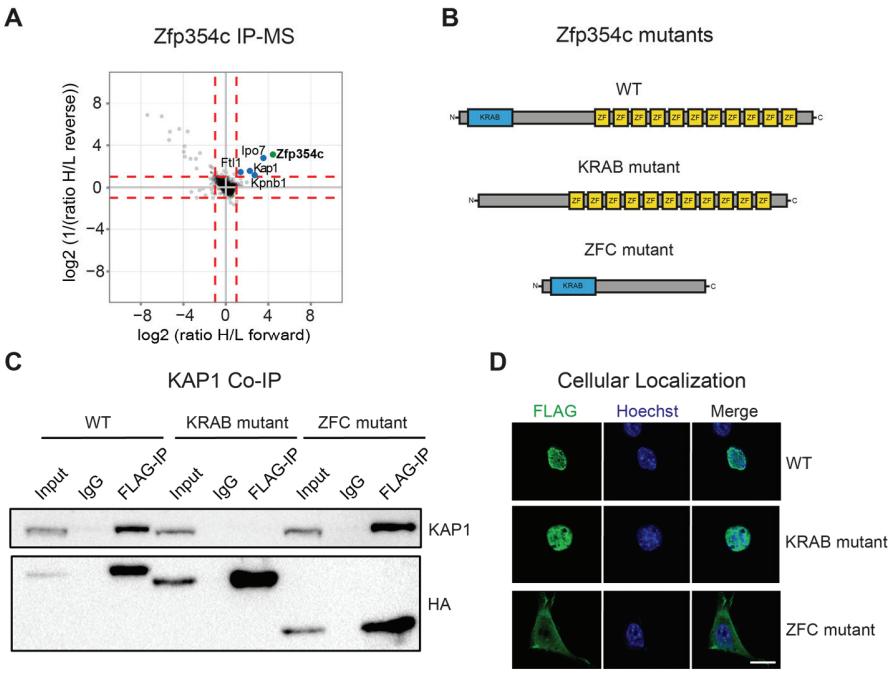


Figure 65: The KRAB domain of Zfp354c is crucial for the interaction with KAP1 while its zinc-finger cluster is essential for nuclear localization.
(A) Scatter blot of Zfp354c interaction partners (blue dots) enriched in SILAC forward as well as reverse IP-MS in Neuro2a cells 48h after transfection with pHAFlagZfp354c vector. Zfp354c interacted e.g. with KAP1 in Neuro2a cells. The heavy (H) and light (L) SILAC medium was supplemented with 20 µM retinoic acid 24 h after transfection to induce neuronal differentiation for 24 h. The bait protein (HA-FLAG-Zfp354c) is depicted with a green dot. The experiment was replicated in biological triplicates. (B) Schemes showing the domain structure of wild type (WT) Zfp354c and two domain mutants: The KRAB mutant lacks the KRAB domain whereas the zinc-finger cluster mutant (ZFC) lacks all eleven zinc-fingers. These three Zfp354c constructs were generated with an N-terminal HA-FLAG-tag (pHAFlagZfp354c, pHAFlagZfp354c-KRAB-mutant and pHAFlagZfp354c-ZFC-mutant). (C) Western blot assay for KAP1 and HA-tag on FLAG-IPs performed in Neuro2a cells 48 h after transfection of the three HA-FLAG Zfp354c constructs from (B). The KRAB domain of Zfp354c was essential for the interaction with KAP1 as no Co-IP was detected after pull down of the HA-FLAG KRAB mutant construct. The transfected cells were exposed to 20 µM retinoic acid after 24 h to induce neuronal differentiation for 24 h. Western blot detection was carried out with KAP1 and HA antibodies after SDS-PAGE in a 8% gel where 10% of the respective input material was loaded. (D) Immunocytochemistry of Neuro2a cells for Hoechst (nucleus) and the FLAG-tag of the three HA-FLAG-Zfp354c constructs from (B) 48 h after transfection with the respective vectors. Whereas the WT Zfp354c and KRAB mutant construct are localized in the nucleus, the ZFC mutant shows a strong enrichment in the extra-nuclear space (cytoplasm). Scale bar: 10 µm. Data derived by JJ.

4.3 *Zfp354c* is critical for neuronal differentiation *in vitro*

The identification of Zfp354c as an unexplored neuronal KRAB-ZFP interacting with KAP1 raised the question to which extend this protein might be important for neuronal differentiation. In order to study the role of *Zfp354c* during neurogenesis, an established culture system for neuronal differentiation of Neuro2a cells by applying retinoic acid (RA) as well as serum starvation was employed [120]. An expression analysis for *Zfp354c* during this *in vitro* differentiation was performed using a RT-qPCR assay (Figure 66A). Interestingly, *Zfp354c* was transcriptionally induced during neuronal differentiation of Neuro2a cells. Intrigued by its expression profile, the next question was whether *Zfp354c* depletion would impair neuronal differentiation *in vitro*. Therefore, a short hairpin RNA (shRNA) against the 3'-UTR of *Zfp354c* and a

scrambled control shRNA were designed in order to knock down (KD) *Zfp354c* by utilizing the cellular mechanism of RNA interference (RNAi) [127-132]. These shRNAs were cloned into the shRNA expression vector pRNAT-H1.1/Hygro (data not shown), therefore generating the pH1.shNTC (scrambled) and pH1.shZfp354c (shZfp354c) vectors. This vector facilitated the expression of a shRNA from a H1 promoter whereas the GFP marker is expressed from a second promoter. After an initial test of the KD efficiency *in vitro* (data not shown), Neuro2a cells stably expressing the shRNA against *Zfp354c* (shZfp354c) or the scrambled control shRNA (scrambled) were generated (data not shown).

These stable cell lines were differentiated for five days with RA and the KD efficiency of *Zfp354c* was assessed by RT-qPCR analysis comparing the scrambled and the shZfp354c cells (Figure 66B). The expression level of *Zfp354c* was depleted by around 70%, suggesting that these cells serve as a suitable model to study potential deficits in neurogenesis mediated by diminished levels of *Zfp354c*. Next, the distribution of the neuronal marker TUJ1 in differentiated scrambled and shZfp354c Neuro2a cells was assessed with an immunocytochemistry assay (Figure 66C). Interestingly, whereas the scrambled cells exhibited neuron-like characteristics by formation of TUJ1-positive processes upon five days of neuronal differentiation, this morphology was absent in *Zfp354c*-depleted cells, which appeared almost completely roundish. In addition, an increased membranous localization of phalloidin (F-actin) was observed in the same shZfp354c cells by immunocytochemistry analysis (data not shown), suggesting an increased epithelial nature of differentiated *Zfp354c*-depleted Neuro2a cells.

Furthermore, this immunocytochemistry analysis showed that the shZfp354c Neuro2a cells exhibited a substantial increase in nuclear area in comparison to the scrambled control as measured by Hoechst staining (Figure 66D). As the previous IP-MS analysis strengthened the link to the post-translational histone modification H3K9me3 via the interaction partner KAP1 (see Figure 65), the question arose whether the observed phenotype was influenced by alterations in the levels of this heterochromatin mark. A western blot analysis was carried out for H3K9me3 in scrambled and shZfp354c cells for two conditions: after day five of differentiation (d5) and in undifferentiated (d0) cells (Figure 66E). Interestingly, although the total level of H3K9me3 increased during differentiation in both cell lines, there was no difference detectable between control cells and *Zfp354c* knock down cells. The next question addressed was whether the nuclear distribution of H3K9me3 might be altered upon *Zfp354c* KD during *in vitro* differentiation. To examine this, an immunocytochemistry analysis for H3K9me3 was carried out in scrambled and shZfp354c cells after five days of RA mediated differentiation. Interestingly, an altered nuclear localization of this mark in knock down cells was observed as the speckled pattern of H3K9me3 foci found in control cells was completely lost (Figure 66F - G).

Taken together, these observations suggested that *Zfp354c* is induced during and required for neuronal differentiation *in vitro.* In addition, *Zfp354c* depletion had an impact on H3K9me3-marked heterochromatin domains.

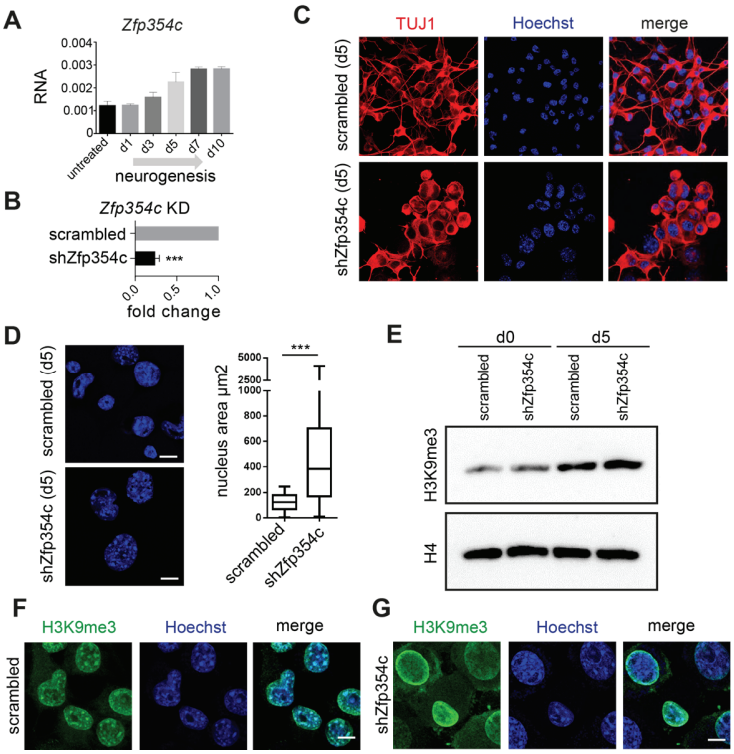


Figure 66: *Zfp354c* is transcriptionally induced during *in vitro* neurogenesis and essential for neuronal process formation as well as nuclear distribution of H3K9me3 foci.
(A) RT-qPCR analysis for *Zfp354c* expression during retinoic acid mediated differentiation of Neuro2a cells. *Zfp354c* is transcriptionally induced during the differentiation of Neuro2a cells. (B) Expression of *Zfp354c* in five day differentiated Neuro2a cells which are stably expressing shRNAs, either scrambled control (grey) or shZfp354c (black). The expression of a shRNA against *Zfp354c* (shZfp354c) led to a *Zfp354c* knock down (KD) of around 70% on RNA level. (C) Immunocytochemistry of Neuro2a cells (scrambled and shZfp354c) for TUJ1 (neuronal marker, red) and Hoechst (nucleus, blue) after five days of retinoic acid mediated differentiation. The KD of *Zfp354c* impaired the formation of TUJ1-positive process morphology which was observed for the scrambled control. (D) (Left panel) Representative immunocytochemistry image of nuclei (Hoechst) from scrambled and shZfp354c Neuro2a cells on day five of differentiation. Scale bar: 10µm. (Right panel) Bar plot showing the measurement of the nuclear area in µm² plotted on the y-axis for four biological replicates of scrambled and shZfp354c Neuro2a cells at day five of differentiation. The nuclear area significantly increased upon KD of *Zfp354c* in Neuro2a cells. (E) Western blot analysis for H3K9me3 and H4 (loading control) of scrambled and shZfp354c Neuro2a cells at d0 and d5 of differentiation. Whereas the global levels of H3K9me3 increased upon differentiation, no difference was detected in these levels upon *Zfp354c* KD. (F - G) Immunocytochemistry of Neuro2a cells [scrambled (F) and shZfp354c (G)] for H3K9me3 (green) and Hoechst (nucleus, blue) after five days of differentiation. The global distribution of H3K9me3 was altered upon *Zfp354c* KD as H3K9me3 foci (green spots in scrambled cells) were depleted. Scale bar: 10µm.
Figure information: RNA reflects the relative gene expression normalized to a housekeeping gene (Rpl19). The fold change in (B) represents the RNA expression in the KD condition relative to the corresponding scrambled control. Error bars reflect standard error of the mean from three (A) or four (B) biological replicates. Significance was determined by t-test with *** $P < 0.001$. Scrambled control refers to Neuro2a.shNTC cells and shZfp354c to Neuro2a.shZfp354c cells. Data in (A) derived by JJ, data in (C - G) derived by AG.

4.4 Zfp354c is involved in the repression of repetitive elements during neurogenesis *in vitro*

Having observed a phenotype of altered *in vitro* neuritogenesis and nuclear distribution of H3K9me3 upon *Zfp354c* KD, the next aim was to perform a more mechanistic investigation of Zfp354c's function on

chromatin level. As previous studies reported a link between KRAB-ZFPs and the regulation of endogenous retroviruses (ERVs) [110, 111, 122, 133-135], it was intriguing to explore a potential relationship between Zfp354c and the regulation of repetitive elements during neurogenesis.

4.4.1 Zfp354c binds to repetitive elements *in vitro*

In order to unravel the global genomic targets of the nuclear Zfp354c protein, chromatin immunoprecipitation for Zfp354c followed by next-generation sequencing (ChIP-seq) was carried out in differentiated Neuro2a cells after ectopic expression of HA-FLAG-Zfp354c. Computational analyses revealed that Zfp354c was highly enriched at repetitive elements (Figure 67A). However, due to the repetitive nature of these target sequences, the ChIP-seq enrichment could only be calculated for the different groups of repetitive elements instead for their single genomic locations. For validation of genomic target regions, a ChIP-qPCR analysis for FLAG-Zfp354c was carried out in Neuro2a cells. Therefore, qPCR primers were designed for specific groups of repetitive elements, including LINE1 (Long interspersed nuclear element 1) as well as ERV elements like IAP (Intracisternal A-Particle) or MMVL30 (*Mus musculus* virus-like 30). The primer sequences were either taken from published resources [122, 135, 136] or designed based on the repbase consensus sequence for specific groups of repetitive elements [137]. The ChIP-qPCR analysis validated that Zfp354c was occupying several groups of repetitive elements, including ERVs, to various degrees (Figure 67B). To get further insight in Zfp354c function, a *de novo* motif analysis was performed on the Zfp354c ChIP-seq peaks in Neuro2a cells (Figure 67C). This analysis suggested that Zfp354c targets repetitive elements likely in a sequence specific manner but that these motifs were also present in non-repeat regions.

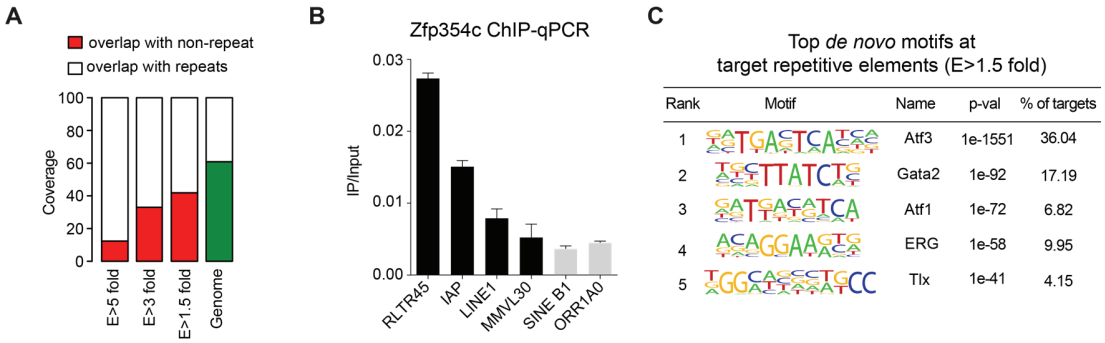


Figure 67: Zfp354c binds to repetitive elements encompassing ERV retrotransposons *in vitro*.
(A) Bar plot representing distribution of Zfp354c ChIP-seq enriched peaks among repetitive as well as non-repetitive genomic regions with different enrichment (E) cut-offs. The murine genome includes around 40% repetitive elements, which are increasingly occupied by Zfp354c by applying more stringent enrichment cut-offs. ChIP-seq was performed in Neuro2a cells after five days of RA induced differentiation and transfection with the pHAFlagZfp354c. (B) ChIP-qPCR analysis of Zfp354c binding at target repetitive elements (black) and control elements (grey) with an enrichment < 1.5 fold. (C) Top five *de novo* predicted DNA motifs at target repetitive elements with a Zfp354c enrichment > 1.5 from (A). Around 36% of target repetitive elements exhibited a *de novo* motif which was closely related to the binding motif of the transcription factor Atf3.
Figure information: Biological replicates of FLAG-Zfp354c ChIP-seq experiments were used for analysis. Error bars in (B) reflect standard error of the mean from three biological replicates. Computational analyses performed by AP, data generated by JJ.

4.4.2 Knock down of *Zfp354c* depletes H3K9me3 at target repetitive elements *in vitro*

In order to investigate the role of Zfp354c at these repetitive elements it was intriguing to explore if *Zfp354c* depletion in Neuro2a cells would lead to altered H3K9me3 levels and possibly expression of target repetitive elements. To this end, the previously used scrambled control and shZfp354c Neuro2a cells were differentiated for five days and their transcriptome subsequently subjected to next-generation sequencing (RNA-seq). A computational analysis comparing *Zfp354c*-depleted Neuro2a cells with the scrambled control revealed that several groups of repetitive elements were differentially expressed (DE, data not shown). By comparing the RNA-seq data of shZfp354c/scrambled Neuro2a cells with the previously generated Zfp354c ChIP-seq data in Neuro2a cells it was shown that several of these DE repetitive elements were found to be targeted by Zfp354c (Figure 68A). The percentage of repetitive elements belonging to the class of LTRs (long terminal repeat containing retrotransposons; this class contains several ERV families) in the upregulated elements was higher in comparison to downregulated repetitive elements (Figure 68B). However, it was surprising to find nearly the same quantity of target Zfp354c elements being either up- or downregulated. Given the previous results of global mislocalized H3K9me3 heterochromatin upon *Zfp354c* KD *in vitro* (Figure 66F - G), it was next explored whether this could be explained by altered levels of H3K9me3 at repetitive elements. ChIP assays for H3K9me3 were performed in day five differentiated shZfp354c and scrambled control Neuro2a cells and the precipitated material was subjected to next-generation sequencing (ChIP-seq). Computational analyses revealed that the majority of Zfp354c bound repetitive elements lost H3K9me3 enrichment upon *Zfp354c* KD *in vitro* (n = 137) whereas only a minor fraction (n = 33) gained H3K9me3 enrichment (Figure 68C). Interestingly further, the majority of the repetitive elements which lost H3K9me3 heterochromatic state belonged to the class of LTRs (Figure 68D), which contain several families of ERVs.

Taken together, these data suggested that upon KD of *Zfp354c in vitro* several of its target repetitive elements are depleted for H3K9me3, especially in the class of LTR elements.

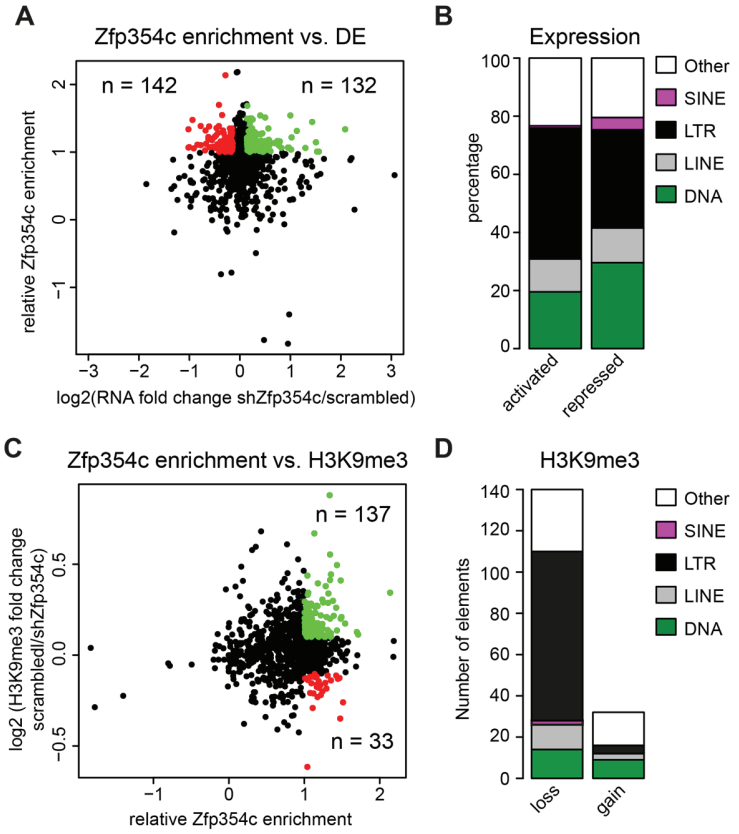


Figure 68: *Zfp354c* depletion accompanies a loss of H3K9me3 at target loci belonging to the class of LTRs.
(A) Scatter plot depicting *in vitro* differentially expressed repetitive element families upon depletion of *Zfp354c* in comparison to scrambled control in d5 differentiated Neuro2a cells in relation to their Zfp354c ChIP-seq enrichment. The x-axis depicts the RNA fold change upon *Zfp354c* depletion (shZfp354c) in comparison to scrambled control cells plotted as a log2 scale. The y-axis shows the relative Zfp354c enrichment in Neuro2a ChIP-seq assay. (B) Barplot showing the percentage distribution of several classes of repetitive elements among repressed (n = 142) as well as activated (n = 132) targets from (A). (C) Scatter plot depicting *in vitro* H3K9me3 dynamics upon KD of *Zfp354c* at repetitive elements in d5 differentiated Neuro2a cells in relation to their Zfp354c ChIP-seq enrichment. The x-axis depicts the relative Zfp354c enrichment in Neuro2a ChIP-seq assay. The y-axis shows the fold change of H3K9me3 enrichment upon *Zfp354c* KD (shZfp354c) in comparison to scrambled control Neuro2a ChIP-seq assay plotted as a log2 scale. (D) Bar plot showing the number of repetitive element groups among the Zfp354c target elements loosing (n = 137) as well as gaining H3K9me3 enrichment (n = 33) from (C).The majority of Zfp354c target repetitive elements belonged to the class of LTR retrotransposons.
Figure information: Biological replicates were used for analysis. Scrambled control cells refer to Neuro2a.shNTC and shZfp354c cells to Neuro2a.shZfp354c. DE, differential expression; SINE, short interspersed nuclear elements; LINE, long interspersed nuclear elements; LTR, long terminal repeat containing retrotransposons. Computational analyses performed by AP, H3K9me3 ChIP performed by AG, HA-Zfp354c ChIP performed by JJ.

4.5 Zfp354c represses repetitive elements via KAP1-mediated heterochromatin formation

The observed binding of Zfp354c to repetitive elements paralleled by a decrease of H3K9me3 levels upon KD preferentially at LTR elements led to the hypothesis that this phenomenon might be linked to the identified interaction partner KAP1, which is known to interact with a heterochromatin machinery in order to catalyze the formation of H3K9me3 at ERVs [110, 111, 122, 136, 138].

4.5.1 Depletion of *Zfp354c* leads to a transcriptional induction of target repetitive elements due to loss of H3K9me3 *in vitro*

In order to address whether KAP1 is recruited to repetitive elements via Zfp354c, a ChIP assay with an endogenous antibody against KAP1 was performed in scrambled and *Zfp354c* KD Neuro2a cells upon neuronal differentiation. The ChIP-qPCR analysis revealed that KAP1 enrichment was significantly reduced at Zfp354c target repetitive elements like LINE1 or the ERVs MMVL30 and IAP upon KD (Figure 69A). This observation suggested that *Zfp354c* depletion *in vitro* leads to a loss of recruitment of its interaction partner KAP1 to target repetitive elements. To validate that these repetitive elements were indeed loosing H3K9me3 in parallel, an independent ChIP-qPCR assay for H3K9me3 was performed in scrambled and *Zfp354c* KD Neuro2a cells upon neuronal differentiation (Figure 69B). This analysis validated the previous observation of the ChIP-seq experiment (see Figure 68C - D) as the levels of H3K9me3 were reduced by more than 50% at Zfp354c target repetitive elements upon KD of this KRAB-ZFP. Following this observation, the expression of target repetitive elements was profiled next in differentiated scrambled as well as shZfp354c cells by RT-qPCR analysis. As the previously designed ChIP-qPCR primers were directed against the consensus sequence of target repetitive elements, they could also be used for detecting their expression levels by RT-qPCR. Consistent with the loss of this repressive heterochromatin mark H3K9me3, the expression of several Zfp354c target repetitive elements was induced upon *Zfp354c* KD during neuronal differentiation *in vitro* in comparison to scrambled control cells (Figure 69C).

These observations suggested that Zfp354c recruits a heterochromatin machinery to repetitive elements via its interaction with KAP1 in order to silence them during neuronal differentiation *in vitro.*

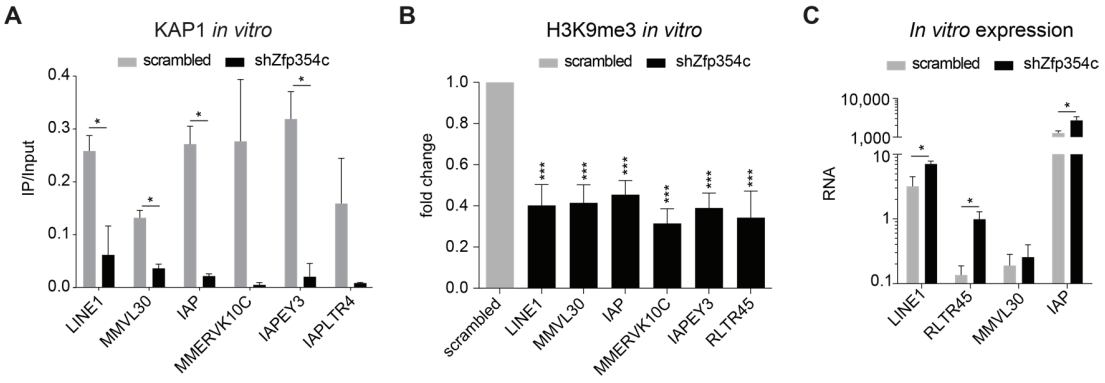


Figure 69: Zfp354c recruits KAP1 to repetitive elements for H3K9me3-mediated heterochromatin formation and silencing *in vitro*.

(A) ChIP-qPCR analysis for KAP1 enrichment at Zfp354c target repetitive elements in Neuro2a cells comparing shZfp354c (black) and scrambled control cells (grey). KAP1 was replaced at target repetitive elements upon *Zfp354c* depletion *in vitro*. The y-axis shows the ChIP DNA normalized to respective input DNA (IP/Input). (B) ChIP-qPCR analysis for the fold change of H3K9me3 enrichment at Zfp354c target repetitive elements in Neuro2a cells comparing shZfp354c cells (black) to scrambled control cells (grey). The repressive histone mark H3K9me3 was reduced at repetitive elements upon *Zfp354c* depletion *in vitro*. The fold change of ChIP-qPCR enrichment from shZfp354c to scrambled cells of the average enrichment of bound DNA normalized to the respective genomic DNA input is plotted on the y-axis. (C) RT-qPCR analysis for the expression of target repetitive elements in scrambled (grey) and shZfp354c Neuro2a cells (black). The target elements gained expression upon *Zfp354c* depletion *in vitro*. RNA is plotted as the relative expression normalized to a housekeeping gene (*Rpl19*) on the y-axis.

Figure information: The scrambled and shZfp354c Neuro2a cells were differentiated for five days (d5) with 20 µM retinoic acid prior performance of experiments. Error bars reflect standard error of the mean from three (A, C) and five (B) biological replicates. Significance was determined by t-test with * P < 0.05 and *** P < 0.001. Scrambled control cells refer to Neuro2a.shNTC and shZfp354c cells to Neuro2a.shZfp354c. Data in (B) derived by AG, data in (A) and (C) derived by JJ.

4.6 *Zfp354c* depletion impaires *in vivo* neurogenesis

4.6.1 Zfp354c target LTRs are repressed upon neuronal differentiation *in vivo*

Having observed the relationship of Zfp354c targeting to LTR elements and their transcriptional silencing during neuronal differentiation *in vitro*, it was tempting to investigate whether this mode of action could also be recapitulated in an *in vivo* developmental context. To this end, the expression of *Zfp354c* and several of its target ERVs was measured by RT-qPCR in the embryonic murine cortex at two developmental time points (E14.5 and E16.5). Interestingly, whereas the cortical expression of *Zfp354c* increased during embryonic development, the RNA levels of several target ERVs like IAP or MMERVK10C decreased significantly (Figure 70A). These data suggested that increasing levels of *Zfp354c* during cortical development *in vivo* might be involved in the transcriptional repression of target LTR elements. In order to further explore this relationship during in *vivo* development, the *in utero* electroporation (IUE) methodology was employed like previously described (see 3.8) in order to deplete *Zfp354c in vivo*. Either the pH1.shNTC (scrambled) or the pH1.shZfp354c (shZfp354c) vector were electroporated at E14.5 into the apical progenitor population of the embryonic cortex. At E18.5, the expression of the green fluorescent protein (GFP) from the pH1 vector backbone was used to FAC-sort the cell population exposed to either the expression of the scrambled or shZfp354c shRNA (data not shown). The two purified cell populations were subjected to RT-qPCR analysis, confirming a significant knock down of *Zfp354c* (Figure 70B). Although the tested ERVs did not exhibit a significant upregulation upon *Zfp354c* KD *in vivo*, a tendency for transcriptional induction was observed for some elements like MMVL30 or MMERVK10C. Taken together, this data suggested that Zfp354c might have an impact on ERV silencing during *in vivo* cortical development.

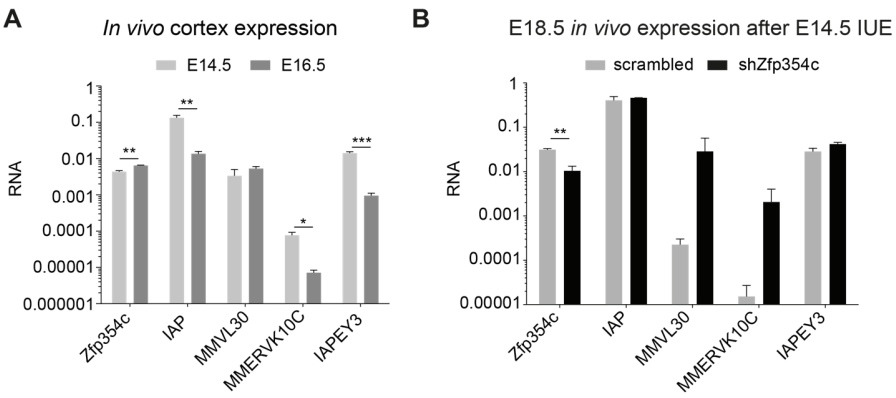


Figure 70: *Zfp354c* depletion increases the expression of target repetitive elements *in vivo*.
(A) RT-qPCR analysis for the expression of *Zfp354c* and representative target repetitive elements in the developing murine embryonic cortex. The expression of *Zfp354c* in the cortex increased from E14.5 (light grey) to E16.5 (dark grey), which was paralleled by a decrease of target repetitive element expression. (B) RT-qPCR analysis for the expression of

Zfp354c and representative target repetitive elements in E18.5 FAC-sorted cortical cells four days post IUE. The expression of *Zfp354c* was significantly reduced in the cell population exposed to the shRNA against *Zfp354c* (shZfp354c, black) in comparison to the control population expressing the scrambled control (grey). The expression levels of some Zfp354c target repetitive elements showed a tendency towards increased expression upon *Zfp354c* depletion *in vivo*. The IUE of pH1.shNTC (scrambled) and pH1.shZfp354c (shZfp354c) was performed at E14.5 and transfected cell populations were isolated at E18.5 by employing FACS methodology with a gating strategy for green (GFP) fluorescence.
Figure information: Error bars reflect standard error of the mean from three biological replicates. RNA reflects relative expression normalized to a housekeeping gene (*Rpl19*) on the logarithmic y-axis. Significance was determined by t-test with * P < 0.05, ** P < 0.01 and *** P < 0.001. IUE and FAC-sorting performed by FN and VR, data derived by JJ.

4.6.2 Impaired *in vivo* neurogenesis upon *Zfp354c* depletion depends on the KRAB domain

The phenotypic and chromatin analysis *in vitro* suggested that Zfp354c might be essential for proper neuronal differentiation due to heterochromatin-mediated silencing of repetitive elements. In addition, the *in vivo* expression analysis suggested an alteration in ERV expression upon *Zfp354c* depletion. However, it remained elusive whether *Zfp354c* depletion in apical progenitors indeed impacts on *in vivo* neuronal development. Towards this, *Zfp354c* was depleted during *in vivo* cortical development using the same *in utero* electroporation strategy at E14.5 as described above and the phenotypic characterisation was carried out four days later (E18.5). An immunohistochemical analysis was carried out on coronal cryosections of scrambled control and shZfp354c electroporated brains by staining for the manipulated cells (GFP-positive) and Tbr1 as a marker of the lowest neuronal layer. Strikingly, almost no *Zfp354c*-depleted GFP-positive cells of were found in the cortical plate (Figure 71A). A quantification of the cortical distribution taking the Tbr1-positive neuronal layer as border for the CP showed that 95% of *Zfp354c*-depleted cells were retained in lower cortical layers in comparison to 65% found in scrambled electroporated brains (Figure 71B). In addition, an immunohistochemical analysis for GFP-positive cells and the basal-progenitor marker Tbr2 revealed, that the co-localization of *Zfp354c*-depleted cells within the Tbr2-positive progenitor pool was increased upon *Zfp354c* KD in comparison to control (data not shown). This increase in Tbr2-positive cells in the *Zfp354c*-depleted pool further supported the observed retention phenotype. These results suggested that *Zfp354c*-depleted cells can transit from apical to basal progenitors but fail to reach the cortical plate. It was tempting to investigate whether the few electroporated KD cells which still reached the cortical plate would exhibit an altered morphology in comparison to the scrambled control cells. To this end, the shape of manipulated cells in the CP of both populations was analysed: cells devoid of *Zfp354c* exhibited a much rounder shape than control cells, which were more elongated (data not shown). This observation suggested an altered polarity of cells reaching the CP upon *Zfp354c* KD. This observation was in line with the previous *in vitro* findings (see Figure 66C). These data supported a functional role of *Zfp354c* during embryonic neurogenesis *in vivo*.
The next question was, whether the observed phenotype of reduced cell numbers in the CP upon *Zfp354c* depletion could be rescued and whether this is depended on the KRAB domain of Zfp354c. To address this question, the pH1.shZfp354c vector (co-expressing GFP) was co-electroporated at E14.5 with a pCIDRE vector (co-expressing RFP) either containing the CDS for WT Zfp354c (pCIDRE-Zfp354c) or for the KRAB

mutant (pCIDRE-Zfp354c-KRAB-mutant). The pCIDRE vector without an additional CDS was used as control (RFP control). As shZfp354c was directed against the 3'-UTR of *Zfp354c*, only the endogenous *Zfp354c* was targeted in electroporated cells whereas the co-electroporated WT and KRAB mutant were ectopically expressed. At E18.5, the co-electroporated brains were subjected to immunohistochemical analysis. The localization of cells co-electroporated with shZfp354c (GFP, green) and the respective pCIDRE vector (RFP, red) was visualized by their yellow fluorescent signal (Figure 71C - E). The *Zfp354c*-depleted cells which were co-electroporated with the control vector (RFP control) exhibited a similar retention in the VZ/SVZ and IZ layer as previously observed for the single shZfp354c electroporated population (Figure 71E, see Figure 71B). Strikingly, the proportion of *Zfp354c*-depleted cells crossing the Tbr1 border into the cortical plate upon the co-expression of WT Zfp354c was almost completely reverted to the level previously observed for scrambled control cells (Figure 71E, see Figure 71B). This observation suggested that the observed inability of cells to mature towards CP localized cells is indeed caused by the depletion of *Zfp354c* in apical progenitors. Furthermore, the analysis of *Zfp354c*-depleted cells co-electroporated with the Zfp354c KRAB mutant showed that in contrast to the WT construct, these double positive cells were similarly retained in the VZ/SVZ and IZ as observed for the control vector co-electroporation (Figure 71E). As the previous data has shown that the KRAB domain of Zfp354c interacts with KAP1 (see Figure 65C), this new observation supported the notion that impaired neuronal development upon *Zfp354c* KD might be caused by an altered KAP1 function. Interestingly, cells which only overexpressed the Zfp354c WT construct showed an increase of 30% in the localization to the CP in comparison to scrambled control cells (data not shown), suggesting an enhanced generation of cortical cells upon increased levels of *Zfp354c*.

These findings suggested that a fine tuned level of Zfp354c is necessary to enable proper neurogenesis *in vivo*.

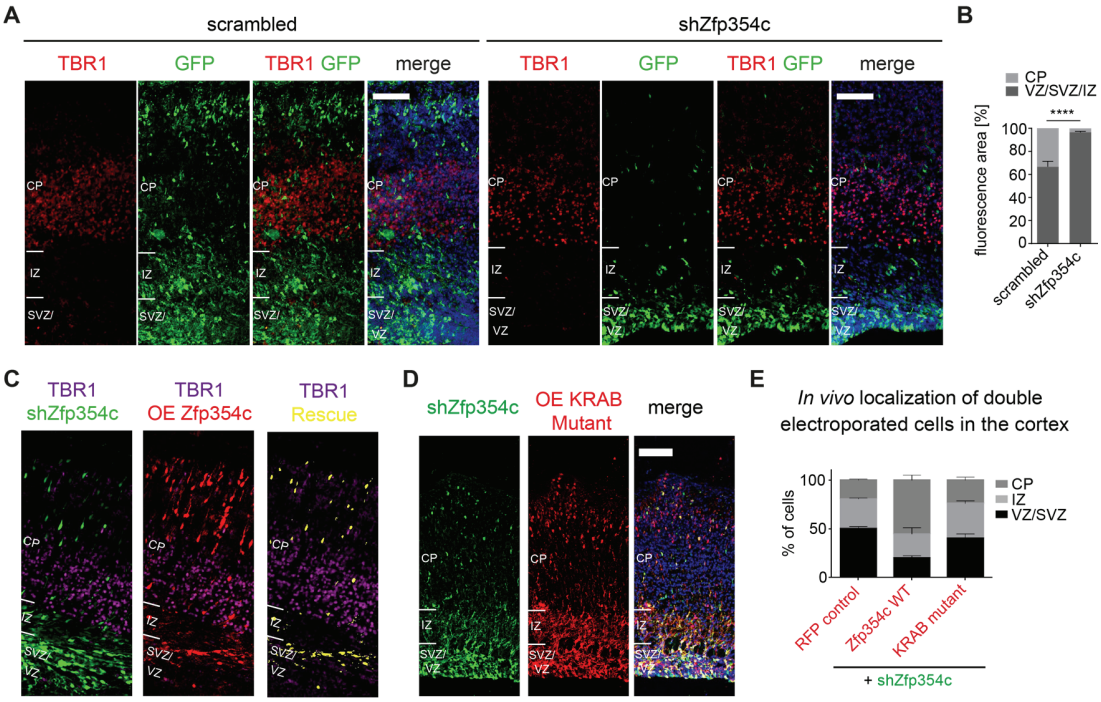


Figure 71: Depletion of *Zfp354c* impairs neurogenesis *in vivo* which cannot be the rescued by the KRAB mutant. (A) Immunohistochemistry for GFP (green, marks shZfp354c or scrambled transfected cells) and Tbr1 (lower layer neuronal marker, red) of coronal brain sections at E18.5. Representative images revealed that almost no cells depleted for *Zfp354c* localized in the cortical plate four days after IUE. (B) Quantification of the green fluorescence signal above (CP) or below the Tbr1 layer (VZ/SVZ/IZ) as illustrated in (A). Error bars reflect standard error of the mean from three representative regions each of three independently electroporated brains. Significance was determined by t-test with **** $P < 0.001$. (C) Same as in (A) but for brains co-electroporated with shZfp354c (green) and a rescue construct expressing the CDS of Zfp354c (red). The yellow signal showed only co-electroporated (rescued) cells. (D) Same as in (C) but the co-electroporation was carried out with a construct expressing the CDS for the Zfp354c KRAB mutant. (E) Quantification of the spatial distribution of double electroporated cells (shZfp354c + respective pCIDRE plasmid) in E18.5 coronal brain sections. The distribution of cells across the cortical layers upon *Zfp354c* depletion was rescued by the overexpression of WT Zfp354c but not the Zfp354c KRAB mutant as almost 50% of double electroporated cells were found in the VZ/SVZ like for the control condition (RFP control).
Figure information: *In utero* electroporations were performed at E14.5 and phenotypic characterisation was performed at E18.5 including staining for DNA (Hoechst). The green fluorescent protein (GFP) is expressed from the backbone of the transfected shRNA vectors pH1.shNTC (scrambled) and pH1.shZfp354c (shZfp354c). The red fluorescent protein is expressed from the backbone of the rescue vectors coding for WT Zfp354c (pCIDRE-Zfp354c), the Zfp354c KRAB mutant (pCIDRE-Zfp354c-KRAB-mutant) or the empty RFP control vector (pCIDRE). CDS, coding sequence; VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP, cortical plate. IUE performed by VR and FN, data generated by AG.

4.7 Working hypothesis

This study employed data mining methodology to generate a list of known and putative zinc-finger protein encoding genes in the mouse transcriptome. Using expression datasets from multiple cell types, lineage-specifically expressed zinc-finger proteins were identified. This collection can serve as a resource to delineate a potential lineage-specific function of these ZFPs during development considering several tissues. The KRAB-containing zinc-finger protein Zfp354c was identified as a functional unexplored factor highly expressed in neuronal tissues. This protein exhibited a nuclear localization which was dependent on its cluster of eleven zinc-fingers. Domain mutagenesis and Co-IP experiments further showed that Zfp354c

interacts with the nuclear corepressor KAP1 via its KRAB domain. Strikingly, depletion experiments for *Zfp354c* utilizing shRNAs revealed impaired neuronal differentiation *in vitro* as well as *in vivo*. By performing ChIP-seq analysis for Zfp354c it was observed that this KRAB-ZFP binds to several repetitive elements in the murine genome. Furthermore, especially the class of LTR elements which contains families of endogenous retroviruses (ERVs) was found to be regulated via H3K9me3-dependent heterochromatin formation due to the Zfp354c-mediated recruitment of KAP1 during neuronal differentiation *in vitro* and likely also *in vivo*. Altogether, these observations established Zfp354c as a novel epigenetic regulator of embryonic neurogenesis by controlling the repression of the non-coding genome encompassing repetitive elements during ontogeny.

4.8 References Results Part A and B

1. Pataskar, A., et al., *NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program.* EMBO J, 2016. **35**(1): p. 24-45.
2. Guillemot, F., *Spatial and temporal specification of neural fates by transcription factor codes.* Development, 2007. **134**(21): p. 3771-80.
3. Spitz, F. and E.E. Furlong, *Transcription factors: from enhancer binding to developmental control.* Nat Rev Genet, 2012. **13**(9): p. 613-26.
4. Jones, S., *An overview of the basic helix-loop-helix proteins.* Genome Biol, 2004. **5**(6): p. 226.
5. Hevner, R.F., et al., *Transcription factors in glutamatergic neurogenesis: conserved programs in neocortex, cerebellum, and adult hippocampus.* Neurosci Res, 2006. **55**(3): p. 223-33.
6. Aprea, J., M. Nonaka-Kinoshita, and F. Calegari, *Generation and characterization of Neurod1-CreER(T2) mouse lines for the study of embryonic and adult neurogenesis.* Genesis, 2014. **52**(10): p. 870-8.
7. Pang, Z.P., et al., *Induction of human neuronal cells by defined transcription factors.* Nature, 2011. **476**(7359): p. 220-3.
8. Guo, Z., et al., *In vivo direct reprogramming of reactive glial cells into functional neurons after brain injury and in an Alzheimer's disease model.* Cell Stem Cell, 2014. **14**(2): p. 188-202.
9. Mohn, F., et al., *Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors.* Mol Cell, 2008. **30**(6): p. 755-66.
10. Magnusdottir, E., et al., *Combinatorial control of cell fate and reprogramming in the mammalian germline.* Curr Opin Genet Dev, 2012. **22**(5): p. 466-74.
11. Wamstad, J.A., et al., *Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage.* Cell, 2012. **151**(1): p. 206-20.
12. Xie, W., et al., *Epigenomic analysis of multilineage differentiation of human embryonic stem cells.* Cell, 2013. **153**(5): p. 1134-48.
13. Jaenisch, R. and R. Young, *Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming.* Cell, 2008. **132**(4): p. 567-82.
14. O'Shea, K.S., *Embryonic stem cell models of development.* Anat Rec, 1999. **257**(1): p. 32-41.
15. Bibel, M., et al., *Embryonic stem cell-derived neurons as a novel cellular model system to study neurodegenerative and neuroregenerative processes in vitro.* J Stem Cells Regen Med, 2007. **2**(1): p. 62-3.
16. Bibel, M., et al., *Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells.* Nat Protoc, 2007. **2**(5): p. 1034-43.
17. Bibel, M., et al., *Differentiation of mouse embryonic stem cells into a defined neuronal lineage.* Nat Neurosci, 2004. **7**(9): p. 1003-9.
18. Iacovino, M., et al., *Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells.* Stem Cells, 2011. **29**(10): p. 1580-8.
19. Jung, J., *Gene Regulation by Topoisomerase IIa: A Dual Role in Embryonic Stem Cells.* Johannes Gutenberg-University, Mainz, 2013.
20. Sternberg, N. and D. Hamilton, *Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites.* J Mol Biol, 1981. **150**(4): p. 467-86.
21. Sternberg, N., et al., *Site-specific recombination and its role in the life cycle of bacteriophage P1.* Cold Spring Harb Symp Quant Biol, 1981. **45 Pt 1**: p. 297-309.
22. Sternberg, N., D. Hamilton, and R. Hoess, *Bacteriophage P1 site-specific recombination. II. Recombination between loxP and the bacterial chromosome.* J Mol Biol, 1981. **150**(4): p. 487-507.
23. Sauer, B. and N. Henderson, *Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1.* Proc Natl Acad Sci U S A, 1988. **85**(14): p. 5166-70.
24. Lee, G. and I. Saito, *Role of nucleotide sequences of loxP spacer region in Cre-mediated recombination.* Gene, 1998. **216**(1): p. 55-65.
25. Ghosh, K., F. Guo, and G.D. Van Duyne, *Synapsis of loxP sites by Cre recombinase.* J Biol Chem, 2007. **282**(33): p. 24004-16.
26. Grindley, N.D., *Site-specific recombination: synapsis and strand exchange revealed.* Curr Biol, 1997. **7**(10): p. R608-12.
27. Van Duyne, G.D., *A structural view of cre-loxp site-specific recombination.* Annu Rev Biophys Biomol Struct, 2001. **30**: p. 87-104.

28. Gossen, M. and H. Bujard, *Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.* Proc Natl Acad Sci U S A, 1992. **89**(12): p. 5547-51.

29. Gossen, M., et al., *Transcriptional activation by tetracyclines in mammalian cells.* Science, 1995. **268**(5218): p. 1766-9.

30. Wutz, A., T.P. Rasmussen, and R. Jaenisch, *Chromosomal silencing and localization are mediated by different domains of Xist RNA.* Nat Genet, 2002. **30**(2): p. 167-74.

31. Bronson, S.K., et al., *Single-copy transgenic mice with chosen-site integration.* Proc Natl Acad Sci U S A, 1996. **93**(17): p. 9067-72.

32. Hooper, M., et al., *HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells.* Nature, 1987. **326**(6110): p. 292-5.

33. Friedrich, G. and P. Soriano, *Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice.* Genes Dev, 1991. **5**(9): p. 1513-23.

34. Zambrowicz, B.P., et al., *Disruption of overlapping transcripts in the ROSA beta geo 26 gene trap strain leads to widespread expression of beta-galactosidase in mouse embryos and hematopoietic cells.* Proc Natl Acad Sci U S A, 1997. **94**(8): p. 3789-94.

35. Iacovino, M., et al., *A conserved role for Hox paralog group 4 in regulation of hematopoietic progenitors.* Stem Cells Dev, 2009. **18**(5): p. 783-92.

36. Kyba, M., R.C. Perlingeiro, and G.Q. Daley, *HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors.* Cell, 2002. **109**(1): p. 29-37.

37. Lee, J.E., et al., *Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein.* Science, 1995. **268**(5212): p. 836-44.

38. Lee, J.K., et al., *Expression of neuroD/BETA2 in mitotic and postmitotic neuronal cells during the development of nervous system.* Dev Dyn, 2000. **217**(4): p. 361-7.

39. Schwab, M.H., et al., *Neuronal basic helix-loop-helix proteins (NEX, neuroD, NDRF): spatiotemporal expression and targeted disruption of the NEX gene in transgenic mice.* J Neurosci, 1998. **18**(4): p. 1408-18.

40. Fietz, S.A., et al., *Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal.* Proc Natl Acad Sci U S A, 2012. **109**(29): p. 11836-41.

41. Shen, Y., et al., *A map of the cis-regulatory sequences in the mouse genome.* Nature, 2012. **488**(7409): p. 116-20.

42. Yue, F., et al., *A comparative encyclopedia of DNA elements in the mouse genome.* Nature, 2014. **515**(7527): p. 355-64.

43. Florio, M. and W.B. Huttner, *Neural progenitors, neurogenesis and the evolution of the neocortex.* Development, 2014. **141**(11): p. 2182-94.

44. Taverna, E., M. Gotz, and W.B. Huttner, *The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex.* Annu Rev Cell Dev Biol, 2014. **30**: p. 465-502.

45. Kriegstein, A. and A. Alvarez-Buylla, *The glial nature of embryonic and adult neural stem cells.* Annu Rev Neurosci, 2009. **32**: p. 149-84.

46. Lienert, F., et al., *Genomic prevalence of heterochromatic H3K9me2 and transcription do not discriminate pluripotent from terminally differentiated cells.* PLoS Genet, 2011. **7**(6): p. e1002090.

47. Stadler, M.B., et al., *DNA-binding factors shape the mouse methylome at distal regulatory regions.* Nature, 2011. **480**(7378): p. 490-5.

48. Tiwari, V.K., et al., *A chromatin-modifying function of JNK during stem cell differentiation.* Nat Genet, 2011. **44**(1): p. 94-100.

49. Tiwari, V.K., et al., *Target genes of Topoisomerase IIbeta regulate neuronal survival and are defined by their chromatin state.* Proc Natl Acad Sci U S A, 2012. **109**(16): p. E934-43.

50. Thakurela, S., et al., *Gene regulation and priming by topoisomerase IIalpha in embryonic stem cells.* Nat Commun, 2013. **4**: p. 2478.

51. Chen, W., et al., *Lentivirus carrying the NeuroD1 gene promotes the conversion from glial cells into neurons in a spinal cord injury model.* Brain Res Bull, 2017.

52. Brulet, R., et al., *NEUROD1 Instructs Neuronal Conversion in Non-Reactive Astrocytes.* Stem Cell Reports, 2017. **8**(6): p. 1506-1515.

53. Boyer, L.A., et al., *Core transcriptional regulatory circuitry in human embryonic stem cells.* Cell, 2005. **122**(6): p. 947-56.

54. Loh, Y.H., et al., *The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells.* Nat Genet, 2006. **38**(4): p. 431-40.

55. Rodda, D.J., et al., *Transcriptional regulation of nanog by OCT4 and SOX2.* J Biol Chem, 2005. **280**(26): p. 24731-7.

56. Silva, J., et al., *Nanog is the gateway to the pluripotent ground state.* Cell, 2009. **138**(4): p. 722-37.

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

58. Li, M. and J.C. Belmonte, *Ground rules of the pluripotency gene regulatory network.* Nat Rev Genet, 2017. **18**(3): p. 180-191.

59. Creyghton, M.P., et al., *Histone H3K27ac separates active from poised enhancers and predicts developmental state.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21931-6.

60. Rada-Iglesias, A., et al., *A unique chromatin signature uncovers early developmental enhancers in humans.* Nature, 2011. **470**(7333): p. 279-83.

61. Bonn, S., et al., *Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development.* Nat Genet, 2012. **44**(2): p. 148-56.

62. Zhu, Y., et al., *Predicting enhancer transcription and activity from chromatin modifications.* Nucleic Acids Res, 2013. **41**(22): p. 10032-43.

63. Shlyueva, D., G. Stampfel, and A. Stark, *Transcriptional enhancers: from properties to genome-wide predictions.* Nat Rev Genet, 2014. **15**(4): p. 272-86.

64. Jin, L. and R.V. Lloyd, *In situ hybridization: methods and applications.* J Clin Lab Anal, 1997. **11**(1): p. 2-9.

65. Kim, W.Y., *NeuroD regulates neuronal migration.* Mol Cells, 2013. **35**(5): p. 444-9.

66. Sahu, S.K., et al., *JNK-dependent gene regulatory circuitry governs mesenchymal fate.* EMBO J, 2015. **34**(16): p. 2162-81.

67. Lamouille, S., J. Xu, and R. Derynck, *Molecular mechanisms of epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol, 2014. **15**(3): p. 178-96.

68. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.

69. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* J Clin Invest, 2009. **119**(6): p. 1420-8.

70. Keller, G.M., *In vitro differentiation of embryonic stem cells.* Curr Opin Cell Biol, 1995. **7**(6): p. 862-9.

71. Keller, G., *Embryonic stem cell differentiation: emergence of a new era in biology and medicine.* Genes Dev, 2005. **19**(10): p. 1129-55.

72. Gaspar-Maia, A., et al., *Open chromatin in pluripotency and reprogramming.* Nat Rev Mol Cell Biol, 2011. **12**(1): p. 36-47.

73. Adam, R.C. and E. Fuchs, *The Yin and Yang of Chromatin Dynamics In Stem Cell Fate Selection.* Trends Genet, 2016. **32**(2): p. 89-100.

74. Kobayashi, H. and N. Kikyo, *Epigenetic regulation of open chromatin in pluripotent stem cells.* Transl Res, 2015. **165**(1): p. 18-27.

75. Allis, C.D. and T. Jenuwein, *The molecular hallmarks of epigenetic control.* Nat Rev Genet, 2016. **17**(8): p. 487-500.

76. Weidgang, C.E., et al., *TBX3 Directs Cell-Fate Decision toward Mesendoderm.* Stem Cell Reports, 2014. **2**(5): p. 747.

77. Xue, Y., et al., *NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities.* Mol Cell, 1998. **2**(6): p. 851-61.

78. Le Guezennec, X., et al., *MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties.* Mol Cell Biol, 2006. **26**(3): p. 843-51.

79. Basta, J. and M. Rauchman, *The nucleosome remodeling and deacetylase complex in development and disease.* Transl Res, 2015. **165**(1): p. 36-47.

80. Aprea, J., et al., *Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment.* EMBO J, 2013. **32**(24): p. 3145-60.

81. Stancik, E.K., et al., *Heterogeneity in ventricular zone neural precursors contributes to neuronal fate diversity in the postnatal neocortex.* J Neurosci, 2010. **30**(20): p. 7028-36.

82. Martello, G. and A. Smith, *The nature of embryonic stem cells.* Annu Rev Cell Dev Biol, 2014. **30**: p. 647-75.

83. Williams, R.L., et al., *Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells.* Nature, 1988. **336**(6200): p. 684-7.

84. Boeuf, H., et al., *Leukemia inhibitory factor-dependent transcriptional activation in embryonic stem cells.* J Cell Biol, 1997. **138**(6): p. 1207-17.

85. Niwa, H., et al., *Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3.* Genes Dev, 1998. **12**(13): p. 2048-60.

86. Niwa, H., et al., *A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells.* Nature, 2009. **460**(7251): p. 118-22.

87. Imayoshi, I. and R. Kageyama, *Oscillatory control of bHLH factors in neural progenitors.* Trends Neurosci, 2014. **37**(10): p. 531-8.

88. Imayoshi, I., et al., *Oscillatory control of factors determining multipotency and fate in mouse neural progenitors.* Science, 2013. **342**(6163): p. 1203-8.

89. Imayoshi, I. and R. Kageyama, *bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells.* Neuron, 2014. **82**(1): p. 9-23.

90. Paridaen, J.T. and W.B. Huttner, *Neurogenesis during development of the vertebrate central nervous system.* EMBO Rep, 2014. **15**(4): p. 351-64.

91. Asami, M., et al., *The role of Pax6 in regulating the orientation and mode of cell division of progenitors in the mouse cerebral cortex.* Development, 2011. **138**(23): p. 5067-78.

92. Gotz, M., A. Stoykova, and P. Gruss, *Pax6 controls radial glia differentiation in the cerebral cortex.* Neuron, 1998. **21**(5): p. 1031-44.

93. Hatakeyama, J., et al., *Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation.* Development, 2004. **131**(22): p. 5539-50.

94. Martynoga, B., D. Drechsel, and F. Guillemot, *Molecular control of neurogenesis: a view from the mammalian cerebral cortex.* Cold Spring Harb Perspect Biol, 2012. **4**(10).

95. Nieto, M., et al., *Expression of Cux-1 and Cux-2 in the subventricular zone and upper layers II-IV of the cerebral cortex.* J Comp Neurol, 2004. **479**(2): p. 168-80.

96. Osumi, N., et al., *Concise review: Pax6 transcription factor contributes to both embryonic and adult neurogenesis as a multifunctional regulator.* Stem Cells, 2008. **26**(7): p. 1663-72.

97. Sansom, S.N. and F.J. Livesey, *Gradients in the brain: the control of the development of form and function in the cerebral cortex.* Cold Spring Harb Perspect Biol, 2009. **1**(2): p. a002519.

98. Sugitani, Y., et al., *Brn-1 and Brn-2 share crucial roles in the production and positioning of mouse neocortical neurons.* Genes Dev, 2002. **16**(14): p. 1760-5.

99. Tarabykin, V., et al., *Cortical upper layer neurons derive from the subventricular zone as indicated by Svet1 gene expression.* Development, 2001. **128**(11): p. 1983-93.

100. Tuoc, T.C., R. Narayanan, and A. Stoykova, *BAF chromatin remodeling complex: cortical size regulation and beyond.* Cell Cycle, 2013. **12**(18): p. 2953-9.

101. Custo Greig, L.F., et al., *Molecular logic of neocortical projection neuron specification, development and diversity.* Nat Rev Neurosci, 2013. **14**(11): p. 755-69.

102. Sweatt, J.D., *The emerging field of neuroepigenetics.* Neuron, 2013. **80**(3): p. 624-32.

103. Laity, J.H., B.M. Lee, and P.E. Wright, *Zinc finger proteins: new insights into structural and functional diversity.* Curr Opin Struct Biol, 2001. **11**(1): p. 39-46.

104. Najafabadi, H.S., et al., *C2H2 zinc finger proteins greatly expand the human regulatory lexicon.* Nat Biotechnol, 2015. **33**(5): p. 555-62.

105. Jen, J. and Y.C. Wang, *Zinc finger proteins in cancer progression.* J Biomed Sci, 2016. **23**(1): p. 53.

106. Stubbs, L., Y. Sun, and D. Caetano-Anolles, *Function and Evolution of C2H2 Zinc Finger Arrays.* Subcell Biochem, 2011. **52**: p. 75-94.

107. Emerson, R.O. and J.H. Thomas, *Adaptive evolution in zinc finger transcription factors.* PLoS Genet, 2009. **5**(1): p. e1000325.

108. Urrutia, R., *KRAB-containing zinc-finger repressor proteins.* Genome Biol, 2003. **4**(10): p. 231.

109. Thomas, J.H. and S. Schneider, *Coevolution of retroelements and tandem zinc finger genes.* Genome Res, 2011. **21**(11): p. 1800-12.

110. Wolf, G., D. Greenberg, and T.S. Macfarlan, *Spotting the enemy within: Targeted silencing of foreign DNA in mammalian genomes by the Kruppel-associated box zinc finger protein family.* Mob DNA, 2015. **6**: p. 17.

111. Ecco, G., M. Imbeault, and D. Trono, *KRAB zinc finger proteins.* Development, 2017. **144**(15): p. 2719-2729.

112. Kasaai, B., M.H. Gaumond, and P. Moffatt, *Regulation of the bone-restricted IFITM-like (Bril) gene transcription by Sp and Gli family members and CpG methylation.* J Biol Chem, 2013. **288**(19): p. 13278-94.

113. Takahashi, T., et al., *Autoregulatory mechanism of Runx2 through the expression of transcription factors and bone matrix proteins in multipotential mesenchymal cell line, ROB-C26.* J Oral Sci, 2005. **47**(4): p. 199-207.

114. Takahashi, T., et al., *The effect of retinoic acid on a zinc finger transcription factor, AJ18, during differentiation of a rat clonal preosteoblastic cell line, ROB-C20, into osteoblasts.* Arch Oral Biol, 2008. **53**(1): p. 87-94.

115. Sun, Y. and Z. Xu, *Regulation of neural stem cell by bone morphogenetic protein (BMP) signaling during brain development.* Front. Biol., 2010. **5**(5): p. 380-385.

116. Yu, S., et al., *Retinoic acid induces neurogenesis by activating both retinoic acid receptors (RARs) and peroxisome proliferator-activated receptor beta/delta (PPARbeta/delta).* J Biol Chem, 2012. **287**(50): p. 42195-205.

117. Gao, L., et al., *Cloning and characterization of a novel human zinc finger gene, hKid3, from a C2H2-ZNF enriched human embryonic cDNA library.* Biochem Biophys Res Commun, 2004. **325**(4): p. 1145-52.

118. Watson, R.P., N. Tekki-Kessaris, and C.A. Boulter, *Characterisation, chromosomal localisation and expression of the mouse Kid3 gene.* Biochim Biophys Acta, 2000. **1490**(1-2): p. 153-8.

119. Olmsted, J.B., et al., *Isolation of microtubule protein from cultured mouse neuroblastoma cells.* Proc Natl Acad Sci U S A, 1970. **65**(1): p. 129-36.

120. Sato, C., T. Matsuda, and K. Kitajima, *Neuronal differentiation-dependent expression of the disialic acid epitope on CD166 and its involvement in neurite formation in Neuro2A cells.* J Biol Chem, 2002. **277**(47): p. 45299-305.

121. LePage, K.T., et al., *On the use of neuro-2a neuroblastoma cells versus intact neurons in primary culture for neurotoxicity studies.* Crit Rev Neurobiol, 2005. **17**(1): p. 27-50.

122. Rowe, H.M., et al., *KAP1 controls endogenous retroviruses in embryonic stem cells.* Nature, 2010. **463**(7278): p. 237-40.

123. Jakel, S., et al., *The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1.* EMBO J, 1999. **18**(9): p. 2411-23.

124. Knockenhauer, K.E. and T.U. Schwartz, *The Nuclear Pore Complex as a Flexible and Dynamic Gate.* Cell, 2016. **164**(6): p. 1162-1171.

125. Mosammaparast, N. and L.F. Pemberton, *Karyopherins: from nuclear-transport mediators to nuclear-function regulators.* Trends Cell Biol, 2004. **14**(10): p. 547-56.

126. Powers, M.A. and D.J. Forbes, *Cytosolic factors in nuclear transport: what's importin?* Cell, 1994. **79**(6): p. 931-4.

127. Fire, A., et al., *Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans.* Nature, 1998. **391**(6669): p. 806-11.

128. Elbashir, S.M., W. Lendeckel, and T. Tuschl, *RNA interference is mediated by 21- and 22-nucleotide RNAs.* Genes Dev, 2001. **15**(2): p. 188-200.

129. Maniataki, E. and Z. Mourelatos, *A human, ATP-independent, RISC assembly machine fueled by pre-miRNA.* Genes Dev, 2005. **19**(24): p. 2979-90.

130. Kutter, C. and P. Svoboda, *miRNA, siRNA, piRNA: Knowns of the unknown.* RNA Biol, 2008. **5**(4): p. 181-8.

131. Moore, C.B., et al., *Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown.* Methods Mol Biol, 2010. **629**: p. 141-58.

132. Castel, S.E. and R.A. Martienssen, *RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond.* Nat Rev Genet, 2013. **14**(2): p. 100-12.

133. Rowe, H.M. and D. Trono, *Dynamic control of endogenous retroviruses during development.* Virology, 2011. **411**(2): p. 273-87.

134. Wolf, D. and S.P. Goff, *Embryonic stem cells use ZFP809 to silence retroviral DNAs.* Nature, 2009. **458**(7242): p. 1201-4.

135. Wolf, G., et al., *The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses.* Genes Dev, 2015. **29**(5): p. 538-54.

136. Fasching, L., et al., *TRIM28 represses transcription of endogenous retroviruses in neural progenitor cells.* Cell Rep, 2015. **10**(1): p. 20-8.

137. Bao, W., K.K. Kojima, and O. Kohany, *Repbase Update, a database of repetitive elements in eukaryotic genomes.* Mob DNA, 2015. **6**: p. 11.

138. Sripathy, S.P., J. Stevens, and D.C. Schultz, *The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression.* Mol Cell Biol, 2006. **26**(22): p. 8623-38.

5 Discussion

The process of neuronal cell fate acquisition and terminal differentiation is regulated by multiple cell-intrinsic and -extrinsic mechanisms in a spatio-temporal manner. The interplay of TFs and epigenetic regulators that become available at distinct phases of embryonic development is increasingly appreciated as a driving force in this process. The environment-specific inputs are integrated by *cis*-regulatory elements to configure the transcriptional output of a cell during ontogeny [1-4]. These elements are embedded in a chromatin landscape, which is regulated at several layers, for example by chromatin remodeling complexes, histone modifying enzymes, DNA methylation dynamics as well as non-coding RNAs [5]. Sequence-specific nuclear factors, especially pioneer transcription factors [6-8], are involved at the forefront of the acquisition of a permissive chromatin state conferring cell fate alterations. The concerted action of these gene regulatory layers during embryonic neurogenesis controls the transcriptome of distinct cell states as they progress along neuronal fate into mature neurons [9-11]. Major players therein are activators of neuronal and repressors of non-neuronal genomic elements. It is essential to recognise the fine-tuned interplay at this various regulatory levels to gather an understanding of how cell fate determination is achieved. Two nuclear factors and their impact on the chromatin landscape during neuronal cell fate acquisition were characterised in this thesis: the bHLH factor NeuroD1 (activator, see 3) and the KRAB-ZFP Zfp354c (repressor, see 4). Strikingly, these two nuclear factors belong to TF families that were found to exhibit an unusually high gene gain rate in *Mus musculus*, possibly by single gene duplication events during evolution, which might explain the huge functional diversification among the members of these TF families [12-16].

5.1 NeuroD1 reprograms chromatin and transcription factor landscape to induce the neuronal program

Embryonic stem cells are an asset to decipher developmental processes *in vitro* [17-23]. Several key principles of the epigenetic regulatory landscape are derived from studies of ESCs and from their differentiation into cell types of all three germ layers, especially into neural cells [24-29]. The knowledge about the chromatin landscape and TF occupancy in these pluripotent cells laid the foundation for this thesis chapter. The bHLH TF NeuroD1 was chosen in this thesis as an ectopic neuronal differentiation stimulus in order to convert the cell fate of mESCs into neuronal lineage. This protein was originally identified as an activator of *ins* (insulin) expression in the murine pancreas [30] and as a neurogenic differentiation factor that initiated the conversion of ectodermal cells into neurons as well as premature neuronal differentiation of neuronal precursors upon ectopic expression in *Xenopus* [31]. Further research showed that NeuroD1 is important for neurogenesis along with neuronal migration during development [32-37]. However, its underlying functional mechanism and which genomic regulatory elements NeuroD1 targets during neuronal lineage commitment were not fully understood. Given the described lethality upon *NeuroD1* knock out in mice [30, 38], an overexpression approach was chosen in this thesis to study the molecular basis of NeuroD1's function. Furthermore, recent reports observed that NeuroD1 can transdifferentiate glial cell types into neurons *in vivo* [39-41], which initiated the investigation on how this bHLH factor might interact

with the chromatin landscape in order to convey cell fate changes. The mechanistic insights derived from this thesis can help to understand the molecular basis of such transdifferentiation and might improve induction approaches of neuronal fate also *in vivo* in the light of therapeutic avenues with regard to e.g. neurodegenerative diseases [6, 42-45].

5.1.1 NeuroD1 targets *cis*-regulatory elements of critical neuronal genes to induce neuronal cell fate commitment

Experiments for the induction kinetics in A2lox.GFP mESCs followed by FACS revealed that around 95% of transgenic cells robustly express the transgene from the tetracycline inducible gene expression locus (Diploma thesis Johannes Jung [46]). The doxycycline inducible A2lox cells were therefore used in this thesis to ectopically express *NeuroD1* in a reproducible fashion. This system additionally avoided possible deviations in transfection efficiency using plasmids encoding for *NeuroD1*. Strikingly, the ectopic expression of *NeuroD1* in these cells was sufficient to induce their differentiation into neuronal-like cells which exhibit TUJ1-positive cellular processes like neuronal cells *in vivo*. This phenomenon was accompanied by a massive induction of a neurogenesis-associated transcriptome, as NeuroD1-induced genes displayed a significant enrichment for neuronal gene ontologies. Furthermore, these genes were shown to depend on *NeuroD1* expression as they were repressed in cells without *NeuroD1* transcription, notably tissues of non-neuronal lineages as well as during neuronal development. Interestingly, the downregulated genes upon ectopic *NeuroD1* induction were not either physiologically expressed during *in vitro* and *in vivo* neurogenesis. This physiological expression profile of differentially expressed genes in A2lox.NeuroD1 cells upon *NeuroD1* induction strongly supports the hypothesis that this bHLH TF promotes a differentiation program for neuronal development. The neuronal cell fate commitment and differentiation of mESCs induced by ectopic *NeuroD1* is furthermore remarkable because pluripotency stimuli from the mESC culture media were overwritten, evident by the repression of core pluripotency genes like *Oct4*, *Klf4* and *Nanog*. However, it has been shown that mESCs conventionally cultured with serum and LIF appear as heterogeneous subpopulations with fluctuations in the expression levels of important pluripotency-associated TFs like *Nanog, T-box 3* (*Tbx3*) or *estrogen related receptor beta* (*Esrrb*) [47-51]. This conventional pluripotency state might mediate the cellular response to NeuroD1's neurogenic activity in a mESC subpopulation which is in a more permissive cell state for differentiation [52], although almost all cells express the transgene. This might explain the occurrence of TUJ1-negative cells in the immunofluorescence analyses of A2lox.NeuroD1 cells after 48h of *NeuroD1* induction. Interestingly, previous reports showed that bone morphogenetic proteins (BMPs), especially BMP4, are able to block the differentiation of mESCs in combination with LIF through the induction of *Id* genes via Smad transcription factors [53-55]. In contrast, the activation of extracellular signal-regulated kinase (Erk) signaling pathway by the fibroblast growth factor 4 (Fgf4) was found to promote the differentiation of mESCs [56-58]. Remarkably further, recent findings established that the combination of two chemical inhibitors (2i), one selective for the glycogen synthase kinase-3 (GSK3) and one for the MAPK/ERK kinase (MEK) pathway, enables the culture of mESCs which resemble the *in vivo* ground state (naïve pluripotency) in the ICM more closely and further robustly self-

renew as a homogeneous cell population [59-65]. Therefore, it would be interesting to investigate whether the efficiency of NeuroD1-induced mESC differentiation into neuronal cells could be altered by employing different media compositions in order to change the pluripotency state of A2lox.NeuroD1 mESCs prior ectopic *NeuroD1* expression as well as external differentiation stimuli.

A direct relationship of *NeuroD1* overexpression and the initiation of a neuronal developmental program was evident by the occupancy of ectopic NeuroD1 at *cis*-regulatory elements, both at enhancers and promoters, associated to a number of the upregulated neuronal genes. Strikingly, a significant fraction of the here *in vitro* identified enhancers are known to be cortex- or brain-specific in *Mus musculus* [66]. This finding strengthens the hypothesis that NeuroD1 plays an important role during brain development by enhancing gene expression. This finding was additionally corroborated by the sequence-specific nature of NeuroD1 binding as its target elements harboured the established E-box motif for bHLH factors [67] directly at the peak summit. Although several variations of E-box motifs exist, one of the top three detected motifs in this thesis represents the exact sequence as published for the NeuroD1/E47 hetero-dimer [68]. The bHLH factors are known to act in a concerted manner with co-factors to integrate the spatio-temporal context, especially at enhancers [69, 70]. It would be interesting to further investigate whether NeuroD1 indeed occupies its targets in A2lox.NeuroD1 mESCs on the one hand as a hetero-dimer with the E47 [68] or with other dimerisation partners that might influence the preference for DNA binding sites [71, 72]. On the other hand, NeuroD1's possible cooperation with other nuclear factors to induce the neurogenic program would be of interest, like it has been shown for the bHLH TF *Ascl1* [69, 73]. This is fascinating in the light of neuronal subtype specification which has been linked to the cooperativity of bHLH TFs during embryonic development [74-78]. It would be furthermore interesting to investigate whether differential co-partners of NeuroD1 influence the repertoire of bound versus non-bound NeuroD1 motifs in the genome. By using a refined ChIP-seq methodology like ChIP-exo [79] it would be possible to exactly determine the genomic location of NeuroD1 occupancy for example at proximal or distal promoter elements. This has experimentally not been explored in this thesis as a promoter occupancy of NeuroD1 was generally referred to as occurring between -800 bp and +200 bp from the TSS.

Interestingly, a significant fraction of NeuroD1's direct targets encode for TFs as well as epigenetic regulators, uncovering that NeuroD1 itself activates gene regulatory factors and therefore governs the neurogenic transcription program on various levels. This is in line with a recent report that similarly showed an induced set of transcriptional regulators, including *NeuroD1*, as direct targets of the proneural bHLH TFs *NGN1* and *NGN2,* whose simultaneous overexpression in human iPS cells prompted neuronal differentiation [80]. Apart from the global analysis of RNA-seq data in this thesis, several of the NeuroD1-induced transcriptional regulators have also been validated in their neuronal-specific expression pattern by RT-qPCR measurements in embryonic tissues as well as with publicly available ISH images of the Allan Brain Atlas (brain-map.org/api/index.html; [81]). Several of the identified nuclear factors like *Aff3*, *Rnf182* or *Zfand5* have not been deeply investigated in the context of embryonic neurogenesis and warrant further analysis. Nonetheless, *Aff3* and *Rnf182* have been implicated in the context of brain malfunctions like intellectual disability or Alzheimer's disease [82-85], supporting the functional relevance of NeuroD1's

targets. Interestingly, a significant part of NeuroD1-induced genes is also upregulated in cells during developmental epithelial-to-mesenchymal transition [86], a process that is closely related to neuronal migration [87, 88]. These upregulated genes encompass several TFs that are established EMT markers for mesenchymal fate such as *Snai1*, *Twist1* or *Zeb2* [89-91]. Importantly, another hallmark regulator which represses E-cadherin and induces EMT, *Snai2* [92-95], is directly bound by NeuroD1 at its promoter. Interestingly, the invasiveness of carcinoma cells has been attributed in part to the concerted action of bHLH, Snai and Zeb factors [96] and NeuroD1 has recently been implicated in cell motility regulation of neuroendocrine cancers [97-99]. Taken together, these data suggest that NeuroD1 regulates the migration of newborn neurons by directly activating the underlying transcriptional program.

Previous reports showed that mESCs could be committed to a glutamatergic neuronal cell fate by overexpression of *Ngn1* or *Ngn2* [100, 101], which are transcriptional activators of *NeuroD1* [102, 103]. Velkey & O'Shea overexpressed *Ngn1* in mESCs of a similar (trans-)genetic background as A2lox mESCs [101]. The authors proposed that the neuronal differentiation of mESCs was majorly induced by Ngn1 through the activation of *NeuroD1* apart from six additional key neuronal TFs, including *Hes6*, *Nhlh1* and *NeuroD4* [101, 104]. Strikingly, the expression of *NeuroD1*, *Nhlh1* and *Hes6* is repressed in mice with *Ngn1* or *Ngn2* mutations [105-107]. This thesis presents that *Hes6*, *Nhlh1* and *NeuroD4* are directly regulated by NeuroD1, proposing the neuronal cell fate commitment and differentiation of mESCs induced by ectopic expression of neurogenins might therefore be caused by NeuroD1. Interestingly, *Hes6* is expressed in immature neurons but not induced via Notch signaling [107]. Its function to promote neuronal differentiation in a positive feedback-loop with neurogenins supports its exclusive role among Hes family members as for example Hes1 and Hes5 promote NSC self-renewal and astrocyte differentiation [107-110]. Indeed, although Hes6 possesses a DNA-binding ability, it was shown to display its neurogenic activity by directly inhibiting the Notch signaling effector Hes1 due to the formation of a non-functional heterodimer and supporting its proteolytic degradation [110-113]. Hes6 supports thereby the switch from self-renewal to neuronal differentiation of NSCs. The findings of this thesis strengthen the role of Hes6 as an essential downstream effector of NeuroD1's neurogenic ability [114]. The bHLH TF *Nhlh1* has been reported as a NeuroD1 target and shown to function in the context of neurogenesis [115-117]. Its neuronal expression pattern in the embryonic cortex mimics that of *NeuroD1* [118], which is in line with the observations made in this thesis. However, *Nhlh1* mutant mice are viable and seem not to exhibit neurological malfunctions [119]. This data suggest that *Nhlh1* is not an indispensable regulator of embryonic neurogenesis but might play an important role as bHLH dimerisation partner during neuronal differentiation. Another NeuroD1-induced target gene shown in this thesis is *Dll3.* This gene was first described to function during murine embryonic development by promoting neurogenesis in the spinal cord following the transcriptional reduction of the Notch ligand *Dll1* [120]. Also in the developing cortex, the sequential expression of *Dll1* followed by *Dll3* has been observed as RG cells in the VZ become committed towards neuronal fate and their progeny migrates towards the CP [121, 122]. Interestingly, Dll3 does not act in *trans* by binding to Notch receptors exposed on the cell membrane of other cells [123]. Ladi and colleagues instead proposed that Dll3 acts in *cis* by blocking Notch receptors on co-expressing cells and hence promotes their differentiation along

neuronal fate in addition to preventing astrocytic differentiation [123]. Taken together these data support the hypothesis that a major part of the downstream events initiated by the proneural bHLH factors *Ngn1* or *Ngn2* might be mediated by the activity of NeuroD1 itself, which is in line with previous reports [124, 125]. Interestingly, the induction of ectopic *NeuroD1* in A2lox.NeuroD1 mESCs also increased the transcription of the endogenous *NeuroD1* gene, suggesting a possible positive feedback loop. However, despite constant supply of doxycycline, the protein level of ectopic NeuroD1 decreased from 24 h to 48 h of Dox induction in A2lox.NeuroD1 mESCs, preceded by a drop of its RNA levels. This observation suggests that there are cell intrinsic mechanisms that counteract the differentiation stimulus of NeuroD1 once stem cell fate has progressed along neuronal lineage, possibly mediated by the action of other bHLH proteins [69, 70, 126]. Interestingly, studies with *NeuroD1* knock out mice rescued with *NeuroD1* expression in the pancreas proposed that there might be a functional redundancy of NeuroD1's function in the neocortex with other nuclear factors. The loss of neurons in these mice was majorly restricted to the cerebellum, hippocampus or sensory regions like the retina and inner ear, causing behavioural deficits [124, 127-129]. This is strengthened by the observation that humans with *NeuroD1* mutations possess a neocortex, although they exhibited severe neuronal deficits and only two individuals are known who survived [130]. These individuals might have experienced a compensation of the normally lethal pancreatic malfunctions encountered shortly after embryonic development. These observations suggest that several regulatory mechanisms have evolved that facilitate the coordinated formation of the brain in order to prevent detrimental effects of a loss-of-function mutations in a single gene. It is not clear which factors might compensate NeuroD1's function during embryonic neurogenesis and future research should be focused in this direction.

Nonetheless, the findings presented in this thesis support a prominent role for NeuroD1 as a regulator of neuronal development by promoting neuronal migration and differentiation due to directly occupying *cis*-regulatory elements of key neuronal genes to initiate a neuronal developmental program. However, the NeuroD1-induced neuronal cells have neither been further characterised for their electro-physiological properties nor for their ability to form functional synapses. Nevertheless, several marker genes for excitatory glutamatergic neurons but not GABAergic interneurons were significantly upregulated after 48 h of Dox induction in A2lox.NeuroD1 mESCs (see 8.6). These findings suggest that NeuroD1 is important for the cell fate of neuronal cells born in the dorsal telencephalon, which is in line with its expression pattern in the embryonic cortex [37]. Further investigations should be focused on unravelling the detailed neuronal identity of NeuroD1-induced cells and moreover explore its interdependency with neural signaling pathways like Notch-Hes signaling in the context of neuronal fate acquisition.

5.1.2 NeuroD1 initiates chromatin and transcription factor landscape reprogramming to promote neurogenesis

The direct target sites of NeuroD1, both promoters and enhancers, exhibited a different chromatin and nuclear factor landscape in mESCs in comparison to non-target sites linked to NeuroD1-induced genes. The identified target regulatory regions were silenced by distinct combinations of histone PTMs and regulatory factors. The ectopic induction of NeuroD1 led to its occupancy at these target sites which initiated

a remodeling of the chromatin and nuclear factor landscapes towards a euchromatin state, mediating transcriptional activation of associated genes.

At promoters, NeuroD1 occupancy initiated a decrease in the repressive polycomb mark H3K27me3 which was paralleled by an increase in the activating modification H3K27ac and chromatin accessibility at the same sites. The loss of H3K27me3 during (ectopic) neuronal differentiation of mESCs suggests that these promoters might have been poised in mESCs [131-133], but the levels of H3K4me3 enrichment were not experimentally investigated to support this hypothesis. A time course analysis of the chromatin dynamics at target promoters for *NeuroD1*, *Hes6* and *Dll3* showed a time dependent recruitment of RNA-Polymerase II and hence transcriptional activation of the associated genes. Similarly, NeuroD1 occupancy at enhancers increased their chromatin accessibility as well as H3K27ac enrichment, both hallmarks of active enhancers [134-138]. Nevertheless, acetylation at other histone residues apart from lysine 27 of histone H3 is increasingly appreciated to support the identification of active enhancers [139] and could be considered for further investigations. The Bayesian model predicted higher H3K4me1 levels at NeuroD1 enhancer targets in comparison to control regions associated with NeuroD1-induced genes. These data in combination with their less accessible chromatin suggest that the probed enhancers might be poised in mESCs [134, 135]. However, the H3K27me3 dynamics at these *cis*-regulatory elements should be investigated in order to strengthen this hypothesis in addition to their DNA-methylation cycle [134-136, 140]. Strikingly, a time course expression analysis following NeuroD1 induction showed that the genes *NeuroD4*, *Nhlhl1* and *Ncam1*, which are associated to target enhancers, were transcriptionally activated much later in comparison to the probed promoter targets. Although the H3K27ac increase at these enhancers mimicked the temporal dynamics upon NeuroD1 occupancy as observed at promoters, these data suggest that a differential regulatory cascade might operate at these two groups of NeuroD1-targeted *cis*-regulatory elements. It would be interesting to further investigate whether promoter-enhancer or enhancer-gene body interactions differ between these two entities of NeuroD1-induced targets and to which extend this might be caused by distinct combinations of harboured TF motifs [3, 141-145]. It is unclear whether the regulation of NeuroD1-target elements might be influenced on a larger scale by differential nuclear architecture or higher order chromatin structures in the nuclear environment. It has to be noted that the HA-NeuroD1 ChIP-seq in A2lox.NeuroD1 cells was performed after 24 h of induction but the time-resolved binding analysis showed NeuroD1 binding already after 6 h. Although the levels of NeuroD1 occupancy increased at target *cis*-regulatory elements after 6 h, the experimental time point of 24 h already reflected a phase of decreasing NeuroD1 occupancy. Several studies have shown that the residence times of TFs at their target sites is dynamic [146-148]. In the context of the global transcriptome profiling of a steady-state after 48 h of NeuroD1 induction in A2lox.NeuroD1 cells, it is possible that a potential population of early target genes was not identified in this thesis. Additionally, potential target regions of endogenous NeuroD1 are not covered in the computational analyses as NeuroD1 ChIP-seq was performed for the ectopically induced HA-tagged protein.

It has recently been shown during neuronal differentiation of human ESCs that stage-specific TFs prepare the chromatin landscape at *cis*-regulatory elements for other nuclear factors to act during cell fate commitment [149]. In line with these observations it was exciting to find that NeuroD1 targeted *cis*-regulatory

elements embedded in a silent chromatin environment and initiated the conversion of their chromatin landscape towards a more accessible and active state. This findings suggest that NeuroD1 is a potential pioneer transcription factor [6-8] that conveys transcriptional competence during neuronal cell fate acquisition. This hypothesis is strengthened by the observation that NeuroD1 was able to target the identified *cis*-regulatory regions also upon overexpression in murine fibroblast and that associated genes were subsequently transcriptionally activated. This sequence-specific targeting of NeuroD1 was also observed in *in vitro* derived early neuronal cells (TN d1) as well as directly in the murine embryonic cortex. These observations furthermore support that the *in vitro* identified targets upon ectopic expression of HA-tagged NeuroD1 were authentic *in vivo* targets, as all of them showed NeuroD1 enrichment in the embryonic cortex in comparison to control regions. Interestingly, the *in vivo* overexpression of *NeuroD1* in RG cells promoted their neurogenic differentiation emphasised by the findings that almost no *NeuroD1*-positive cells were found in the VZ but an increased population resided in the IZ in comparison to control condition. This "depletion" of the progenitor pool overexpressing *NeuroD1* is possibly mediated by altered Notch signaling in these progenitors due to the action of NeuroD1 target genes such as *Hes6* and *Dll3*. These observations suggest that NeuroD1-positive cells prematurely migrated from the VZ into upper cortical layers. This is in line with previous reports showing that NeuroD1 is important for neuronal migration [32, 34] and the observations in this thesis that it directly regulates critical EMT genes. However, in developmental context neither neuronal migration directly [150] nor underlying cell biological dynamics [151, 152] were investigated in this thesis in order to strengthen this hypothesis. Nonetheless, the *in vivo* population of NeuroD1-positive cells displayed an upregulated expression of NeuroD1-regulated genes similar as observed during *in vitro* neurogenesis. Remarkably, also H3K27ac levels at their identified NeuroD1-target regulatory elements increased, suggesting that NeuroD1 initiated the euchromatin conversion of *cis*-regulatory elements associated with critical neuronal genes during *in vivo* embryonic neurogenesis. This observation is among the first to show the impact of increased levels of a neurogenic TF on the chromatin landscape during *in vivo* cortical development that consecutively determines the cell fate of neural stem cells towards neuronal lineage. Taken together, these observations implicate that NeuroD1 might be a PTF for neuronal development in line with its described ability to induce neuronal transdifferentiation [39-41]. However, further experimentation is needed to fully understand NeuroD1's function. The *in vivo* investigation should be extended as the gene delivery by *in utero* electroporation (IUE) is a transient approach and IUE efficiencies exhibit experimental variations. These characteristics of IUE might explain the appearance of Pax6-positive RG cells in the VZ of NeuroD1-IRES-RFP electroporated brains because not all NSCs might have expressed the ectopic CDS. Performing *in vivo* IUE rescue experiments by combining an endogenous *NeuroD1* knock down in parallel to ectopic *NeuroD1* overexpression could shed light to the relevance of the proposed positive-feedback loop in relation to the overall expression levels of this bHLH factor. Although the *in vivo* phenotype was reproducible, employing refined *in vivo* experimental approaches [153-155] and an increase of biological replicates especially for the immunohistochemical analyses would strengthen the hypothesis of NeuroD1's function in committing multipotent NSCs towards neuronal lineage.

Interestingly, a recent study in *Caenorhabditis elegans* proposed the recruitment of Pol II to *cis*-regulatory elements as a novel role for PTFs [156]. This would be in line with the observed Pol II kinetics at NeuroD1-target promoters. Since this concept was derived from an invertebrate, it should be further investigated especially in the context of the complex mammalian enhancer landscape. Additionally, one validation of NeuroD1's proposed activity as a PTF should be further mechanistically investigated: the ability of PTFs to directly occupy DNA target sites on *in vitro* nucleosomes sequence-specifically [6, 157]. This approach would improve the understanding of how PTFs recognise their target elements on nucleosomal DNA [158-160] as the bHLH DBD domain seems not to appear completely folded in solution apart from its DNA association [161]. In line with this, a recent study showed that the bHLH factor c-Myc can recognize a partial E-box motif and adapts its DBD domain to the nucleosomal surface in cooperation with other nuclear factors [159]. It should be further investigated if this mechanism holds true for NeuroD1 as well. Interestingly, NeuroD1 was not able to confer transcriptional competence to *NeuroD4* in murine fibroblasts. This gene was transcriptionally silent in wild type fibroblast, which suggests that it might possibly be repressed by H3K9me3 containing heterochromatin domains in concert with DNA methylation and therefore be embedded in a repellent chromatin landscape leaving NeuroD1 unable to bind as a pioneering factor [6, 162-164]. If NeuroD1's proposed action as a PTF does impact on DNA methylation turnover at *cis*-regulatory elements [165] is still an open question.

The mechanisms by which NeuroD1 convers the downstream chromatin remodeling events upon its occupancy of *cis*-regulatory elements are not fully unravelled. A recent report engaged a CRISPR/Cas9-based methodology [166-173] to demonstrate that an ectopic increase of H3K27ac levels at *cis*-regulatory elements mediated by targeting the HAT p300 to these sequences is sufficient to induce the expression of associated genes [174]. This line of observations is in agreement with the proposed function of NeuroD1 to activate *cis*-regulatory elements important for neuronal development as its recruitment triggered an increase of H3K27ac levels as well as transcriptional induction of associated genes. Strikingly, NeuroD1-target *cis*-regulatory elements and their associated genes exhibited similar dynamics with regard to chromatin landscape and transcriptional profile during physiological *in vitro* and *in vivo* neurogenesis following endogenous *NeuroD1* induction. It was not investigated in this thesis if NeuroD1 directly recruits H3K27-specific HATs to its target sites during neuronal development. However, previous reports proposed that NeuroD1 directly interacts with p300 [175-178]. Although these observation were derived in a different cellular context apart from neurogenesis, the interaction of p300-NeuroD1 would explain the observed H3K27ac dynamics following NeuroD1 occupancy at the tested genomic elements. This hypothesis could be validated by co-IP or IP-MS experiments in the applied cellular systems as well as during *in vivo* neurogenesis. A recent study [179] altered *in vivo* neurogenesis by using CRISPR/Cas9 technology to specifically target the HMT Ezh2 to the *Tbr2* locus, which led to ectopically increased H3K27me3 levels and hence transcriptional repression of this important regulator of indirect neurogenesis [180, 181]. Strikingly, a similar phenotype was previously observed with a catalytically active Cas9 enzyme targeting the *Tbr2* locus [182]. These observations show that the CRISPR/Cas9 technology is a powerful tool for genome- as well as epigenome editing [183], especially during *in vivo* development. Such observations provoke the question

whether the neuronal cell fate induction in RG cells could be recapitulated by specifically activating NeuroD1 target genes *in vivo* or in the employed mESC differentiation system *in vitro* using CRISPR/Cas9. On the other hand this technology would allow a screening for the functional importance of specific NeuroD1 target genes during neuronal cell fate acquisition by combining H3K27me3-mediated silencing with ectopic NeuroD1 overexpression *in vivo*. In the context of data presented in this thesis such approaches could further be applied to the identified distal regulatory elements that were shown to regulate essential neuronal genes. Furthermore, the manipulation of endogenous *NeuroD1* levels during *in vivo* development as well as in combination with its ectopic induction would enhance the understanding of the proposed NeuroD1 positive feedback loop during neuronal differentiation of NSCs. A functional interplay of NeuroD1 with erasers of histone PTMs has been indicated for the H3K9me2-demethylation activity of LSD1 in order to promote transcriptional induction in a different cellular background [184], but further investigation is needed to elucidate a possible link to neurogenesis.

It is well accepted that the accessibility of *cis*-regulatory elements to nuclear factors and the transcriptional machinery is a critical layer of gene regulation [5, 10, 185-187]. The role of (lineage-specific) TFs, especially PTFs, in regulating these DNA motility events is increasingly appreciated [8, 188-190]. One essential chromatin remodeling complex (CRC) for neuronal development is the BAF complex [191], which might be responsible for the increase of chromatin accessibility at NeuroD1-target *cis*-regulatory elements as measured by FAIRE-qPCR in this thesis. This hypothesis is supported by a study showing that the neurogenic activity of NeuroD1 was abolished in a *Brg1*-depleted embryonic cell line (P19) [124]. Preliminary data for NeuroD1 IP-MS experiments in A2lox.NeuroD1 mESCs showed an interaction of NeuroD1 with several components of the esBAF complex that are also present in npBAF and nBAF, including Brg1 (data not shown). This data suggests that npBAF/nBAF CRCs might confer chromatin accessibility changes by NeuroD1-dependend recruitment to *cis*-regulatory elements, but further validation is needed. A recent study published single nucleus ATAC-seq data of chromatin accessibility dynamics in the developing murine brain [192]. Interestingly, the authors found the NeuroD1 motif specifically enriched in open chromatin areas associated with excitatory neurons, suggesting a prominent role for NeuroD1 in regulating chromatin accessibility during embryonic neurogenesis in the dorsal telencephalon as proposed in this thesis. These time-resolved chromatin accessibility data should be investigated in more detail with regard to the developmental chromatin compaction dynamics at the here identified NeuroD1 target elements.

This thesis presents that NeuroD1 not only modulates histone PTM patterns but also sets changes to the landscape of nuclear proteins at its genomic target loci. It was shown that the occupancy of NeuroD1 at target promoters displaced for example Tbx3 [193] in order to activate gene expression. Tbx3 has been implicated in promoting the pluripotent state of ESCs by mediating for example the downstream effects of LIF-signaling [51, 194, 195]. However, its importance for the pluripotency gene regulatory network is under debate as it might be dispensable for mESC maintenance [196, 197]. On the other hand, Tbx3 has also been demonstrated to promote cell fate acquisition processes towards mesendoderm lineage, especially during ESC differentiation [198-202]. These data suggest that Tbx3 might represses ectodermal

differentiation of mESCs [201, 202] and that its removal at critical neuronal genes initiated by NeuroD1 supports transcriptional competence for neuronal differentiation. However, Tbx3 has apart from its repressive function also been described as a transcriptional activator [197, 203-206]. Interestingly, a recent report showed that Tbx3 associates with Jmjd3, a H3K27me3 demethylases [207], to promote endoderm differentiation for example by *Tbr2* activation [208]. These findings suggest an activating function for Tbx3 during endodermal lineage specification also on an epigenetic level, which might be linked to a specific cellular context [209]. The NeuroD1-mediated antagonism of Tbx3 at promoters of critical neuronal genes should therefore be further investigated in more mechanistic details, especially in the light of a newly proposed ability of TBX3 in promoting human ESC differentiation into neuroepithelial cells [210].

In addition to Tbx3, another nuclear factor, Mbd3 [211], was specifically enriched at NeuroD1-target enhancers prior its induction in mESCs. Mbd3 was identified as a subunit of the nucleosome remodeling deacetylase (NuRD) complex regulating chromatin accessibility due to its CHD3/4 and HDAC subunits [212-217]. The Mbd3-NuRD complex might keep the identified NeuroD1-target enhancers in a hypoacetylated, closed chromatin landscape. After its dissociation initiated by NeuroD1-binding the landscape is converted into euchromatin, as discussed above. However, distinct NuRD complexes exist concerning their subunit composition and functional specificity [212, 218]. Mbd3 was reported as an important epigenetic regulator contributing to the differentiation ability of mESCs as well as to the success of somatic cell reprogramming into induced pluripotent states in a context dependent manner [219-222]. Interestingly, the Mbd3 subunit might only to some extent be responsible for targeting NuRD to methylated DNA loci in order to trigger heterochromatin formation as it is unable to bind 5mC but might exhibit an affinity to 5-hmC [223, 224]. These observations currently lead to controversies about Mbd3's functional occupancy at genomic loci, which also comprise non-methylated regions as well as lowly methylated CpG islands [225-228]. Additionally, a recent study observed that the emergence of bivalent chromatin in mESCs is partially facilitated by the NuRD-depended recruitment of PRC2 at promoters of developmental genes [229]. These data suggest that NuRD activity prepares the chromatin landscape of pluripotent cells for developmental cell fate commitment. This is further supported by the observations that the NuRD complex modulates the transcriptional competence of *cis*-regulatory elements associated to pluripotency genes to enable their rapid silencing as mESCs receive adequate differentiation cues [229, 230]. These findings further advocate that neurogenic ability of NeuroD1 demonstrated in this thesis might have been enhanced by the action of NuRD. As these observations were drawn from mESCs in a conventional, heterogeneous pluripotency state it would be interesting to recapitulate these findings in other culture conditions as proposed above. Nonetheless, these findings exemplify the possibility of a fine-tuned functional interplay of CRCs (NuRD) and a PTF (NeuroD1) during lineage commitment [189, 229, 231, 232], which should be further explored in the context of neuronal cell fate acquisition as well as cyclic DNA methylation dynamics.

5.1.3 NeuroD1 induces an epigenetic memory for neuronal lineage during cell fate commitment

All above discussed *in vitro* observations in mESCs were derived in culture conditions that promote pluripotency [49, 61]. However, several reports showed that NeuroD1 is capable to prompt neuronal differentiation of somatic cells by initiating their (trans-)differentiation into neurons [31, 39-41]. The demonstrated NeuroD1-dependend conversion of glial cell types into neurons directly *in vivo* [39-41] initiated the investigation in this thesis on NeuroD1's ability to induce neuronal cell fate in a different cellular milieu. The above discussed *in vitro* findings of this study should be reproducible in culture conditions that favour neuronal differentiation, which furthermore might even enhance NeuroD1's function by additionally providing neurogenic extracellular stimuli. The here developed iTN (induced terminal differentiated neurons) procedure was derived from established culture conditions of an *in vitro* mESC differentiation protocol published by Bibel and colleagues [233, 234]. In contrast, iTN differentiation omits the cellular aggregate formation and instead uses NeuroD1-induction for direct cell fate commitment of mESCs into neuronal-like cells. It was apparent that the culture environment promoted neuronal differentiation *per se* as the N2 media in combination with the provided ECM on the culture dishes affected the expression profile of A2lox.NeuroD1 mESCs as such*:* for example, the expression of the neuronal marker *Tubb3* increased in the control condition from ESC state to iTN day one without augmented *NeuroD1* levels and only dropped after three days. Furthermore, the expression of *Sox2* was also induced in non-induced cells (-Dox), suggesting that the neurogenic environment directly promoted its transcriptional upregulation. This increase of *Sox2* was in contrast to the unaltered expression levels in ESC culture conditions and the functional implication of this induction remains to be investigated. However, apart from its established role in the pluripotency network in ESCs [49, 61], several studies implicated an important role for *Sox2* in the regulation of neural progenitor identity versus their cell fate commitment towards neuronal lineage [235-240]. Furthermore, Sox2 has been used as a single factor to transdifferentiate fibroblast into NSCs, which supports its described function as PTF [158, 159]. Interestingly, a recent study proposed that SOX2 promotes the bivalent chromatin landscape at important neuronal genes, including *Ngn2* and *NeuroD1*, for their activation during neurogenesis by preventing PRC2 function [241]. However, this observation was derived during adult neurogenesis in the hippocampus and it should be further investigated if Sox2 exhibits a similar function in the developmental context studied in this thesis. It has to be noted that especially in the context of the proposed interaction of Sox2 with lncRNAs [242] and a negative-feedback loop with NeuroD1 [243], the regulatory role of Sox2 might be complex. Nonetheless, the removal of pluripotency stimuli in parallel to providing a neurogenic environment enhanced the effects of NeuroD1 as proposed. The immunocytochemistry analysis revealed that almost all iTN cells exhibited a TUJ1-positive neuronal-like morphology in contrast to induced cells in LIF- and serum containing medium (pluripotency condition). Additionally, it was observed that reduced RNA levels of hallmark pluripotency genes were achieved earlier in NeuroD1-induced iTN cells in comparison to the non-induced control condition, especially prominent for *Nanog*. However, it has to be noted that the iTN culture in neuronal media was extended over a longer time period in comparison to the pluripotency condition. Furthermore, the used iTN conditions led to a low cell

density, possibly as a result of cells acquiring a post-mitotic state during neuronal cell fate acquisition whereas non-induced cells continued to proliferate. However, the previously identified NeuroD1 target *cis*-regulatory elements were also occupied during iTN formation. Furthermore, the NeuroD1-induction led to similar euchromatin conversion of target loci and hence the transcriptional induction of associated neuronal genes. A direct comparison of the total levels might not be ideal as the iTN profiling was performed after seven days in contrast to two days in the pluripotency conditions. However, the induction of neuronal fate is evident from the presented data. A cell density kinetics as well as a time-resolved chromatin landscape profiling should be performed in the future to revisit the demonstrated findings during iTN formation.

Nonetheless, time-resolved expression kinetics for ectopic *NeuroD1* as well as four neuronal marker genes (*Tubb3*, *MapT*, *vGlut2* and *Syp*) suggested that the induction of neuronal target genes depends on *NeuroD1* levels. Whereas doxycycline concentrations above 500 ng/ml led to a strong time-depended induction of all tested genes, the lowest tested Dox concentration (50 ng/ml) only minimally induced *NeuroD1* and neuronal genes. This observations suggest that a certain threshold of *NeuroD1* expression has to be met in order to enable this bHLH factor to induce neuronal differentiation. Interestingly, the expression of bHLH factors of the *Hes* family, and in particular *Hes1*, has been demonstrated to oscillate in many cell types, especially in neural progenitors [110, 244-247]. A recent study identified several other bHLH factors, including *Ascl1*, whose oscillatory expression in NSCs is critical for their maintenance in a multipotent, self-renewing state [248]. The authors showed that a sustained expression of *Ascl1* in NSCs initiated their acquisition of neuronal fate and that *Ascl1* levels decreased in a subpopulation of immature neuronal cells after they express early neuronal markers [248]. These findings suggest that the requirements for cell fate acquisition might be met once a critical fate-determining bHLH factor has reached a certain expression level; after progression into the determined fate its sustained expression is not needed anymore. In the light of these findings it is plausible that a similar causality might underly the observed neurogenic ability of *NeuroD1* upon ectopic induction with the higher doxycycline concentrations, which might be enhanced by increased endogenous *NeuroD1* levels. Some studies implicated that *NeuroD1* levels might be important for its function [249] and that its expression is oscillating [250, 251], but these data are derived from a different developmental context and species. Overall, the exact molecular mechanism how the fine-tuned interplay of several oscillating bHLH factors determines NSC identity versus fate-commitment [69, 244, 252] has to be further elucidated, also in concert with other nuclear factors. It should be taken into account that not only the expression dynamics of bHLH factors is regulated. Several studies observed a post-transcriptional regulation for bHLH factors at the protein level, either by PTMs [250, 253, 254] or protein degradation [245, 255, 256], as well as at the level of RNA regulation [257-262]. These studies also suggest that profiling cells in a steady state might overlook the potential dynamics in underlying gene regulatory networks and epigenetic layers on a smaller time scale. Recent developments of NGS approaches in single cells [263-266] might be an additional asset to decipher these regulatory layers without potentially diluting observed effects due to conventional profiling of (heterogeneous) cell populations. The *NeuroD1* induction kinetics for iTN formation should be refined with doxycycline concentrations between 50 ng/ml and 500 ng/ml in order to precisely define the proposed *NeuroD1* expression threshold for neuronal differentiation. These

experiments should be complemented with an immunocytochemistry analysis for TUJ1 to monitor the occurrence of a neuron-like morphology.

The detected physiological expression of *NeuroD1* during both, *in vitro* and *in vivo* neurogenesis, was transiently induced. These findings presented in this thesis are in line with previous reports [31, 267, 268]. Furthermore, many of the above identified targets are transcribed and marked with H3K27ac at developmental time points where NeuroD1 expression has already declined. These data suggest that NeuroD1 is required for a specific developmental period and possibly orchestrates neuronal differentiation by engraining an epigenetic memory for the neuronal transcriptome on the chromatin landscape. The employed A2lox.NeuroD1 mESCs were ideal to mimic expression dynamics due to the inducibility of ectopic *NeuroD1* and the reversible nature of this ectopic induction: the exclusion of doxycycline from the culture medium after two days of iTN formation led to a rapid decrease of ectopic *NeuroD1* levels within 24 h. Strikingly, this transient induction of NeuroD1 was sufficient to confer the euchromatin conversion of target *cis*-regulatory elements that persisted until the examined time point of iTN day seven. The repressive H3K27me3 mark as well as Tbx3 and Mbd3 occupancy were diminished, in line with the previous findings in iTNs using sustained *NeuroD1* induction for seven days. Although the increase of H3K27ac and chromatin accessibility were not as prominent as during conventional iTN formation, the expression of associated genes was still significantly induced. These data suggest that the transient cell fate-determining action of *NeuroD1* is stored in an epigenetic landscape of neuronal fate as a long-term acquired phenotype. In combination with the sequence-specific nature of NeuroD1's occupancy at target regulatory elements of neuronal genes, these findings support the concept that genetic determinants may guide the dynamics of epigenetic landscape acquisition to ensure a stable transcriptome of acquired cell fate during neuronal development. The interplay between the chromatin landscape and fate-determining TFs during cell fate acquisition is well established [2, 269] and the concept of an epigenetic memory for the acquired cellular phenotype increasingly appreciated [137, 269-271]. Additionally, also DNA methylation has been shown to contribute to the formation of an epigenetic memory [272-275], but this aspect was not investigated in this thesis. Interestingly, it has been proposed that the acquisition of an epigenetic cell fate memory already originates during very early murine development separating the trophectoderm from the derivatives of the inner cell mass [276]. These findings suggest that NeuroD1 promotes the progression of the ectodermal epigenetic landscape into a neuronal silhouette. Taken together, the observations presented here suggest that NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. It has to be noted that the term "epigenetic memory" used in this thesis does not refer to heritable changes in gene expression across generations (transgenerational memory, [277]). The term is rather used to describe the chromatin landscape shape that is acquired by a concerted action of multiple layers of epigenetic regulators and TFs during the developmental journey of a cell towards its terminal differentiated neuronal state (cellular memory, [277]).

The presented findings in this thesis describe the molecular events underlying the acquisition of a neuronal fate from a pluripotent cellular origin. It should be investigated if these mechanism are also underlying direct reprograming approaches of somatic cells into neuronal lineage [9, 278-280]. Especially the observations

of NeuroD1-binding and target gene induction in murine fibroblast presented here as well as previous *in vivo* reprogramming reports [39-41] indicate a similar functional ability of NeuroD1 in differentiated cells. A transdifferentiation ability has also been described for other bHLH factors. Vierbuchen and colleagues demonstrated that overexpression of the bHLH factor *Ascl1* in combination with *Brn2* and *Myt1l* (BAM factors) can initiate the transdifferentiation of murine fibroblast into neurons [281]. Furthermore, these MEF-derived neurons were able to integrate into co-cultures of neonatal cortical neurons and formed functional synapses *in vitro*. The iTN system would be ideal to test whether mESC-derived neuronal cells exhibit the same properties to further validate the functional aspects of the NeuroD1-induced cell fate shift into neuronal lineage. This is also interesting in the light of a recent report showing that embryonic neurons could successfully integrate into adult neuronal networks in the murine visual cortex [282]. A follow up study on the above mentioned report used *NeuroD1* in combination with the BAM factors to increase the reprogramming efficiency in human fibroblasts [104]. Strikingly, the authors also observed that a transient induction of the used ectopic factors was sufficient to fully convert fibroblast into induced neurons, which is in agreement with the findings presented in this thesis. It would be interesting to explore if the transient expression of other single neuronal fate determining factors could resemble the observed findings in this thesis. An exciting candidate would be *Ascl1*, which has been successfully used as a single factor in direct neuronal lineage conversions [278, 283, 284] and has already been indicated as a PTF [6, 285]. Additionally, a recent study investigated the transcriptional events upon *Ascl1*-induced fibroblast-to-neuron conversion on a single-cell time course level [286]. This study showed that silencing of *Ascl1* during the early reprogramming phase led to failed lineage conversion. This furthermore supports the hypothesis that a critical expression threshold of a fate-determining factor has to be reached over a certain developmental period in order to acquire neuronal cell fate. The transient necessity of *NeuroD1* expression could harbour a great potential for regenerative medicine. Temporary cell manipulations would omit potentially harmful properties of methodologies using sustained transgene expression from viral vectors, for example [43, 45, 287, 288]. However, such transient expression of a TF is technically and ethically challenging in *Homo sapiens* [287, 289]. Furthermore, it should be taken into account that reprogramming approaches using somatic cells harbour several risks, for example the accumulation of age-dependent mutations [290]. Nonetheless, several new approaches hold the promise for transdifferentiating somatic cells into neurons without the need of supplying exogenous transgenes. For example, small molecules were recently used as reprogramming stimuli for chemically induced neurons or NSCs by supplementing them to the culture media [288, 291-294]. It has been shown that molecules like certain HDAC inhibitors as well as p300/CBP-specific activators are able to induce neurogenesis in the adult hippocampus, probably by activating *NeuroD1* [295-297]. However, the exact molecular mechanism underlying this chemically-induced cell fate conversion remains elusive. These findings put forward a promising ground for future investigations on the potential for therapeutic interventions which directly manipulate endogenous levels of transcriptional regulators by altering the epigenetic landscape to convert cell fates into desired destinies.

5.2 The KRAB-zinc-finger protein Zfp354c is involved retrotransposon repression during neuronal differentiation

The family of zinc-finger proteins (ZFPs) comprises more genes than any other TF family in eukaryotes and its members have been implicated in numerous cellular processes [298-301]. However, a comprehensive picture of their lineage-specificity and functional role during cell fate acquisition is not fully unravelled. Several ZFPs have been identified as regulators of neurogenesis [302-307], but the majority of factors are poorly characterised. The applied comprehensive bioinformatic approach presented in this thesis identified the expression profile of zinc-finger containing factors (ZFPs) in the murine transcriptome across several embryonic murine tissues (ZincOme). These differential expression profiles of ZFPs among different tissues suggest that ZincOme factors play distinct roles in lineage-specific developmental processes. However, the accuracy of this approach should be refined by a comparison with other studies proposing differential expression patterns for distinct groups of ZFPs [308-310]. Nonetheless, this comprehensive resource can serve as the basis for future research towards functional characterisation of tissue-specific ZFPs during development.

5.2.1 Zfp354c is a nuclear KRAB-ZFP with a neuronal-specific expression pattern

The group of neuronal-specific ZFPs (445 genes; NeuroZincOme; see 8.7) comprised several KRAB-domain containing ZFPs. KRAB-ZFPs are TFs exclusively found in tetrapods [298, 311-314]. They have been described as sequence-specific recruiters of the KAP1/SETDB1 machinery and some KRAB-ZFPs were established as mediators of ERV silencing in ESCs [315-321]. Especially the proposed co-evolution of KRAB-ZFPs and ERVs [322-324] and the implication of transposable elements as evolutionary contributors to brain complexity [325, 326] intrigued to investigate the KRAB-ZFP/KAP1/SETDB1 system in the context of embryonic neurogenesis. Interestingly, only few attempts have been made in this direction so far [327-329].

This thesis investigated Zfp354c as a novel regulator involved in *in vivo* and *in vitro* neuronal differentiation. This KRAB-ZFP was selected due to its observed expression pattern which showed a specifically high induction from VZ to CP in the murine cortex. Interestingly, this finding is supported by recent reports for *Homo sapiens* which showed a high *ZNF354c* expression in foetal brain [330, 331]. Interestingly, although this factor is transcriptionally induced during mouse and human corticogenesis, *Zfp354c* has not been functionally explored for a potential role during neuronal development. Overall, very little is known about this protein [332-334], supporting the degree of novelty presented in this thesis. Despite several attempts, no commercially available antibody worked in immunohistochemistry analyses to detect Zfp354c in embryonic brain slices (data not shown). The generation of a custom antibody holds the promise to validate the observed transcriptional upregulation from VZ to CP at E14.5 also on protein level in the future. It would be furthermore intriguing to investigate the distribution of Zfp354c in the murine cortex throughout embryonic development as well as to identify potential interaction partners *in vivo*.

The neuroblastoma cell line Neuro2a [335, 336] was used in this thesis to get insights into properties and molecular mechanisms of *Zfp354c*. These cells exhibit the capacity to proliferate in culture and upon adequate stimuli to differentiate into neuronal cell types [335, 337], what marks them as a versatile cellular model system for molecular neurobiology. However, observations from immortalized cell lines cannot be transferred directly to *in vivo* neurodevelopmental contexts. Nevertheless, the findings of this thesis validate Zfp354c as a nuclear factor, strengthening its link to a functional role in the context of chromatin. This role was further supported by the analysis of Zfp354c domain mutants, which demonstrated that the KRAB domain of Zfp354c is essential for the interaction with the KRAB-associated protein-1 (KAP1). This is in line with the proposed function of KRAB-domains to mediate the interaction with KAP1 [338-342], which in turn recruits the methyltransferase SETDB1 to establish the heterochromatic histone modification H3K9me3 at target genomic loci [343-345]. These data implicate Zfp354c as a novel member of the KRAB-ZFP/KAP1/SETDB1 system which is known to initiate site-specific heterochromatin formation at target loci [312, 314, 323, 344, 346-348]. The nuclear localization [349] of Zfp354c depended on a sequence within the cluster of Zfp354c's eleven C2H2 zinc-fingers (ZFC). Strikingly, several reports observed that ZFCs, including their linker DNA, can mediate nuclear import of ZFPs due to interaction with nuclear importins [350-357]. In support of this observations are the *in vitro* derived IP-MS data from Neuro2a cells overexpressing HA-FLAG-Zfp354c. Two importins, Ipo7 (importin 7) and Kpnb1 (karyopherin beta 1; importin subunit beta-1) [358-363], were found as interaction partners of Zfp354c. Co-IPs with full length Zfp354c and the Zfp354c ZFC-mutant should be performed to validate that the ZFC is essential for such interaction with importins. The versatility of zinc-fingers, which are likely to not only function as a DBD but are also able to facilitate protein-protein interactions or to recognize RNA [364-369], suggests that this proposed interaction is in general possible. Further mutations of the ZFC by deleting each zinc-finger separately could be considered to break down which specific zinc-finger(s) might be responsible for this interaction. This functional versatility of C2H2 domains harbours the challenge of precisely predicting which ligand will be recognized by a particular zinc-finger [364-366, 370, 371].

5.2.2 Zfp354c might repress repetitive elements via KAP1-mediated formation of heterochromatin during *in vitro* neuronal differentiation

The retinoic acid mediated neuronal differentiation of Neuro2a cells is a model system to study neurogenesis-related processes *in vitro* [372-376]. Zfp354c was transcriptionally induced during neuronal differentiation of Neuro2a cells, suggesting a possible function during this process. To study the functional impact of *Zfp354c* during neuronal differentiation, cells with a stable *Zfp354c* depletion were generated by employing RNAi mediated mRNA degradation using shRNA [194, 377]. The successful depletion of *Zfp354c* by a single shRNA against the 3' UTR of *Zfp354c* (shZfp354c) severely influenced the generation of a neuronal cell morphology *in vitro*: The retinoic acid mediated differentiation of Neuro2a cells did not lead to the formation of TUJ1-positive processes in the manipulated cell population in comparison to scrambled control condition. Interestingly, several studies investigated quite a few regulators of neurite outgrowth during differentiation of Neuro2a cells [374, 378-382]. However, the observed phenotype upon *Zfp354c*

depletion served here as an indicator for altered neuronal differentiation but the underlying cause was not investigated in this thesis. Although the phenotype was reproducible in two clones of shZfp354c Neuro2a cells, shRNAs are reported to exhibit several off-target effects [383, 384]. Therefore at least one additional shZfp354c, transient transfection of siRNA against *Zfp354c* or RNA decoy [385] should be used in the future to validate the observed *in vitro* phenotype. Nonetheless, it was striking to find that the nuclei of *Zfp354c*-depleted Neuro2a cells exhibited a greater size upon differentiation in comparison to control conditions. Some reports suggest that nuclear size is altered upon perturbation of nuclear import [386-388]. As discussed above, the interaction partners of Zfp354c comprised two importin factors, supporting a link to the observed nuclear phenotype. However, the underlying mechanism is not clear and the regulation of nuclear size as such not fully understood [389, 390]. Nonetheless, increasing evidence shows that the nuclear envelope and pore complexes are functional regulators of the genome [391, 392] and a possible role for Zfp354c in this context should be further explored.

The alteration of nuclear size prompted to investigate the distribution of H3K9me3 in the nucleus as the IP-MS data linked Zfp354c to the KAP1/SETDB1 machinery [338, 340, 341, 346, 347, 393, 394]. Indeed, depletion of *Zfp354c* led to alterations of constitutive heterochromatin localization in the nucleus: whereas the global H3K9me3 levels were not affected upon knock down of *Zfp354c*, the appearance of H3K9me3 dense foci [395-397] on chromatin was ablated in the immunocytochemical analysis. This data suggests that Zfp354c might be essential for the locus specific accumulation of H3K9me3 dense heterochromatin foci. It further indicates that Zfp354c might directly bind DNA to allow such function. This hypothesis was supported by Zfp354c ChIP-seq data in Neuro2a cells, which revealed a high enrichment of this KRAB-ZFP at repetitive DNA sequences, especially LTR elements. It has to be noted that the computational analyses of repetitive sequences in NGS data is still challenging and no common guidelines are established [398-400]. The NGS data should therefore be interpreted with caution to avoid false claims. Although several specific target ERVs were validated by qPCR, the repetitive nature of these elements makes it technically challenging to investigate TF occupancy: no single genomic instance of the investigated ERVs is PCR amplified like for promoters of coding genes. Instead, all elements of the same ERV type are probed together with the used primers and PCR artefacts are likely to appear [401, 402]. However, the qPCR results for target ERVs were derived using mostly published primers [319, 321, 327] in order to be consistent with the literature. The *de novo* motif prediction at target repetitive elements showed significant enrichments for certain DNA motifs, suggesting a sequence specific nature of Zfp354c's occupancy. However, the primary sequence of target motifs was diverse and only a subset of target elements exhibited certain predictions. This data does therefore not permit to derive one specific motif for Zfp354c targeting to repetitive elements. The global profiling of the transcriptome and H3K9me3 occupancy in *Zfp354c*-depleted Neuro2a cells suggested that majorly Zfp354c target LTR elements loose heterochromatin silencing and become de-repressed in comparison to the control condition. The alteration of H3K9me3 levels and hence ERV reactivation is in line with previous reports for the KRAB-ZFP/KAP1/SETDB1 system in mESCs [320, 321, 323, 403, 404]. However, also several other transposons like LINE elements were influenced by *Zfp354c*-depletion *in vitro*, advocating that also the non-coding genome apart from ERV target regions exhibited

alterations of the repressive chromatin landscape. At this point it is not clear whether Zfp354c does indeed impact on the global H3K9me3 distribution on chromatin overall or if the detected de-repression of LTR elements leads to secondary effects on the nuclear chromatin landscape. Nonetheless, a global alteration of H3K9me3 distribution is supported by the above discussed immunocytochemical analysis, although the cause needs to be elucidated. Furthermore, an interplay of the KRAB-ZFP/KAP1 system with DNA methylation has been proposed as part of heterochromatin-mediated silencing of repetitive elements [318, 405]. The regulatory layer of DNA methylation-mediated silencing of ERVs should therefore be further explored in the context of *Zfp354c* function during neuronal differentiation. Taken together, the global NGS data is not unambiguous at this point to clearly delineate the function of Zfp354c at specific genomic loci. Future NGS approaches should employ longer reads (150 bp) or paired-end reads to increase the mapability of reads from repetitive elements against the reference genome [399, 400]. This new approach would improve the significance of computational analyses and potentially generate a clearer picture of Zfp354c's genomic target regions and hence the primary loci of chromatin landscape dynamics mediated by this KRAB-ZFP during neuronal differentiation.

Nonetheless, the NGS data suggested that the majority of Zfp354c target repetitive elements that lost H3K9me3 enrichment upon its depletion in Neuro2a cells belong to the class of LTR retrotransposons. Several of these LTR elements were therefore further investigated by ChIP-qPCR analyses. Strikingly, the tested ERV entities displayed a significant loss of KAP1 enrichment upon *Zfp354c*-depletion during neuronal differentiation of Neuro2a cells. Furthermore, a decrease of H3K9me3 deposition as well as a transcriptional increase of associated ERV elements paralleled this observation. These findings suggest that the failed acquisition of H3K9me3 levels during neuronal differentiation is caused by the decrease of KAP1 recruitment and hence its interaction partner SETDB1, a HMT for H3K9 [346, 347, 406]. The functional relationship between KRAB-ZFPs and SETDB1-mediated H3K9me3 deposition is widely supported in the literature [312, 319, 323, 329, 403]. However, the changes in HMT occupancy at identified target ERVs remains to be experimentally proven during neuronal differentiation of *Zfp354c*-depleted Neuro2a cells. Interestingly, among the studied Zfp354c target ERVs were Intracisternal A-Particle (IAP) as well as MMERVK10C elements, which have been shown to be de-repressed upon KAP1 depletion in neural progenitors [327]. IAP elements are among the most recent acquired retrotransposons in the murine genome and are able to actively transpose [407-411]. MMERVK10C elements are also believed to be recent intruders of the murine genome, but they are not well characterised till date [327, 412]. Interestingly, several LTR retrotransposon, including the Zfp354c target ERVs IAP, MMVL30 and RLTR45, have been described to be repressed via polycomb repressive complexes in ESCs [316, 412]. Furthermore, the histone variant H3.3 has been implicated in the KAP1/SETDB1 mediated silencing especially of IAP elements in mESCs [413, 414], but this was not further investigated in this thesis. These data suggest that these ERV elements might become embedded in a more stable H3K9me3-marked heterochromatin landscape during differentiation upon the cell type-specific action of KRAB-ZFPs. If these elements are indeed silenced by Zfp354c-mediated heterochromatin formation upon neuronal differentiation of mESCs should be further investigated to strengthen the functional relevance of the observation derived in the employed Neuro2a

model system. Taken together, these data put forward that at the selected target LTR elements Zfp354c might function as a sequence-specific mediator to silence these recently invaded ERVs by recruiting a heterochromatin machinery [312, 323]. These observations are among the first to indicate this function for a KRAB-ZFP during *in vitro* neuronal differentiation, but further investigation is needed to strengthen this claim. Interestingly, another class of transposable elements potentially regulated by Zfp354c are LINE1 elements. These non-LTR retrotransposons have been found to be transcriptionally active and to retrotranspose in NSCs [415-417], leading to somatic mosaicism in neurons [325]. As the expression of Zfp354c is specifically high in the embryonic cortex, this KRAB-ZFP might regulate LINE1 elements during *in vivo* neuronal differentiation. Although some reports indicated KRAB-ZFPs in the regulation of LINE1 elements [418, 419], the functional relevance has to be further explored. Additionally, the effect of *Zfp354c* depletion on coding genes has not been studied in this thesis. Zfp354c might directly regulate coding genes as indicated in an osteogenesis study [332]. Interestingly, Fasching and colleagues observed that KAP1 depletion and hence de-repression of IAP and MMERVK10C elements influenced gene expression in neural progenitors, especially lncRNAs [327]. This observation suggests that in the neuronal context at least an indirect effect on gene regulation might appear upon *Zfp354c* depletion. It is intriguing to further delineate which cellular responses [420] are triggered by *Zfp354c* depletion during *in vitro* neuronal differentiation apart from the observed alteration in TUJ1-positive process formation. However, in what way the identified target LTR elements are possibly regulated in other developmental contexts remains elusive.

5.2.3 *In vivo* neurogenesis is impaired upon *Zfp354c* depletion

The RT-qPCR expression profiling of *Zfp354c* and several of the identified target ERVs during *in vivo* corticogenesis exhibited an anti-correlation between the developmental increase of *Zfp354c* and target ERV expression. The employed IUE methodology allowed for the *in vivo* manipulation of *Zfp354c* levels in neural progenitor to test the functional relationship of the proposed role for *Zfp354c* in silencing ERVs. Strikingly, several target ERVs showed a tendency of transcriptional induction upon *Zfp354c* depletion during *in vivo* neurogenesis. These expression data suggest that *Zfp354c* might play a role during *in vivo* cortical development by silencing ERV elements. This hypothesis is further supported by the finding that literally no *Zfp354c*-depleted cells were found in cortical plate in comparison to the IUE with control constructs. It has to be noted that this data does not allow the claim that *Zfp354c* depletion causes a neuronal migration phenotype as no migration assays were performed [150]. Interestingly, a recent report proposed that some shRNAs exhibit an off-target toxicity which itself leads to neuronal migration defects by misregulation of endogenous miRNAs [421]. Although the here observed *in vivo* phenotype was dependent on *Zfp354c* depletion, the IUEs should be repeated with additional Zfp354c-specific shRNAs [422], as already discussed for the *in vitro* experiments above. Furthermore, it is not clear if the knock down-depended upregulation of transposons as such leads to increased cell death in the manipulated progenitor population by possibly compromising genome integrity [420]. This hypothesis should be assessed by evaluating cell viability upon IUE of shZfp354c by immunohistochemical analysis of cleaved caspase-3, for example [403, 423-427]. Nonetheless, the performed *in vivo* rescue experiments advocate for a KAP1-mediated cause of the

observed phenotype: Only the combination of full length Zfp354c together with the shZfp354c construct led to a cortical plate distribution of co-expressing cells as observed for the scrambled control. The KRAB-mutant was not able to rescue the shZfp354c-induced retention of cells in the VZ/SVZ as also observed for the control co-electroporation. These data suggest that a fine-tuned level of Zfp354c is critical for *in vivo* neurogenesis, possibly due to the coordinated regulation of KRAB-ZFP/KAP1/SETDB1-mediated repression of non-coding genomic regions. The importance of this silencing mechanism for detrimental genomic invaders during embryonic development is evident by the observed embryonic lethality upon knockout of SETDB1 or KAP1 [319, 403]. However, the mechanistic details for the proposed Zfp354c-mediated regulation of the non-coding genome during *in vivo* neurogenesis has to be further unravelled. Future investigations should encounter how the epigenetic regulation of repetitive elements via KRAB-ZFPs contributes to brain development in health and disease, especially in the light of a potential regulatory function of LTR elements during development [420, 428-430].

5.3 Concluding remarks

The regulation of gene expression and the adaptation of the transcriptome to environmental cues are fundamental principles of cell fate changes during ontogeny of living organisms. A plethora of regulatory layers has evolved whose crosstalk integrates cell-intrinsic as well as cell-extrinsic signals underlying the tremendous diversity of biological processes in multicellular organism. It is fascinating how the marvelous amount of different cellular phenotypes is derived by regulating the functional output of the totipotent zygotic genome during ontogeny. Of special interest in this thesis was the process of embryonic neurogenesis in *Mus musculus*. The functional interplay between nuclear factors and the chromatin landscape in the regulation of the neuronal genomic output was shown here on the level of coding genes for the bHLH transcription factor NeuroD1. The observations of this thesis suggest a novel function for NeuroD1 as a potential pioneer transcription factor for cell fate specification of excitatory neurons in the dorsal telencephalon. Furthermore, the presented findings promote the concept that the transient action of key developmental regulators for a specific lineage sets up alterations to gene regulatory networks as well as to the chromatin landscape during cellular differentiation. This time- and context-dependent action leads to a sustained fate-specific transcriptome, accentuating the importance of key developmental stimuli during ontogeny. Furthermore, this thesis has expanded the view on the importance of non-coding genome regulation during neuronal differentiation. The KRAB-ZFP Zfp354c was proposed to mediate the repression of repetitive elements, especially LTR-containing endogenous retroviruses, during neuronal differentiation. This novel member of KAP1-interacting transcription factors seems to be important for the H3K9me3-dependend heterochromatin formation at recently invaded transposable elements during neuronal cell fate acquisition. Taken together, these findings illustrate the complex dynamics within the nucleus of a cell that undergoes neuronal differentiation.

Several principles of neurodevelopment were discovered in mice and are transferrable to humans, supporting a common evolutionary origin of the mammalian neocortex. However, there are a number of

developmental differences in embryonic neurogenesis between these two species that advocate for separate functional validations of the here presented findings in a human background. Nonetheless, a deeper understanding of the molecular layers orchestrating cell fate determination during development can benefit approaches of regenerative medicine. These approaches routinely aim for ectopically induced cell fate changes in order to restore cell populations affected by disease states. Strikingly, the attempts of cellular reprogramming and transdifferentiation are increasingly challenging the concept of terminal differentiation as cell states appear more plastic than previously believed. A better understanding of the molecular barriers that stabilise these developmentally acquired states under healthy physiological conditions holds the promise to enable their restoration in diseased states like cancer or cellular degeneration. The molecular tools available today are a great asset to decipher the function of specific genomic elements in the complex landscape of gene regulation. Especially the advent of CRISPR/Cas9 technology enables researchers for locus specific manipulations throughout development. However, as exemplified in human evolution, great power comes with great responsibility and preventive measures have to be taken for a reasonable usage of these influential technologies.

“The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.”

Isaac Asimov

“Altering a gene in the gene line to produce improved offspring is likely to be very difficult because of the danger of unwanted side effects. It would also raise obvious ethical problems.”

Francis Crick

“Reliable scientific knowledge is value free and has no moral or ethical value. Science tells us how the world is. … Dangers and ethical issue arise only when science is applied as technology.”

Lewis Wolpert

5.4 References Discussion

1. Schaffner, W., *Enhancers, enhancers - from their discovery to today's universe of transcription enhancers.* Biol Chem, 2015. **396**(4): p. 311-27.
2. Voss, T.C. and G.L. Hager, *Dynamic regulation of transcriptional states by chromatin and transcription factors.* Nat Rev Genet, 2014. **15**(2): p. 69-81.
3. Zabidi, M.A., et al., *Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation.* Nature, 2015. **518**(7540): p. 556-9.
4. Zabidi, M.A. and A. Stark, *Regulatory Enhancer-Core-Promoter Communication via Transcription Factors and Cofactors.* Trends Genet, 2016. **32**(12): p. 801-814.
5. Allis, C.D. and T. Jenuwein, *The molecular hallmarks of epigenetic control.* Nat Rev Genet, 2016. **17**(8): p. 487-500.
6. Iwafuchi-Doi, M. and K.S. Zaret, *Pioneer transcription factors in cell reprogramming.* Genes Dev, 2014. **28**(24): p. 2679-92.
7. Zaret, K.S. and J.S. Carroll, *Pioneer transcription factors: establishing competence for gene expression.* Genes Dev, 2011. **25**(21): p. 2227-41.
8. Zaret, K.S. and S.E. Mango, *Pioneer transcription factors, chromatin dynamics, and cell fate control.* Curr Opin Genet Dev, 2016. **37**: p. 76-81.
9. Amamoto, R. and P. Arlotta, *Development-inspired reprogramming of the mammalian central nervous system.* Science, 2014. **343**(6170): p. 1239882.
10. Perino, M. and G.J. Veenstra, *Chromatin Control of Developmental Dynamics and Plasticity.* Dev Cell, 2016. **38**(6): p. 610-20.
11. Shchuka, V.M., et al., *Chromatin Dynamics in Lineage Commitment and Cellular Reprogramming.* Genes (Basel), 2015. **6**(3): p. 641-61.
12. Amoutzias, G.D., et al., *Convergent evolution of gene networks by single-gene duplications in higher eukaryotes.* EMBO Rep, 2004. **5**(3): p. 274-9.
13. Charoensawan, V., D. Wilson, and S.A. Teichmann, *Lineage-specific expansion of DNA-binding transcription factor families.* Trends Genet, 2010. **26**(9): p. 388-93.
14. de Mendoza, A., et al., *Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages.* Proc Natl Acad Sci U S A, 2013. **110**(50): p. E4858-66.
15. Schmitz, J.F., F. Zimmer, and E. Bornberg-Bauer, *Mechanisms of transcription factor evolution in Metazoa.* Nucleic Acids Res, 2016. **44**(13): p. 6287-97.
16. Maerkl, S.J. and S.R. Quake, *Experimental determination of the evolvability of a transcription factor.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18650-5.
17. Armstrong, L., et al., *Editorial: Our top 10 developments in stem cell biology over the last 30 years.* Stem Cells, 2012. **30**(1): p. 2-9.
18. Jaenisch, R. and R. Young, *Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming.* Cell, 2008. **132**(4): p. 567-82.
19. Nishikawa, S., L.M. Jakt, and T. Era, *Embryonic stem-cell culture as a tool for developmental cell biology.* Nat Rev Mol Cell Biol, 2007. **8**(6): p. 502-7.
20. Niwa, H., *Mouse ES cell culture system as a model of development.* Dev Growth Differ, 2010. **52**(3): p. 275-83.
21. Posfai, E., O.H. Tam, and J. Rossant, *Mechanisms of pluripotency in vivo and in vitro.* Curr Top Dev Biol, 2014. **107**: p. 1-37.
22. Spagnoli, F.M. and A. Hemmati-Brivanlou, *Guiding embryonic stem cells towards differentiation: lessons from molecular embryology.* Curr Opin Genet Dev, 2006. **16**(5): p. 469-75.
23. Turner, D.A., P. Baillie-Johnson, and A. Martinez Arias, *Organoids and the genetically encoded self-assembly of embryonic stem cells.* Bioessays, 2016. **38**(2): p. 181-91.
24. Coskun, V., R. Tsoa, and Y.E. Sun, *Epigenetic regulation of stem cells differentiating along the neural lineage.* Curr Opin Neurobiol, 2012. **22**(5): p. 762-7.
25. Hansen, D.V., J.L. Rubenstein, and A.R. Kriegstein, *Deriving excitatory neurons of the neocortex from pluripotent stem cells.* Neuron, 2011. **70**(4): p. 645-60.
26. Hemberger, M., W. Dean, and W. Reik, *Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal.* Nat Rev Mol Cell Biol, 2009. **10**(8): p. 526-37.
27. Keller, G., *Embryonic stem cell differentiation: emergence of a new era in biology and medicine.* Genes Dev, 2005. **19**(10): p. 1129-55.

28. Muguruma, K. and Y. Sasai, *In vitro recapitulation of neural development using embryonic stem cells: from neurogenesis to histogenesis.* Dev Growth Differ, 2012. **54**(3): p. 349-57.

29. Olynik, B.M. and M. Rastegar, *The genetic and epigenetic journey of embryonic stem cells into mature neural cells.* Front Genet, 2012. **3**: p. 81.

30. Naya, F.J., C.M. Stellrecht, and M.J. Tsai, *Tissue-specific regulation of the insulin gene by a novel basic helix-loop-helix transcription factor.* Genes Dev, 1995. **9**(8): p. 1009-19.

31. Lee, J.E., et al., *Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein.* Science, 1995. **268**(5212): p. 836-44.

32. Aprea, J., M. Nonaka-Kinoshita, and F. Calegari, *Generation and characterization of Neurod1-CreER(T2) mouse lines for the study of embryonic and adult neurogenesis.* Genesis, 2014. **52**(10): p. 870-8.

33. Cho, J.H. and M.J. Tsai, *The role of BETA2/NeuroD1 in the development of the nervous system.* Mol Neurobiol, 2004. **30**(1): p. 35-47.

34. Kim, W.Y., *NeuroD regulates neuronal migration.* Mol Cells, 2013. **35**(5): p. 444-9.

35. Lee, J.E., *NeuroD and neurogenesis.* Dev Neurosci, 1997. **19**(1): p. 27-32.

36. Morrow, E.M., et al., *NeuroD regulates multiple functions in the developing neural retina in rodent.* Development, 1999. **126**(1): p. 23-36.

37. Hevner, R.F., et al., *Transcription factors in glutamatergic neurogenesis: conserved programs in neocortex, cerebellum, and adult hippocampus.* Neurosci Res, 2006. **55**(3): p. 223-33.

38. Naya, F.J., et al., *Diabetes, defective pancreatic morphogenesis, and abnormal enteroendocrine differentiation in BETA2/neuroD-deficient mice.* Genes Dev, 1997. **11**(18): p. 2323-34.

39. Brulet, R., et al., *NEUROD1 Instructs Neuronal Conversion in Non-Reactive Astrocytes.* Stem Cell Reports, 2017. **8**(6): p. 1506-1515.

40. Guo, Z., et al., *In vivo direct reprogramming of reactive glial cells into functional neurons after brain injury and in an Alzheimer's disease model.* Cell Stem Cell, 2014. **14**(2): p. 188-202.

41. Chen, W., et al., *Lentivirus carrying the NeuroD1 gene promotes the conversion from glial cells into neurons in a spinal cord injury model.* Brain Res Bull, 2017.

42. Vermunt, M.W. and M.P. Creyghton, *Transcriptional Dynamics at Brain Enhancers: from Functional Specialization to Neurodegeneration.* Curr Neurol Neurosci Rep, 2016. **16**(10): p. 94.

43. Mertens, J., et al., *Evaluating cell reprogramming, differentiation and conversion technologies in neuroscience.* Nat Rev Neurosci, 2016. **17**(7): p. 424-37.

44. Mahla, R.S., *Stem Cells Applications in Regenerative Medicine and Disease Therapeutics.* Int J Cell Biol, 2016. **2016**: p. 6940283.

45. Jopling, C., S. Boue, and J.C. Izpisua Belmonte, *Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration.* Nat Rev Mol Cell Biol, 2011. **12**(2): p. 79-89.

46. Jung, J., *Gene Regulation by Topoisomerase IIa: A Dual Role in Embryonic Stem Cells.* Johannes Gutenberg-University, Mainz, 2013.

47. Cahan, P. and G.Q. Daley, *Origins and implications of pluripotent stem cell variability and heterogeneity.* Nat Rev Mol Cell Biol, 2013. **14**(6): p. 357-68.

48. Chambers, I., et al., *Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells.* Cell, 2003. **113**(5): p. 643-55.

49. Martello, G. and A. Smith, *The nature of embryonic stem cells.* Annu Rev Cell Dev Biol, 2014. **30**: p. 647-75.

50. MacArthur, B.D. and I.R. Lemischka, *Statistical mechanics of pluripotency.* Cell, 2013. **154**(3): p. 484-9.

51. Niwa, H., et al., *A parallel circuit of LIF signalling pathways maintains pluripotency of mouse ES cells.* Nature, 2009. **460**(7251): p. 118-22.

52. Kalmar, T., et al., *Regulated fluctuations in nanog expression mediate cell fate decisions in embryonic stem cells.* PLoS Biol, 2009. **7**(7): p. e1000149.

53. Davies, O.R., et al., *Tcf15 primes pluripotent cells for differentiation.* Cell Rep, 2013. **3**(2): p. 472-84.

54. Ying, Q.L., et al., *BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3.* Cell, 2003. **115**(3): p. 281-92.

55. Zhang, K., et al., *Distinct functions of BMP4 during different stages of mouse ES cell neural commitment.* Development, 2010. **137**(13): p. 2095-105.

56. Kunath, T., et al., *FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment.* Development, 2007. **134**(16): p. 2895-902.

57. Yuan, H., et al., *Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3.* Genes Dev, 1995. **9**(21): p. 2635-45.

58. Tee, W.W., et al., *Erk1/2 activity promotes chromatin features and RNAPII phosphorylation at developmental promoters in mouse ESCs.* Cell, 2014. **156**(4): p. 678-90.

59. Boroviak, T., et al., *The ability of inner-cell-mass cells to self-renew as embryonic stem cells is acquired following epiblast specification.* Nat Cell Biol, 2014. **16**(6): p. 516-28.

60. Kim, S.H., et al., *ERK1 phosphorylates Nanog to regulate protein stability and stem cell self-renewal.* Stem Cell Res, 2014. **13**(1): p. 1-11.

61. Li, M. and J.C. Belmonte, *Ground rules of the pluripotency gene regulatory network.* Nat Rev Genet, 2017. **18**(3): p. 180-191.

62. Marks, H., et al., *The transcriptional and epigenomic foundations of ground state pluripotency.* Cell, 2012. **149**(3): p. 590-604.

63. Ying, Q.L., et al., *The ground state of embryonic stem cell self-renewal.* Nature, 2008. **453**(7194): p. 519-23.

64. Young, R.A., *Control of the embryonic stem cell state.* Cell, 2011. **144**(6): p. 940-54.

65. Yeo, J.C., et al., *Klf2 is an essential factor that sustains ground state pluripotency.* Cell Stem Cell, 2014. **14**(6): p. 864-72.

66. Shen, Y., et al., *A map of the cis-regulatory sequences in the mouse genome.* Nature, 2012. **488**(7409): p. 116-20.

67. Jones, S., *An overview of the basic helix-loop-helix proteins.* Genome Biol, 2004. **5**(6): p. 226.

68. Longo, A., G.P. Guanga, and R.B. Rose, *Crystal structure of E47-NeuroD1/beta2 bHLH domain-DNA complex: heterodimer selectivity and DNA recognition.* Biochemistry, 2008. **47**(1): p. 218-29.

69. Imayoshi, I. and R. Kageyama, *bHLH factors in self-renewal, multipotency, and fate choice of neural progenitor cells.* Neuron, 2014. **82**(1): p. 9-23.

70. Powell, L.M. and A.P. Jarman, *Context dependence of proneural bHLH proteins.* Curr Opin Genet Dev, 2008. **18**(5): p. 411-7.

71. Flora, A., et al., *The E-protein Tcf4 interacts with Math1 to regulate differentiation of a specific subset of neuronal progenitors.* Proc Natl Acad Sci U S A, 2007. **104**(39): p. 15382-7.

72. Castanon, I., et al., *Dimerization partners determine the activity of the Twist bHLH protein during Drosophila mesoderm development.* Development, 2001. **128**(16): p. 3145-59.

73. Castro, D.S., et al., *Proneural bHLH and Brn proteins coregulate a neurogenic program through cooperative binding to a conserved DNA motif.* Dev Cell, 2006. **11**(6): p. 831-44.

74. Mattar, P., et al., *Basic helix-loop-helix transcription factors cooperate to specify a cortical projection neuron identity.* Mol Cell Biol, 2008. **28**(5): p. 1456-69.

75. Nakada, Y., et al., *Distinct domains within Mash1 and Math1 are required for function in neuronal differentiation versus neuronal cell-type specification.* Development, 2004. **131**(6): p. 1319-30.

76. Saba, R., J.E. Johnson, and T. Saito, *Commissural neuron identity is specified by a homeodomain protein, Mbh1, that is directly downstream of Math1.* Development, 2005. **132**(9): p. 2147-55.

77. Quan, X.J., et al., *Evolution of neural precursor selection: functional divergence of proneural proteins.* Development, 2004. **131**(8): p. 1679-89.

78. Zhou, Q. and D.J. Anderson, *The bHLH transcription factors OLIG2 and OLIG1 couple neuronal and glial subtype specification.* Cell, 2002. **109**(1): p. 61-73.

79. Rhee, H.S. and B.F. Pugh, *ChIP-exo method for identifying genomic location of DNA-binding proteins with near-single-nucleotide accuracy.* Curr Protoc Mol Biol, 2012. **Chapter 21**: p. Unit 21 24.

80. Busskamp, V., et al., *Rapid neurogenesis through transcriptional activation in human stem cells.* Mol Syst Biol, 2014. **10**: p. 760.

81. Lein, E.S., et al., *Genome-wide atlas of gene expression in the adult mouse brain.* Nature, 2007. **445**(7124): p. 168-76.

82. Liu, Q.Y., et al., *A novel brain-enriched E3 ubiquitin ligase RNF182 is up regulated in the brains of Alzheimer's patients and targets ATP6V0C for degradation.* Mol Neurodegener, 2008. **3**: p. 4.

83. Steichen-Gersdorf, E., et al., *Triangular tibia with fibular aplasia associated with a microdeletion on 2q11.2 encompassing LAF4.* Clin Genet, 2008. **74**(6): p. 560-5.

84. Metsu, S., et al., *FRA2A is a CGG repeat expansion associated with silencing of AFF3.* PLoS Genet, 2014. **10**(4): p. e1004242.

85. Moore, J.M., et al., *Laf4/Aff3, a gene involved in intellectual disability, is required for cellular migration in the mouse cerebral cortex.* PLoS One, 2014. **9**(8): p. e105933.

86. Sahu, S.K., et al., *JNK-dependent gene regulatory circuitry governs mesenchymal fate.* EMBO J, 2015. **34**(16): p. 2162-81.

87. Itoh, Y., et al., *Scratch regulates neuronal migration onset via an epithelial-mesenchymal transition-like mechanism.* Nat Neurosci, 2013. **16**(4): p. 416-25.

88. Singh, S. and D.J. Solecki, *Polarity transitions during neurogenesis and germinal zone exit in the developing central nervous system.* Front Cell Neurosci, 2015. **9**: p. 62.

89. Baum, B., J. Settleman, and M.P. Quinlan, *Transitions between epithelial and mesenchymal states in development and disease.* Semin Cell Dev Biol, 2008. **19**(3): p. 294-308.

90. Chaffer, C.L., E.W. Thompson, and E.D. Williams, *Mesenchymal to epithelial transition in development and disease.* Cells Tissues Organs, 2007. **185**(1-3): p. 7-19.

91. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.

92. Barrallo-Gimeno, A. and M.A. Nieto, *The Snail genes as inducers of cell movement and survival: implications in development and cancer.* Development, 2005. **132**(14): p. 3151-61.

93. Carver, E.A., et al., *The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition.* Mol Cell Biol, 2001. **21**(23): p. 8184-8.

94. Oda, H., S. Tsukita, and M. Takeichi, *Dynamic behavior of the cadherin-based cell-cell adhesion system during Drosophila gastrulation.* Dev Biol, 1998. **203**(2): p. 435-50.

95. Zander, M.A., et al., *Snail coordinately regulates downstream pathways to control multiple aspects of mammalian neural precursor development.* J Neurosci, 2014. **34**(15): p. 5164-75.

96. Peinado, H., D. Olmeda, and A. Cano, *Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?* Nat Rev Cancer, 2007. **7**(6): p. 415-28.

97. Huang, P., et al., *The neuronal differentiation factor NeuroD1 downregulates the neuronal repellent factor Slit2 expression and promotes cell motility and tumor formation of neuroblastoma.* Cancer Res, 2011. **71**(8): p. 2938-48.

98. Osborne, J.K., et al., *NeuroD1 mediates nicotine-induced migration and invasion via regulation of the nicotinic acetylcholine receptor subunits in a subset of neural and neuroendocrine carcinomas.* Mol Biol Cell, 2014. **25**(11): p. 1782-92.

99. Osborne, J.K., et al., *NeuroD1 regulation of migration accompanies the differential sensitivity of neuroendocrine carcinomas to TrkB inhibition.* Oncogenesis, 2013. **2**: p. e63.

100. Thoma, E.C., et al., *Ectopic expression of neurogenin 2 alone is sufficient to induce differentiation of embryonic stem cells into mature neurons.* PLoS One, 2012. **7**(6): p. e38651.

101. Velkey, J.M. and K.S. O'Shea, *Expression of Neurogenin 1 in mouse embryonic stem cells directs the differentiation of neuronal precursors and identifies unique patterns of down-stream gene expression.* Dev Dyn, 2013. **242**(3): p. 230-53.

102. Sommer, L., Q. Ma, and D.J. Anderson, *neurogenins, a novel family of atonal-related bHLH transcription factors, are putative mammalian neuronal determination genes that reveal progenitor cell heterogeneity in the developing CNS and PNS.* Mol Cell Neurosci, 1996. **8**(4): p. 221-41.

103. Wilkinson, G., D. Dennis, and C. Schuurmans, *Proneural genes in neocortical development.* Neuroscience, 2013. **253**: p. 256-73.

104. Pang, Z.P., et al., *Induction of human neuronal cells by defined transcription factors.* Nature, 2011. **476**(7359): p. 220-3.

105. Fode, C., et al., *The bHLH protein NEUROGENIN 2 is a determination factor for epibranchial placode-derived sensory neurons.* Neuron, 1998. **20**(3): p. 483-94.

106. Ma, Q., et al., *neurogenin1 is essential for the determination of neuronal precursors for proximal cranial sensory ganglia.* Neuron, 1998. **20**(3): p. 469-82.

107. Koyano-Nakagawa, N., et al., *Hes6 acts in a positive feedback loop with the neurogenins to promote neuronal differentiation.* Development, 2000. **127**(19): p. 4203-16.

108. Gaiano, N. and G. Fishell, *The role of notch in promoting glial and neural stem cell fates.* Annu Rev Neurosci, 2002. **25**: p. 471-90.

109. Kageyama, R., et al., *Roles of bHLH genes in neural stem cell differentiation.* Exp Cell Res, 2005. **306**(2): p. 343-8.

110. Kageyama, R., T. Ohtsuka, and T. Kobayashi, *The Hes gene family: repressors and oscillators that orchestrate embryogenesis.* Development, 2007. **134**(7): p. 1243-51.

111. Bae, S., et al., *The bHLH gene Hes6, an inhibitor of Hes1, promotes neuronal differentiation.* Development, 2000. **127**(13): p. 2933-43.

112. Gratton, M.O., et al., *Hes6 promotes cortical neurogenesis and inhibits Hes1 transcription repression activity by multiple mechanisms.* Mol Cell Biol, 2003. **23**(19): p. 6922-35.

113. Jhas, S., et al., *Hes6 inhibits astrocyte differentiation and promotes neurogenesis through different mechanisms.* J Neurosci, 2006. **26**(43): p. 11061-71.

114. Murai, K., A. Philpott, and P.H. Jones, *Hes6 is required for the neurogenic activity of neurogenin and NeuroD.* PLoS One, 2011. **6**(11): p. e27880.

115. Bao, J., et al., *Regulation of neurogenesis by interactions between HEN1 and neuronal LMO proteins.* Development, 2000. **127**(2): p. 425-35.

116. Kim, W.Y., *NeuroD1 is an upstream regulator of NSCL1.* Biochem Biophys Res Commun, 2012. **419**(1): p. 27-31.

117. Li, C.M., R.T. Yan, and S.Z. Wang, *Misexpression of a bHLH gene, cNSCL1, results in abnormal brain development.* Dev Dyn, 1999. **215**(3): p. 238-47.

118. Murdoch, J.N., et al., *Sequence and expression analysis of Nhlh1: a basic helix-loop-helix gene implicated in neurogenesis.* Dev Genet, 1999. **24**(1-2): p. 165-77.

119. Kruger, M. and T. Braun, *The neuronal basic helix-loop-helix transcription factor NSCL-1 is dispensable for normal neuronal development.* Mol Cell Biol, 2002. **22**(3): p. 792-800.

120. Dunwoodie, S.L., et al., *Mouse Dll3: a novel divergent Delta gene which may complement the function of other Delta homologues during early pattern formation in the mouse embryo.* Development, 1997. **124**(16): p. 3065-76.

121. Campos, L.S., et al., *mDll1 and mDll3 expression in the developing mouse brain: role in the establishment of the early cortex.* J Neurosci Res, 2001. **64**(6): p. 590-8.

122. Sparrow, D.B., et al., *Diverse requirements for Notch signalling in mammals.* Int J Dev Biol, 2002. **46**(4): p. 365-74.

123. Ladi, E., et al., *The divergent DSL ligand Dll3 does not activate Notch signaling but cell autonomously attenuates signaling induced by other DSL ligands.* J Cell Biol, 2005. **170**(6): p. 983-92.

124. Seo, S., et al., *Neurogenin and NeuroD direct transcriptional targets and their regulatory enhancers.* EMBO J, 2007. **26**(24): p. 5093-108.

125. Mattar, P., et al., *A screen for downstream effectors of Neurogenin2 in the embryonic neocortex.* Dev Biol, 2004. **273**(2): p. 373-89.

126. Ghil, S.H., Y.J. Jeon, and H. Suh-Kim, *Inhibition of BETA2/NeuroD by Id2.* Exp Mol Med, 2002. **34**(5): p. 367-73.

127. Miyata, T., T. Maeda, and J.E. Lee, *NeuroD is required for differentiation of the granule cells in the cerebellum and hippocampus.* Genes Dev, 1999. **13**(13): p. 1647-52.

128. Kim, W.Y., et al., *NeuroD-null mice are deaf due to a severe loss of the inner ear sensory neurons during development.* Development, 2001. **128**(3): p. 417-26.

129. Pennesi, M.E., et al., *BETA2/NeuroD1 null mice: a new model for transcription factor-dependent photoreceptor degeneration.* J Neurosci, 2003. **23**(2): p. 453-61.

130. Rubio-Cabezas, O., et al., *Homozygous mutations in NEUROD1 are responsible for a novel syndrome of permanent neonatal diabetes and neurological abnormalities.* Diabetes, 2010. **59**(9): p. 2326-31.

131. Azuara, V., et al., *Chromatin signatures of pluripotent cell lines.* Nat Cell Biol, 2006. **8**(5): p. 532-8.

132. Bernstein, B.E., et al., *A bivalent chromatin structure marks key developmental genes in embryonic stem cells.* Cell, 2006. **125**(2): p. 315-26.

133. Boyer, L.A., et al., *Polycomb complexes repress developmental regulators in murine embryonic stem cells.* Nature, 2006. **441**(7091): p. 349-53.

134. Creyghton, M.P., et al., *Histone H3K27ac separates active from poised enhancers and predicts developmental state.* Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21931-6.

135. Rada-Iglesias, A., et al., *A unique chromatin signature uncovers early developmental enhancers in humans.* Nature, 2011. **470**(7333): p. 279-83.

136. Shlyueva, D., G. Stampfel, and A. Stark, *Transcriptional enhancers: from properties to genome-wide predictions.* Nat Rev Genet, 2014. **15**(4): p. 272-86.

137. Zhu, J., et al., *Genome-wide chromatin state transitions associated with developmental and environmental cues.* Cell, 2013. **152**(3): p. 642-54.

138. Zhu, Y., et al., *Predicting enhancer transcription and activity from chromatin modifications.* Nucleic Acids Res, 2013. **41**(22): p. 10032-43.

139. Pradeepa, M.M., *Causal role of histone acetylations in enhancer function.* Transcription, 2017. **8**(1): p. 40-47.

140. Atlasi, Y. and H.G. Stunnenberg, *The interplay of epigenetic marks during stem cell differentiation and development.* Nat Rev Genet, 2017. **18**(11): p. 643-658.

141. Cusanovich, D.A., et al., *The functional consequences of variation in transcription factor binding.* PLoS Genet, 2014. **10**(3): p. e1004226.

142. Fukaya, T., B. Lim, and M. Levine, *Enhancer Control of Transcriptional Bursting.* Cell, 2016. **166**(2): p. 358-368.

143. Hnisz, D., et al., *A Phase Separation Model for Transcriptional Control.* Cell, 2017. **169**(1): p. 13-23.

144. Lee, K., et al., *Dynamic enhancer-gene body contacts during transcription elongation.* Genes Dev, 2015. **29**(19): p. 1992-7.

145. Zhang, Y., et al., *Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations.* Nature, 2013. **504**(7479): p. 306-310.

146. Coulon, A., et al., *Eukaryotic transcriptional dynamics: from single molecules to cell populations.* Nat Rev Genet, 2013. **14**(8): p. 572-84.

147. Lickwar, C.R., F. Mueller, and J.D. Lieb, *Genome-wide measurement of protein-DNA binding dynamics using competition ChIP.* Nat Protoc, 2013. **8**(7): p. 1337-53.

148. Presman, D.M., et al., *Quantifying transcription factor binding dynamics at the single-molecule level in live cells.* Methods, 2017. **123**: p. 76-88.

149. Ziller, M.J., et al., *Dissecting neural differentiation regulatory networks through epigenetic footprinting.* Nature, 2015. **518**(7539): p. 355-359.

150. Azzarelli, R., et al., *In Vitro, Ex Vivo and In Vivo Techniques to Study Neuronal Migration in the Developing Cerebral Cortex.* Brain Sci, 2017. **7**(5).

151. Evsyukova, I., C. Plestant, and E.S. Anton, *Integrative mechanisms of oriented neuronal migration in the developing brain.* Annu Rev Cell Dev Biol, 2013. **29**: p. 299-353.

152. Taverna, E., M. Gotz, and W.B. Huttner, *The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex.* Annu Rev Cell Dev Biol, 2014. **30**: p. 465-502.

153. dal Maschio, M., et al., *High-performance and site-directed in utero electroporation by a triple-electrode probe.* Nat Commun, 2012. **3**: p. 960.

154. Saito, T., *Electroporation Methods in Neuroscience.* Humana Press, 2015. **102**.

155. Szczurkowska, J., et al., *Targeted in vivo genetic manipulation of the mouse or rat brain by in utero electroporation with a triple-electrode probe.* Nat Protoc, 2016. **11**(3): p. 399-412.

156. Hsu, H.T., et al., *TRANSCRIPTION. Recruitment of RNA polymerase II by the pioneer transcription factor PHA-4.* Science, 2015. **348**(6241): p. 1372-6.

157. Adams, C.C. and J.L. Workman, *Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative.* Mol Cell Biol, 1995. **15**(3): p. 1405-21.

158. Soufi, A., G. Donahue, and K.S. Zaret, *Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome.* Cell, 2012. **151**(5): p. 994-1004.

159. Soufi, A., et al., *Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming.* Cell, 2015. **161**(3): p. 555-568.

160. Zaret, K.S., J. Lerner, and M. Iwafuchi-Doi, *Chromatin Scanning by Dynamic Binding of Pioneer Factors.* Mol Cell, 2016. **62**(5): p. 665-7.

161. Sauve, S., L. Tremblay, and P. Lavigne, *The NMR solution structure of a mutant of the Max b/HLH/LZ free of DNA: insights into the specific and reversible DNA binding mechanism of dimeric transcription factors.* J Mol Biol, 2004. **342**(3): p. 813-32.

162. Becker, J.S., D. Nicetto, and K.S. Zaret, *H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes.* Trends Genet, 2016. **32**(1): p. 29-41.

163. Chen, J., et al., *H3K9 methylation is a barrier during somatic cell reprogramming into iPSCs.* Nat Genet, 2013. **45**(1): p. 34-42.

164. Sridharan, R., et al., *Proteomic and genomic approaches reveal critical functions of H3K9 methylation and heterochromatin protein-1gamma in reprogramming to pluripotency.* Nat Cell Biol, 2013. **15**(7): p. 872-82.

165. Zhu, H., G. Wang, and J. Qian, *Transcription factors as readers and effectors of DNA methylation.* Nat Rev Genet, 2016. **17**(9): p. 551-65.

166. Barrangou, R., et al., *CRISPR provides acquired resistance against viruses in prokaryotes.* Science, 2007. **315**(5819): p. 1709-12.

167. Dominguez, A.A., W.A. Lim, and L.S. Qi, *Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation.* Nat Rev Mol Cell Biol, 2016. **17**(1): p. 5-15.

168. Garneau, J.E., et al., *The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA.* Nature, 2010. **468**(7320): p. 67-71.

169. Gasiunas, G., et al., *Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria.* Proc Natl Acad Sci U S A, 2012. **109**(39): p. E2579-86.

170. Jinek, M., et al., *A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity.* Science, 2012. **337**(6096): p. 816-21.

171. Marraffini, L.A. and E.J. Sontheimer, *CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea.* Nat Rev Genet, 2010. **11**(3): p. 181-90.

172. Qi, L.S., et al., *Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression.* Cell, 2013. **152**(5): p. 1173-83.

173. Wiedenheft, B., S.H. Sternberg, and J.A. Doudna, *RNA-guided genetic silencing systems in bacteria and archaea.* Nature, 2012. **482**(7385): p. 331-8.

174. Hilton, I.B., et al., *Epigenome editing by a CRISPR-Cas9-based acetyltransferase activates genes from promoters and enhancers.* Nat Biotechnol, 2015. **33**(5): p. 510-7.

175. Mutoh, H., et al., *The basic helix-loop-helix protein BETA2 interacts with p300 to coordinate differentiation of secretin-expressing enteroendocrine cells.* Genes Dev, 1998. **12**(6): p. 820-30.

176. Sharma, A., et al., *The NeuroD1/BETA2 sequences essential for insulin gene transcription colocalize with those necessary for neurogenesis and p300/CREB binding protein binding.* Mol Cell Biol, 1999. **19**(1): p. 704-13.

177. Qiu, Y., A. Sharma, and R. Stein, *p300 mediates transcriptional stimulation by the basic helix-loop-helix activators of the insulin gene.* Mol Cell Biol, 1998. **18**(5): p. 2957-64.

178. Qiu, Y., et al., *Insulin gene transcription is mediated by interactions between the p300 coactivator and PDX-1, BETA2, and E47.* Mol Cell Biol, 2002. **22**(2): p. 412-20.

179. Albert, M., et al., *Epigenome profiling and editing of neocortical progenitor cells during development.* EMBO J, 2017. **36**(17): p. 2642-2658.

180. Arnold, S.J., et al., *The T-box transcription factor Eomes/Tbr2 regulates neurogenesis in the cortical subventricular zone.* Genes Dev, 2008. **22**(18): p. 2479-84.

181. Sessa, A., et al., *Tbr2 directs conversion of radial glia into basal precursors and guides neuronal amplification by indirect neurogenesis in the developing neocortex.* Neuron, 2008. **60**(1): p. 56-69.

182. Kalebic, N., et al., *CRISPR/Cas9-induced disruption of gene expression in mouse embryonic brain and single neural stem cells in vivo.* EMBO Rep, 2016. **17**(3): p. 338-48.

183. Thakore, P.I., et al., *Editing the epigenome: technologies for programmable transcription and epigenetic modulation.* Nat Methods, 2016. **13**(2): p. 127-37.

184. Ray, S.K., et al., *CtBP and associated LSD1 are required for transcriptional activation by NeuroD1 in gastrointestinal endocrine cells.* Mol Cell Biol, 2014. **34**(12): p. 2308-17.

185. Luger, K., M.L. Dechassa, and D.J. Tremethick, *New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?* Nat Rev Mol Cell Biol, 2012. **13**(7): p. 436-47.

186. Macadangdang, B.R., et al., *Evolution of histone 2A for chromatin compaction in eukaryotes.* Elife, 2014. **3**.

187. Misteli, T., *The cell biology of genomes: bringing the double helix to life.* Cell, 2013. **152**(6): p. 1209-12.

188. Adam, R.C. and E. Fuchs, *The Yin and Yang of Chromatin Dynamics In Stem Cell Fate Selection.* Trends Genet, 2016. **32**(2): p. 89-100.

189. Clapier, C.R., et al., *Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes.* Nat Rev Mol Cell Biol, 2017. **18**(7): p. 407-422.

190. Stergachis, A.B., et al., *Developmental fate and cellular maturity encoded in human regulatory DNA landscapes.* Cell, 2013. **154**(4): p. 888-903.

191. Staahl, B.T. and G.R. Crabtree, *Creating a neural specific chromatin landscape by npBAF and nBAF complexes.* Curr Opin Neurobiol, 2013. **23**(6): p. 903-13.

192. Preissl, S., et al., *Single nucleus analysis of the chromatin landscape in mouse forebrain development.* bioRxiv, 2017.

193. Chapman, D.L., et al., *Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development.* Dev Dyn, 1996. **206**(4): p. 379-90.

194. Ivanova, N., et al., *Dissecting self-renewal in stem cells with RNA interference.* Nature, 2006. **442**(7102): p. 533-8.

195. Nichols, J. and A. Smith, *Pluripotency in the embryo and in culture.* Cold Spring Harb Perspect Biol, 2012. **4**(8): p. a008128.

196. Russell, R., et al., *A Dynamic Role of TBX3 in the Pluripotency Circuitry.* Stem Cell Reports, 2015. **5**(6): p. 1155-1170.

197. Waghray, A., et al., *Tbx3 Controls Dppa3 Levels and Exit from Pluripotency toward Mesoderm.* Stem Cell Reports, 2015. **5**(1): p. 97-110.

198. Bakker, M.L., et al., *Transcription factor Tbx3 is required for the specification of the atrioventricular conduction system.* Circ Res, 2008. **102**(11): p. 1340-9.

199. Bamshad, M., et al., *Mutations in human TBX3 alter limb, apocrine and genital development in ulnar-mammary syndrome.* Nat Genet, 1997. **16**(3): p. 311-5.

200. Hoogaars, W.M., et al., *Tbx3 controls the sinoatrial node gene program and imposes pacemaker function on the atria.* Genes Dev, 2007. **21**(9): p. 1098-112.

201. Lu, R., A. Yang, and Y. Jin, *Dual functions of T-box 3 (Tbx3) in the control of self-renewal and extraembryonic endoderm differentiation in mouse embryonic stem cells.* J Biol Chem, 2011. **286**(10): p. 8425-36.

202. Weidgang, C.E., et al., *TBX3 Directs Cell-Fate Decision toward Mesendoderm.* Stem Cell Reports, 2014. **2**(5): p. 747.

203. Kunasegaran, K., et al., *Transcriptional repressor Tbx3 is required for the hormone-sensing cell lineage in mammary epithelium.* PLoS One, 2014. **9**(10): p. e110191.

204. Han, J., et al., *Tbx3 improves the germ-line competency of induced pluripotent stem cells.* Nature, 2010. **463**(7284): p. 1096-100.

205. Hoogaars, W.M., et al., *The transcriptional repressor Tbx3 delineates the developing central conduction system of the heart.* Cardiovasc Res, 2004. **62**(3): p. 489-99.

206. Nishiyama, A., et al., *Systematic repression of transcription factors reveals limited patterns of gene expression changes in ES cells.* Sci Rep, 2013. **3**: p. 1390.

207. Hong, S., et al., *Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases.* Proc Natl Acad Sci U S A, 2007. **104**(47): p. 18439-44.

208. Kartikasari, A.E., et al., *The histone demethylase Jmjd3 sequentially associates with the transcription factors Tbx3 and Eomes to drive endoderm differentiation.* EMBO J, 2013. **32**(10): p. 1393-408.

209. Papaioannou, V.E., *The T-box gene family: emerging roles in development, stem cells and cancer.* Development, 2014. **141**(20): p. 3819-33.

210. Esmailpour, T. and T. Huang, *TBX3 promotes human embryonic stem cell proliferation and neuroepithelial differentiation in a differentiation stage-dependent manner.* Stem Cells, 2012. **30**(10): p. 2152-63.

211. Hendrich, B. and A. Bird, *Identification and characterization of a family of mammalian methyl-CpG binding proteins.* Mol Cell Biol, 1998. **18**(11): p. 6538-47.

212. Basta, J. and M. Rauchman, *The nucleosome remodeling and deacetylase complex in development and disease.* Transl Res, 2015. **165**(1): p. 36-47.

213. Lai, A.Y. and P.A. Wade, *Cancer biology and NuRD: a multifaceted chromatin remodelling complex.* Nat Rev Cancer, 2011. **11**(8): p. 588-96.

214. Wade, P.A., et al., *A multiple subunit Mi-2 histone deacetylase from Xenopus laevis cofractionates with an associated Snf2 superfamily ATPase.* Curr Biol, 1998. **8**(14): p. 843-6.

215. Tong, J.K., et al., *Chromatin deacetylation by an ATP-dependent nucleosome remodelling complex.* Nature, 1998. **395**(6705): p. 917-21.

216. Xue, Y., et al., *NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities.* Mol Cell, 1998. **2**(6): p. 851-61.

217. Zhang, Y., et al., *The dermatomyositis-specific autoantigen Mi2 is a component of a complex containing histone deacetylase and nucleosome remodeling activities.* Cell, 1998. **95**(2): p. 279-89.

218. Le Guezennec, X., et al., *MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties.* Mol Cell Biol, 2006. **26**(3): p. 843-51.

219. Dos Santos, R.L., et al., *MBD3/NuRD Facilitates Induction of Pluripotency in a Context-Dependent Manner.* Cell Stem Cell, 2014. **15**(3): p. 392.

220. Kaji, K., et al., *The NuRD component Mbd3 is required for pluripotency of embryonic stem cells.* Nat Cell Biol, 2006. **8**(3): p. 285-92.

221. Luo, M., et al., *NuRD blocks reprogramming of mouse somatic cells into pluripotent stem cells.* Stem Cells, 2013. **31**(7): p. 1278-86.

222. Rais, Y., et al., *Deterministic direct reprogramming of somatic cells to pluripotency.* Nature, 2013. **502**(7469): p. 65-70.

223. Saito, M. and F. Ishikawa, *The mCpG-binding domain of human MBD3 does not bind to mCpG but interacts with NuRD/Mi2 components HDAC1 and MTA2.* J Biol Chem, 2002. **277**(38): p. 35434-9.

224. Yildirim, O., et al., *Mbd3/NURD complex regulates expression of 5-hydroxymethylcytosine marked genes in embryonic stem cells.* Cell, 2011. **147**(7): p. 1498-510.

225. Baubec, T., et al., *Methylation-dependent and -independent genomic targeting principles of the MBD protein family.* Cell, 2013. **153**(2): p. 480-92.

226. Hashimoto, H., et al., *Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation.* Nucleic Acids Res, 2012. **40**(11): p. 4841-9.

227. Shimbo, T., et al., *MBD3 localizes at promoters, gene bodies and enhancers of active genes.* PLoS Genet, 2013. **9**(12): p. e1004028.

228. Spruijt, C.G., et al., *Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives.* Cell, 2013. **152**(5): p. 1146-59.

229. Reynolds, N., et al., *NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression.* EMBO J, 2012. **31**(3): p. 593-605.

230. Reynolds, N., A. O'Shaughnessy, and B. Hendrich, *Transcriptional repressors: multifaceted regulators of gene expression.* Development, 2013. **140**(3): p. 505-12.

231. Hu, G. and P.A. Wade, *NuRD and pluripotency: a complex balancing act.* Cell Stem Cell, 2012. **10**(5): p. 497-503.

232. Laugesen, A. and K. Helin, *Chromatin repressive complexes in stem cells, development, and cancer.* Cell Stem Cell, 2014. **14**(6): p. 735-51.

233. Bibel, M., et al., *Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells.* Nat Protoc, 2007. **2**(5): p. 1034-43.

234. Bibel, M., et al., *Differentiation of mouse embryonic stem cells into a defined neuronal lineage.* Nat Neurosci, 2004. **7**(9): p. 1003-9.

235. Bergsland, M., et al., *Sequentially acting Sox transcription factors in neural lineage development.* Genes Dev, 2011. **25**(23): p. 2453-64.

236. Bani-Yaghoub, M., et al., *Role of Sox2 in the development of the mouse neocortex.* Dev Biol, 2006. **295**(1): p. 52-66.

237. Bylund, M., et al., *Vertebrate neurogenesis is counteracted by Sox1-3 activity.* Nat Neurosci, 2003. **6**(11): p. 1162-8.

238. Cimadamore, F., et al., *SOX2-LIN28/let-7 pathway regulates proliferation and neurogenesis in neural precursors.* Proc Natl Acad Sci U S A, 2013. **110**(32): p. E3017-26.

239. Graham, V., et al., *SOX2 functions to maintain neural progenitor identity.* Neuron, 2003. **39**(5): p. 749-65.

240. Wegner, M. and C.C. Stolt, *From stem cells to neurons and glia: a Soxist's view of neural development.* Trends Neurosci, 2005. **28**(11): p. 583-8.

241. Amador-Arjona, A., et al., *SOX2 primes the epigenetic landscape in neural precursors enabling proper gene activation during hippocampal neurogenesis.* Proc Natl Acad Sci U S A, 2015. **112**(15): p. E1936-45.

242. Ng, S.Y., et al., *The long noncoding RNA RMST interacts with SOX2 to regulate neurogenesis.* Mol Cell, 2013. **51**(3): p. 349-59.

243. Evsen, L., et al., *Progression of neurogenesis in the inner ear requires inhibition of Sox2 transcription by neurogenin1 and neurod1.* J Neurosci, 2013. **33**(9): p. 3879-90.

244. Baek, J.H., et al., *Persistent and high levels of Hes1 expression regulate boundary formation in the developing central nervous system.* Development, 2006. **133**(13): p. 2467-76.

245. Hirata, H., et al., *Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop.* Science, 2002. **298**(5594): p. 840-3.

246. Kobayashi, T., et al., *The cyclic gene Hes1 contributes to diverse differentiation responses of embryonic stem cells.* Genes Dev, 2009. **23**(16): p. 1870-5.

247. Shimojo, H., T. Ohtsuka, and R. Kageyama, *Oscillations in notch signaling regulate maintenance of neural progenitors.* Neuron, 2008. **58**(1): p. 52-64.

248. Imayoshi, I., et al., *Oscillatory control of factors determining multipotency and fate in mouse neural progenitors.* Science, 2013. **342**(6163): p. 1203-8.

249. Laranjeiro, R. and D. Whitmore, *Transcription factors involved in retinogenesis are co-opted by the circadian clock following photoreceptor differentiation.* Development, 2014. **141**(13): p. 2644-56.

250. Castro, A.E., et al., *Expression and cellular localization of the transcription factor NeuroD1 in the developing and adult rat pineal gland.* J Pineal Res, 2015. **58**(4): p. 439-51.

251. Sato, A. and H. Takeda, *Neuronal subtypes are specified by the level of neurod expression in the zebrafish lateral line.* J Neurosci, 2013. **33**(2): p. 556-62.

252. Imayoshi, I. and R. Kageyama, *Oscillatory control of bHLH factors in neural progenitors.* Trends Neurosci, 2014. **37**(10): p. 531-8.

253. Ali, F., et al., *Cell cycle-regulated multi-site phosphorylation of Neurogenin 2 coordinates cell cycling with differentiation during neurogenesis.* Development, 2011. **138**(19): p. 4267-77.

254. Hindley, C., et al., *Post-translational modification of Ngn2 differentially affects transcription of distinct targets to regulate the balance between progenitor maintenance and differentiation.* Development, 2012. **139**(10): p. 1718-23.

255. Sriuranpong, V., et al., *Notch signaling induces rapid degradation of achaete-scute homolog 1.* Mol Cell Biol, 2002. **22**(9): p. 3129-39.

256. Vinals, F., et al., *BMP-2 decreases Mash1 stability by increasing Id1 expression.* EMBO J, 2004. **23**(17): p. 3527-37.

257. Bonev, B., P. Stanley, and N. Papalopulu, *MicroRNA-9 Modulates Hes1 ultradian oscillations by forming a double-negative feedback loop.* Cell Rep, 2012. **2**(1): p. 10-8.

258. Hu, W., J.R. Alvarez-Dominguez, and H.F. Lodish, *Regulation of mammalian cell differentiation by long non-coding RNAs.* EMBO Rep, 2012. **13**(11): p. 971-83.

259. Knuckles, P., et al., *Drosha regulates neurogenesis by controlling neurogenin 2 expression independent of microRNAs.* Nat Neurosci, 2012. **15**(7): p. 962-9.

260. Li, X. and P. Jin, *Roles of small regulatory RNAs in determining neuronal identity.* Nat Rev Neurosci, 2010. **11**(5): p. 329-38.

261. Morris, K.V. and J.S. Mattick, *The rise of regulatory RNA.* Nat Rev Genet, 2014. **15**(6): p. 423-37.

262. Tan, S.L., et al., *MicroRNA9 regulates neural stem cell differentiation by controlling Hes1 expression dynamics in the developing brain.* Genes Cells, 2012. **17**(12): p. 952-61.

263. Trott, J. and A. Martinez Arias, *Single cell lineage analysis of mouse embryonic stem cells at the exit from pluripotency.* Biol Open, 2013. **2**(10): p. 1049-56.

264. Deng, Q., et al., *Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells.* Science, 2014. **343**(6167): p. 193-6.

265. Battich, N., T. Stoeger, and L. Pelkmans, *Control of Transcript Variability in Single Mammalian Cells.* Cell, 2015. **163**(7): p. 1596-610.

266. Grun, D., et al., *Single-cell messenger RNA sequencing reveals rare intestinal cell types.* Nature, 2015. **525**(7568): p. 251-5.

267. Lee, J.K., et al., *Expression of neuroD/BETA2 in mitotic and postmitotic neuronal cells during the development of nervous system.* Dev Dyn, 2000. **217**(4): p. 361-7.

268. Schwab, M.H., et al., *Neuronal basic helix-loop-helix proteins (NEX, neuroD, NDRF): spatiotemporal expression and targeted disruption of the NEX gene in transgenic mice.* J Neurosci, 1998. **18**(4): p. 1408-18.

269. Nashun, B., P.W. Hill, and P. Hajkova, *Reprogramming of cell fate: epigenetic memory and the erasure of memories past.* EMBO J, 2015. **34**(10): p. 1296-308.

270. Jost, D., *Bifurcation in epigenetics: implications in development, proliferation, and diseases.* Phys Rev E Stat Nonlin Soft Matter Phys, 2014. **89**(1): p. 010701.

271. Steffen, P.A. and L. Ringrose, *What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory.* Nat Rev Mol Cell Biol, 2014. **15**(5): p. 340-56.

272. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.

273. Mohn, F. and D. Schubeler, *Genetics and epigenetics: stability and plasticity during cellular differentiation.* Trends Genet, 2009. **25**(3): p. 129-36.

274. Shipony, Z., et al., *Dynamic and static maintenance of epigenetic memory in pluripotent and somatic cells.* Nature, 2014. **513**(7516): p. 115-9.

275. Smith, Z.D., et al., *A unique regulatory phase of DNA methylation in the early mammalian embryo.* Nature, 2012. **484**(7394): p. 339-44.

276. Cambuli, F., et al., *Epigenetic memory of the first cell fate decision prevents complete ES cell reprogramming into trophoblast.* Nat Commun, 2014. **5**: p. 5538.

277. D'Urso, A. and J.H. Brickner, *Mechanisms of epigenetic memory.* Trends Genet, 2014. **30**(6): p. 230-6.

278. Ang, C.E. and M. Wernig, *Induced neuronal reprogramming.* J Comp Neurol, 2014. **522**(12): p. 2877-86.

279. Arlotta, P. and B. Berninger, *Brains in metamorphosis: reprogramming cell identity within the central nervous system.* Curr Opin Neurobiol, 2014. **27**: p. 208-14.

280. Graf, T., *Historical origins of transdifferentiation and reprogramming.* Cell Stem Cell, 2011. **9**(6): p. 504-16.

281. Vierbuchen, T., et al., *Direct conversion of fibroblasts to functional neurons by defined factors.* Nature, 2010. **463**(7284): p. 1035-41.

282. Falkner, S., et al., *Transplanted embryonic neurons integrate into adult neocortical circuits.* Nature, 2016. **539**(7628): p. 248-253.

283. Berninger, B., et al., *Functional properties of neurons derived from in vitro reprogrammed postnatal astroglia.* J Neurosci, 2007. **27**(32): p. 8654-64.

284. Chanda, S., et al., *Generation of induced neuronal cells by the single reprogramming factor ASCL1.* Stem Cell Reports, 2014. **3**(2): p. 282-96.

285. Wapinski, O.L., et al., *Hierarchical mechanisms for direct reprogramming of fibroblasts to neurons.* Cell, 2013. **155**(3): p. 621-35.

286. Treutlein, B., et al., *Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq.* Nature, 2016. **534**(7607): p. 391-5.

287. Gascon, S., et al., *Direct Neuronal Reprogramming: Achievements, Hurdles, and New Roads to Success.* Cell Stem Cell, 2017. **21**(1): p. 18-34.

288. Li, H. and G. Chen, *In Vivo Reprogramming for CNS Repair: Regenerating Neurons from Endogenous Glial Cells.* Neuron, 2016. **91**(4): p. 728-738.

289. Xu, J., Y. Du, and H. Deng, *Direct lineage reprogramming: strategies, mechanisms, and applications.* Cell Stem Cell, 2015. **16**(2): p. 119-34.

290. Mertens, J., et al., *Directly Reprogrammed Human Neurons Retain Aging-Associated Transcriptomic Signatures and Reveal Age-Related Nucleocytoplasmic Defects.* Cell Stem Cell, 2015. **17**(6): p. 705-18.

291. Babos, K. and J.K. Ichida, *Small Molecules Take a Big Step by Converting Fibroblasts into Neurons.* Cell Stem Cell, 2015. **17**(2): p. 127-9.

292. Biswas, D. and P. Jiang, *Chemically Induced Reprogramming of Somatic Cells to Pluripotent Stem Cells and Neural Cells.* Int J Mol Sci, 2016. **17**(2): p. 226.

293. Hu, W., et al., *Direct Conversion of Normal and Alzheimer's Disease Human Fibroblasts into Neuronal Cells by Small Molecules.* Cell Stem Cell, 2015. **17**(2): p. 204-12.

294. Zheng, J., et al., *A combination of small molecules directly reprograms mouse fibroblasts into neural stem cells.* Biochem Biophys Res Commun, 2016. **476**(1): p. 42-8.

295. Chatterjee, S., et al., *A novel activator of CBP/p300 acetyltransferases promotes neurogenesis and extends memory duration in adult mice.* J Neurosci, 2013. **33**(26): p. 10698-712.

296. Hsieh, J., et al., *Histone deacetylase inhibition-mediated neuronal differentiation of multipotent adult neural progenitor cells.* Proc Natl Acad Sci U S A, 2004. **101**(47): p. 16659-64.

297. Swaminathan, A., et al., *Modulation of neurogenesis by targeting epigenetic enzymes using small molecules: an overview*. ACS Chem Neurosci, 2014. **5**(12): p. 1164-77.

298. Weirauch, M.T. and T.R. Hughes, *A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution.* Subcell Biochem, 2011. **52**: p. 25-73.

299. Emerson, R.O. and J.H. Thomas, *Adaptive evolution in zinc finger transcription factors.* PLoS Genet, 2009. **5**(1): p. e1000325.

300. Laity, J.H., B.M. Lee, and P.E. Wright, *Zinc finger proteins: new insights into structural and functional diversity.* Curr Opin Struct Biol, 2001. **11**(1): p. 39-46.

301. Stubbs, L., Y. Sun, and D. Caetano-Anolles, *Function and Evolution of C2H2 Zinc Finger Arrays.* Subcell Biochem, 2011. **52**: p. 75-94.

302. Ali, R.G., H.M. Bellchambers, and R.M. Arkell, *Zinc fingers of the cerebellum (Zic): transcription factors and co-factors.* Int J Biochem Cell Biol, 2012. **44**(11): p. 2065-8.

303. Aruga, J., *The role of Zic genes in neural development.* Mol Cell Neurosci, 2004. **26**(2): p. 205-21.

304. Hegarty, S.V., A.M. Sullivan, and G.W. O'Keeffe, *Zeb2: A multifunctional regulator of nervous system development.* Prog Neurobiol, 2015. **132**: p. 81-95.

305. Eckler, M.J. and B. Chen, *Fez family transcription factors: controlling neurogenesis and cell fate in the developing mammalian nervous system.* Bioessays, 2014. **36**(8): p. 788-97.

306. Morris, D.R. and C.W. Levenson, *Zinc regulation of transcriptional activity during retinoic acid-induced neuronal differentiation.* J Nutr Biochem, 2013. **24**(11): p. 1940-4.

307. Yin, K.J., et al., *Krupple-like factors in the central nervous system: novel mediators in stroke.* Metab Brain Dis, 2015. **30**(2): p. 401-10.

308. Corsinotti, A., et al., *Global and stage specific patterns of Kruppel-associated-box zinc finger protein gene expression in murine early embryonic cells.* PLoS One, 2013. **8**(2): p. e56721.

309. Imbeault, M., P.Y. Helleboid, and D. Trono, *KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks.* Nature, 2017. **543**(7646): p. 550-554.

310. Liu, H., et al., *Deep vertebrate roots for mammalian zinc finger transcription factor subfamilies.* Genome Biol Evol, 2014. **6**(3): p. 510-25.

311. Urrutia, R., *KRAB-containing zinc-finger repressor proteins.* Genome Biol, 2003. **4**(10): p. 231.

312. Ecco, G., M. Imbeault, and D. Trono, *KRAB zinc finger proteins.* Development, 2017. **144**(15): p. 2719-2729.

313. Kauzlaric, A., et al., *The mouse genome displays highly dynamic populations of KRAB-zinc finger protein genes and related genetic units.* PLoS One, 2017. **12**(3): p. e0173746.

314. Lupo, A., et al., *KRAB-Zinc Finger Proteins: A Repressor Family Displaying Multiple Biological Functions.* Curr Genomics, 2013. **14**(4): p. 268-78.

315. Tan, X., et al., *Zfp819, a novel KRAB-zinc finger protein, interacts with KAP1 and functions in genomic integrity maintenance of mouse embryonic stem cells.* Stem Cell Res, 2013. **11**(3): p. 1045-59.

316. Karimi, M.M., et al., *DNA methylation and SETDB1/H3K9me3 regulate predominantly distinct sets of genes, retroelements, and chimeric transcripts in mESCs.* Cell Stem Cell, 2011. **8**(6): p. 676-87.

317. Quenneville, S., et al., *The KRAB-ZFP/KAP1 system contributes to the early embryonic establishment of site-specific DNA methylation patterns maintained during development.* Cell Rep, 2012. **2**(4): p. 766-73.

318. Rowe, H.M., et al., *De novo DNA methylation of endogenous retroviruses is shaped by KRAB-ZFPs/KAP1 and ESET.* Development, 2013. **140**(3): p. 519-29.

319. Rowe, H.M., et al., *KAP1 controls endogenous retroviruses in embryonic stem cells.* Nature, 2010. **463**(7278): p. 237-40.

320. Wolf, D. and S.P. Goff, *Embryonic stem cells use ZFP809 to silence retroviral DNAs.* Nature, 2009. **458**(7242): p. 1201-4.

321. Wolf, G., et al., *The KRAB zinc finger protein ZFP809 is required to initiate epigenetic silencing of endogenous retroviruses.* Genes Dev, 2015. **29**(5): p. 538-54.

322. Friedli, M. and D. Trono, *The developmental control of transposable elements and the evolution of higher species.* Annu Rev Cell Dev Biol, 2015. **31**: p. 429-51.

323. Wolf, G., D. Greenberg, and T.S. Macfarlan, *Spotting the enemy within: Targeted silencing of foreign DNA in mammalian genomes by the Kruppel-associated box zinc finger protein family.* Mob DNA, 2015. **6**: p. 17.

324. Thomas, J.H. and S. Schneider, *Coevolution of retroelements and tandem zinc finger genes.* Genome Res, 2011. **21**(11): p. 1800-12.

325. Erwin, J.A., M.C. Marchetto, and F.H. Gage, *Mobile DNA elements in the generation of diversity and complexity in the brain.* Nat Rev Neurosci, 2014. **15**(8): p. 497-506.

326. Richardson, S.R., S. Morell, and G.J. Faulkner, *L1 retrotransposons and somatic mosaicism in the brain.* Annu Rev Genet, 2014. **48**: p. 1-27.

327. Fasching, L., et al., *TRIM28 represses transcription of endogenous retroviruses in neural progenitor cells.* Cell Rep, 2015. **10**(1): p. 20-8.

328. Ryu, H., et al., *ESET/SETDB1 gene expression and histone H3 (K9) trimethylation in Huntington's disease.* Proc Natl Acad Sci U S A, 2006. **103**(50): p. 19176-81.

329. Tan, S.L., et al., *Essential roles of the histone methyltransferase ESET in the epigenetic control of neural progenitor cells during development.* Development, 2012. **139**(20): p. 3806-16.

330. Gao, L., et al., *Cloning and characterization of a novel human zinc finger gene, hKid3, from a C2H2-ZNF enriched human embryonic cDNA library.* Biochem Biophys Res Commun, 2004. **325**(4): p. 1145-52.

331. Watson, R.P., N. Tekki-Kessaris, and C.A. Boulter, *Characterisation, chromosomal localisation and expression of the mouse Kid3 gene.* Biochim Biophys Acta, 2000. **1490**(1-2): p. 153-8.

332. Kasaai, B., M.H. Gaumond, and P. Moffatt, *Regulation of the bone-restricted IFITM-like (Bril) gene transcription by Sp and Gli family members and CpG methylation.* J Biol Chem, 2013. **288**(19): p. 13278-94.

333. Takahashi, T., et al., *The effect of retinoic acid on a zinc finger transcription factor, AJ18, during differentiation of a rat clonal preosteoblastic cell line, ROB-C20, into osteoblasts.* Arch Oral Biol, 2008. **53**(1): p. 87-94.

334. Takahashi, T., et al., *Autoregulatory mechanism of Runx2 through the expression of transcription factors and bone matrix proteins in multipotential mesenchymal cell line, ROB-C26.* J Oral Sci, 2005. **47**(4): p. 199-207.

335. LePage, K.T., et al., *On the use of neuro-2a neuroblastoma cells versus intact neurons in primary culture for neurotoxicity studies.* Crit Rev Neurobiol, 2005. **17**(1): p. 27-50.

336. Olmsted, J.B., et al., *Isolation of microtubule protein from cultured mouse neuroblastoma cells.* Proc Natl Acad Sci U S A, 1970. **65**(1): p. 129-36.

337. Shastry, P., A. Basu, and M.S. Rajadhyaksha, *Neuroblastoma cell lines--a versatile in vitro model in neurobiology.* Int J Neurosci, 2001. **108**(1-2): p. 109-26.

338. Bellefroid, E.J., et al., *The evolutionarily conserved Kruppel-associated box domain defines a subfamily of eukaryotic multifingered proteins.* Proc Natl Acad Sci U S A, 1991. **88**(9): p. 3608-12.

339. Collins, T., J.R. Stone, and A.J. Williams, *All in the family: the BTB/POZ, KRAB, and SCAN domains.* Mol Cell Biol, 2001. **21**(11): p. 3609-15.

340. Friedman, J.R., et al., *KAP-1, a novel corepressor for the highly conserved KRAB repression domain.* Genes Dev, 1996. **10**(16): p. 2067-78.

341. Margolin, J.F., et al., *Kruppel-associated boxes are potent transcriptional repression domains.* Proc Natl Acad Sci U S A, 1994. **91**(10): p. 4509-13.

342. Vissing, H., et al., *Repression of transcriptional activity by heterologous KRAB domains present in zinc finger proteins.* FEBS Lett, 1995. **369**(2-3): p. 153-7.

343. Feschotte, C. and C. Gilbert, *Endogenous viruses: insights into viral evolution and impact on host biology.* Nat Rev Genet, 2012. **13**(4): p. 283-96.

344. Groner, A.C., et al., *KRAB-zinc finger proteins and KAP1 can mediate long-range transcriptional repression through heterochromatin spreading.* PLoS Genet, 2010. **6**(3): p. e1000869.

345. Vogel, M.J., et al., *Human heterochromatin proteins form large domains containing KRAB-ZNF genes.* Genome Res, 2006. **16**(12): p. 1493-504.

346. Schultz, D.C., et al., *SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins.* Genes Dev, 2002. **16**(8): p. 919-32.

347. Schultz, D.C., J.R. Friedman, and F.J. Rauscher, 3rd, *Targeting histone deacetylase complexes via KRAB-zinc finger proteins: the PHD and bromodomains of KAP-1 form a cooperative unit that recruits a novel isoform of the Mi-2alpha subunit of NuRD.* Genes Dev, 2001. **15**(4): p. 428-43.

348. Sripathy, S.P., J. Stevens, and D.C. Schultz, *The KAP1 corepressor functions to coordinate the assembly of de novo HP1-demarcated microenvironments of heterochromatin required for KRAB zinc finger protein-mediated transcriptional repression.* Mol Cell Biol, 2006. **26**(22): p. 8623-38.

349. Lange, A., et al., *Classical nuclear localization signals: definition, function, and interaction with importin alpha.* J Biol Chem, 2007. **282**(8): p. 5101-5.

350. Bruening, W., et al., *Identification of nuclear localization signals within the zinc fingers of the WT1 tumor suppressor gene product.* FEBS Lett, 1996. **393**(1): p. 41-7.

351. Choi, S., et al., *Structural basis for the selective nuclear import of the C2H2 zinc-finger protein Snail by importin beta.* Acta Crystallogr D Biol Crystallogr, 2014. **70**(Pt 4): p. 1050-60.

352. Fernandez-Martinez, J., et al., *Overlap of nuclear localisation signal and specific DNA-binding residues within the zinc finger domain of PacC.* J Mol Biol, 2003. **334**(4): p. 667-84.

353. Huang, S.M., et al., *Importin alpha1 is involved in the nuclear localization of Zac1 and the induction of p21WAF1/CIP1 by Zac1.* Biochem J, 2007. **402**(2): p. 359-66.

354. Ito, T., et al., *Role of zinc finger structure in nuclear localization of transcription factor Sp1.* Biochem Biophys Res Commun, 2009. **380**(1): p. 28-32.

355. Li, J.Z., et al., *Identification of a functional nuclear localization signal mediating nuclear import of the zinc finger transcription factor ZNF24.* PLoS One, 2013. **8**(11): p. e79910.

356. Quadrini, K.J. and J.J. Bieker, *Kruppel-like zinc fingers bind to nuclear import proteins and are required for efficient nuclear localization of erythroid Kruppel-like factor.* J Biol Chem, 2002. **277**(35): p. 32243-52.

357. Yamasaki, H., et al., *Zinc finger domain of Snail functions as a nuclear localization signal for importin beta-mediated nuclear import pathway.* Genes Cells, 2005. **10**(5): p. 455-64.

358. Jakel, S., et al., *The importin beta/importin 7 heterodimer is a functional nuclear import receptor for histone H1.* EMBO J, 1999. **18**(9): p. 2411-23.

359. Chook, Y.M. and G. Blobel, *Karyopherins and nuclear import.* Curr Opin Struct Biol, 2001. **11**(6): p. 703-15.

360. Kim, Y.H., M.E. Han, and S.O. Oh, *The molecular mechanism for nuclear transport and its application.* Anat Cell Biol, 2017. **50**(2): p. 77-85.

361. Knockenhauer, K.E. and T.U. Schwartz, *The Nuclear Pore Complex as a Flexible and Dynamic Gate.* Cell, 2016. **164**(6): p. 1162-1171.

362. Mosammaparast, N. and L.F. Pemberton, *Karyopherins: from nuclear-transport mediators to nuclear-function regulators.* Trends Cell Biol, 2004. **14**(10): p. 547-56.

363. Powers, M.A. and D.J. Forbes, *Cytosolic factors in nuclear transport: what's importin?* Cell, 1994. **79**(6): p. 931-4.

364. Brayer, K.J., S. Kulshreshtha, and D.J. Segal, *The protein-binding potential of C2H2 zinc finger domains.* Cell Biochem Biophys, 2008. **51**(1): p. 9-19.

365. Brayer, K.J. and D.J. Segal, *Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains.* Cell Biochem Biophys, 2008. **50**(3): p. 111-31.

366. Gamsjaeger, R., et al., *Sticky fingers: zinc-fingers as protein-recognition motifs.* Trends Biochem Sci, 2007. **32**(2): p. 63-70.

367. Hall, T.M., *Multiple modes of RNA recognition by zinc finger proteins.* Curr Opin Struct Biol, 2005. **15**(3): p. 367-73.

368. Leon, O. and M. Roth, *Zinc fingers: DNA binding and protein-protein interactions.* Biol Res, 2000. **33**(1): p. 21-30.

369. Matthews, J.M. and M. Sunde, *Zinc fingers--folds for many occasions.* IUBMB Life, 2002. **54**(6): p. 351-5.

370. Mackay, J.P. and M. Crossley, *Zinc fingers are sticking together.* Trends Biochem Sci, 1998. **23**(1): p. 1-4.

371. Najafabadi, H.S., et al., *C2H2 zinc finger proteins greatly expand the human regulatory lexicon.* Nat Biotechnol, 2015. **33**(5): p. 555-62.

372. Castel, S.E. and R.A. Martienssen, *RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond.* Nat Rev Genet, 2013. **14**(2): p. 100-12.

373. Riboni, L., et al., *A mediator role of ceramide in the regulation of neuroblastoma Neuro2a cell differentiation.* J Biol Chem, 1995. **270**(45): p. 26868-75.

374. Sato, C., T. Matsuda, and K. Kitajima, *Neuronal differentiation-dependent expression of the disialic acid epitope on CD166 and its involvement in neurite formation in Neuro2A cells.* J Biol Chem, 2002. **277**(47): p. 45299-305.

375. Tsuji, S., T. Yamashita, and Y. Nagai, *A novel, carbohydrate signal-mediated cell surface protein phosphorylation: ganglioside GQ1b stimulates ecto-protein kinase activity on the cell surface of a human neuroblastoma cell line, GOTO.* J Biochem, 1988. **104**(4): p. 498-503.

376. Mao, A.J., et al., *Neuronal differentiation and growth control of neuro-2a cells after retroviral gene delivery of connexin43.* J Biol Chem, 2000. **275**(44): p. 34407-14.

377. Moore, C.B., et al., *Short hairpin RNA (shRNA): design, delivery, and assessment of gene knockdown.* Methods Mol Biol, 2010. **629**: p. 141-58.

378. Ma, Q., et al., *Protease Omi facilitates neurite outgrowth in mouse neuroblastoma N2a cells by cleaving transcription factor E2F1.* Acta Pharmacol Sin, 2015. **36**(8): p. 966-75.

379. Ma'ayan, A., et al., *Neuro2A differentiation by Galphai/o pathway.* Sci Signal, 2009. **2**(54): p. cm1.

380. He, J.C., et al., *Role of the Go/i signaling network in the regulation of neurite outgrowth.* Can J Physiol Pharmacol, 2006. **84**(7): p. 687-94.

381. Jankowski, M.P., et al., *SRY-box containing gene 11 (Sox11) transcription factor is required for neuron survival and neurite growth.* Neuroscience, 2006. **143**(2): p. 501-14.

382. Pal, H.A., et al., *Self-assembling soft structures for intracellular NO release and promotion of neurite outgrowth.* Chem Sci, 2017. **8**(9): p. 6171-6175.

383. Rao, D.D., et al., *siRNA vs. shRNA: similarities and differences.* Adv Drug Deliv Rev, 2009. **61**(9): p. 746-59.

384. Singh, S., A.S. Narang, and R.I. Mahato, *Subcellular fate and off-target effects of siRNA, shRNA, and miRNA.* Pharm Res, 2011. **28**(12): p. 2996-3015.

385. Mockenhaupt, S., et al., *Alleviation of off-target effects from vector-encoded shRNAs via codelivered RNA decoys.* Proc Natl Acad Sci U S A, 2015. **112**(30): p. E4007-16.

386. Cohen-Fix, O., *Cell biology: Import and nuclear size.* Nature, 2010. **468**(7323): p. 513-6.

387. Levy, D.L. and R. Heald, *Nuclear size is regulated by importin alpha and Ntf2 in Xenopus.* Cell, 2010. **143**(2): p. 288-98.

388. van Zanten, M., et al., *Seed maturation in Arabidopsis thaliana is characterized by nuclear size reduction and increased chromatin condensation.* Proc Natl Acad Sci U S A, 2011. **108**(50): p. 20219-24.

389. Jevtic, P., et al., *Sizing and shaping the nucleus: mechanisms and significance.* Curr Opin Cell Biol, 2014. **28**: p. 16-27.

390. Webster, M., K.L. Witkin, and O. Cohen-Fix, *Sizing up the nucleus: nuclear shape, size and nuclear-envelope assembly.* J Cell Sci, 2009. **122**(Pt 10): p. 1477-86.

391. Mekhail, K. and D. Moazed, *The nuclear envelope in genome organization, expression and stability.* Nat Rev Mol Cell Biol, 2010. **11**(5): p. 317-28.

392. Raices, M. and M.A. D'Angelo, *Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions.* Nat Rev Mol Cell Biol, 2012. **13**(11): p. 687-99.

393. Witzgall, R., et al., *The Kruppel-associated box-A (KRAB-A) domain of zinc finger proteins mediates transcriptional repression.* Proc Natl Acad Sci U S A, 1994. **91**(10): p. 4514-8.

394. Meroni, G., *Genomics and evolution of the TRIM gene family.* Adv Exp Med Biol, 2012. **770**: p. 1-9.

395. Black, J.C. and J.R. Whetstine, *Chromatin landscape: methylation beyond transcription.* Epigenetics, 2011. **6**(1): p. 9-15.

396. Gaspar-Maia, A., et al., *Chd1 regulates open chromatin and pluripotency of embryonic stem cells.* Nature, 2009. **460**(7257): p. 863-8.

397. Ugarte, F., et al., *Progressive Chromatin Condensation and H3K9 Methylation Regulate the Differentiation of Embryonic and Hematopoietic Stem Cells.* Stem Cell Reports, 2015. **5**(5): p. 728-40.

398. Day, D.S., et al., *Estimating enrichment of repetitive elements from high-throughput sequence data.* Genome Biol, 2010. **11**(6): p. R69.

399. Reinert, K., et al., *Alignment of Next-Generation Sequencing Reads.* Annu Rev Genomics Hum Genet, 2015. **16**: p. 133-51.

400. Treangen, T.J. and S.L. Salzberg, *Repetitive DNA and next-generation sequencing: computational challenges and solutions.* Nat Rev Genet, 2011. **13**(1): p. 36-46.

401. Hommelsheim, C.M., et al., *PCR amplification of repetitive DNA: a limitation to genome editing technologies and many other applications.* Sci Rep, 2014. **4**: p. 5052.

402. Walker, J.A., et al., *Quantitative PCR for DNA identification based on genome-specific interspersed repetitive elements.* Genomics, 2004. **83**(3): p. 518-27.

403. Matsui, T., et al., *Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET.* Nature, 2010. **464**(7290): p. 927-31.

404. Rowe, H.M., et al., *TRIM28 repression of retrotransposon-based enhancers is necessary to preserve transcriptional dynamics in embryonic stem cells.* Genome Res, 2013. **23**(3): p. 452-61.

405. Turelli, P., et al., *Interplay of TRIM28 and DNA methylation in controlling human endogenous retroelements.* Genome Res, 2014. **24**(8): p. 1260-70.

406. Harte, P.J., et al., *Assignment of a novel bifurcated SET domain gene, SETDB1, to human chromosome band 1q21 by in situ hybridization and radiation hybrids.* Cytogenet Cell Genet, 1999. **84**(1-2): p. 83-6.

407. Dewannieux, M., et al., *Identification of autonomous IAP LTR retrotransposons mobile in mammalian cells.* Nat Genet, 2004. **36**(5): p. 534-9.

408. Ribet, D., et al., *An infectious progenitor for the murine IAP retrotransposon: emergence of an intracellular genetic parasite from an ancient retrovirus.* Genome Res, 2008. **18**(4): p. 597-609.

409. Li, J., et al., *Mouse endogenous retroviruses can trigger premature transcriptional termination at a distance.* Genome Res, 2012. **22**(5): p. 870-84.

410. Qin, C., et al., *Intracisternal A particle genes: Distribution in the mouse genome, active subtypes, and potential roles as species-specific mediators of susceptibility to cancer.* Mol Carcinog, 2010. **49**(1): p. 54-67.

411. Zhang, Y., et al., *Genome-wide assessments reveal extremely high levels of polymorphism of two active families of mouse endogenous retroviral elements.* PLoS Genet, 2008. **4**(2): p. e1000007.

412. Reichmann, J., et al., *Microarray analysis of LTR retrotransposon silencing identifies Hdac1 as a regulator of retrotransposon expression in mouse embryonic stem cells.* PLoS Comput Biol, 2012. **8**(4): p. e1002486.

413. Elsasser, S.J., et al., *Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells.* Nature, 2015. **522**(7555): p. 240-244.

414. Talbert, P.B. and S. Henikoff, *Histone variants on the move: substrates for chromatin dynamics.* Nat Rev Mol Cell Biol, 2017. **18**(2): p. 115-126.

415. Muotri, A.R., et al., *L1 retrotransposition in neurons is modulated by MeCP2.* Nature, 2010. **468**(7322): p. 443-6.

416. Muotri, A.R., et al., *Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition.* Nature, 2005. **435**(7044): p. 903-10.

417. Coufal, N.G., et al., *L1 retrotransposition in human neural progenitor cells.* Nature, 2009. **460**(7259): p. 1127-31.

418. Castro-Diaz, N., et al., *Evolutionally dynamic L1 regulation in embryonic stem cells.* Genes Dev, 2014. **28**(13): p. 1397-409.

419. Jacobs, F.M., et al., *An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons.* Nature, 2014. **516**(7530): p. 242-5.

420. Chuong, E.B., N.C. Elde, and C. Feschotte, *Regulatory activities of transposable elements: from conflicts to benefits.* Nat Rev Genet, 2017. **18**(2): p. 71-86.

421. Baek, S.T., et al., *Off-target effect of doublecortin family shRNA on neuronal migration associated with endogenous microRNA dysregulation.* Neuron, 2014. **82**(6): p. 1255-1262.

422. Svoboda, D.S., et al., *Induction of protein deletion through in utero electroporation to define deficits in neuronal migration in transgenic models.* J Vis Exp, 2015(95): p. 51983.

423. Blankenberg, F.G., *In vivo detection of apoptosis.* J Nucl Med, 2008. **49 Suppl 2**: p. 81S-95S.

424. Brundin, P., O. Isacson, and A. Bjorklund, *Monitoring of cell viability in suspensions of embryonic CNS tissue and its use as a criterion for intracerebral graft survival.* Brain Res, 1985. **331**(2): p. 251-9.

425. Cummings, B.S. and R.G. Schnellmann, *Measurement of cell death in mammalian cells.* Curr Protoc Pharmacol, 2004. **Chapter 12**: p. Unit 12 8.

426. Yamaguchi, Y. and M. Miura, *Programmed cell death in neurodevelopment.* Dev Cell, 2015. **32**(4): p. 478-90.

427. Kaushal, V., et al., *Caspase protocols in mice.* Methods Mol Biol, 2014. **1133**: p. 141-54.

428. Douville, R., et al., *Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis.* Ann Neurol, 2011. **69**(1): p. 141-51.

429. Gifford, W.D., S.L. Pfaff, and T.S. Macfarlan, *Transposable elements as genetic regulatory substrates in early development.* Trends Cell Biol, 2013. **23**(5): p. 218-26.

430. Thompson, P.J., T.S. Macfarlan, and M.C. Lorincz, *Long Terminal Repeats: From Parasitic Elements to Building Blocks of the Transcriptional Regulatory Repertoire.* Mol Cell, 2016. **62**(5): p. 766-76.

6 Materials

6.1 Biological material

Cells	Reference/Supplier
159.2 mESCs	Mohn et al., 2008 [1]
A2Lox.Cre mESCs	Iacovino et al., 2011 [2]
A2Lox.NeuroD1 mESCs	Personally generated
E.coli DH5alpha	Invitrogen
NIH/3T3	ATCC
MEF DR4	ATCC
MEF WT	Kind gift of Niehrs lab, IMB Mainz
Neuro2a	ATCC
Neuro2a.shNTC	Generated by Dr. Angela Garding
Neuro2a.shZfp354c	Generated by Dr. Angela Garding

6.2 Chemicals and biochemicals

Material	Supplier
3-(N-morpholino)propanesulfonic acid (MOPS), $C_7H_{15}NO_4S$	Sigma-Aldrich
30% Acrylamid/Bis-acrylamide Solution 19:1	Bio-Rad
Acetic acid 100%, CH_3COOH	AppliChem
Agarose Basic	AppliChem
Ammonium persulfate (APS), $(NH_4)_2S_2O_8$	Sigma-Aldrich
Anti-Flag M2 affinity gel	Sigma-Aldrich
Bacillol AF	NeoLab
Boric acid, H_3BO_3	Sigma-Aldrich
Bromphenol blue	Sigma-Aldrich
Calcium chloride, $CaCl_2$	Sigma-Aldrich
Chloroform	AppliChem
Chloroform:Isoamyl:Alcohol 24:1	Sigma-Aldrich
cOmplete, EDTA-free Protease Inhibitor Cocktail	Roche
Deoxycholic acid (DOC)	AppliChem
Di(N-succinimidyl) glutarate (DSG)	Sigma-Aldrich
Dimethylsulphoxide (DMSO)	Sigma-Aldrich
dNTPs solution mix, 8 mM each	New England Biolabs
DTT-solution (1M) molecular biology grade	AppliChem
EDTA	Sigma-Aldrich

EGTA	AppliChem
Ethanol, absolute ≥99.8% (GC)	Sigma-Aldrich
Formaldehyde solution for molecular biology, 36.5 - 38% in H_2O	Sigma-Aldrich
Glycerol	Sigma-Aldrich
Glycine BioUltra, for molecular biology, ≥99.0% (NT)	Sigma-Aldrich
Glycogen, from mussels	Roche
Guanidine hydrochloride, ≥99%, $NH_2C(=NH)NH_2 \cdot HCl$	Sigma-Aldrich
HEPES buffer solution (1M)	PAA Laboratories
HEPES, $C_8H_{18}N_2O_4S$	Sigma-Aldrich
Hygromycin B	Sigma-Aldrich
Hydrochloric acid 32%, p.A., HCl	AppliChem
IGEPAL CA-630 (NP-40)	Sigma-Aldrich
Isopropanol, 2-Propanol	Sigma-Aldrich
Lithium chloride, LiCl	Sigma-Aldrich
Magnesium chloride, $MgCl_2$	Sigma-Aldrich
Methanol, CHROMASOLV, for HPLC, ≥99.9%, CH_3OH	Sigma-Aldrich
Nitric acid, HNO_3	AppliChem
N-Lauroylsarcosine sodium salt, ≥94%	Sigma-Aldrich
Nonfat dried milk powder	AppliChem
NP-40	Sigma-Aldrich
Nuclease-free water	Life Technologies
NuPAGE LDS sample buffer (4x)	Life Technologies
Orange G loading buffer	Biozol
Paraformaldehyd reagent grade, crystalline	Sigma-Aldrich
Phenol solution, equilibrated with 10 mM Tris HCl, pH 8.0, 1 mM EDTA	Sigma-Aldrich
Phenol:Chloroform:Isoamyl Alcohol 25:24:1 saturated with 10 mM Tris, pH 8.0, 1 mM EDTA	Sigma-Aldrich
Phenylmethylsulfonyl fluoride (PMSF)	Sigma-Aldrich
PhosSTOP phosphatase inhibitor cocktail tablets	Roche
Potassium acetate, CH_3COOK	Sigma-Aldrich
Potassium chloride, KCl	Sigma-Aldrich
Potassium dihydrogen phosphate, KH_2PO_4	Sigma-Aldrich
Potassium hydroxide, KOH	Sigma-Aldrich
Protein A agarose	Merck Millipore
Protein G agarose	Merck Millipore
Quick Start Bradford 1x dye reagent	Bio-Rad
Rubidiumchloride, RbCl	Sigma-Aldrich

SDS solution (20%)	Sigma-Aldrich
Shandon Immu-Mount	Thermo Scientific
Sodium acetate, $C_2H_3NaO_2$	Sigma-Aldrich
Sodium azide, NaN_3	Sigma-Aldrich
Sodium bicarbonate, $NaHCO_3$	Sigma-Aldrich
Sodium chlorid, NaCl	Sigma-Aldrich
Sodium citrate tribasic dihydrate, $HOC(COONa)(CH_2COONa)_2 \cdot 2H_2O$	Sigma-Aldrich
Sodium deoxycholate, $C_{24}H_{39}NaO_4$	Sigma-Aldrich
Sodium hydroxide, NaOH	Sigma-Aldrich
Sodium orthovanadate, Na_3VO_4	Sigma-Aldrich
Sodium phosphate dibasic dehydrate, $Na_2HPO_4 \cdot 2H_2O$	Sigma-Aldrich
Sodium sulfate, Na_2SO_4	Sigma-Aldrich
SYBR Green PCR master mix	Applied Biosystems
SYBR Safe DNA gel stain	Invitrogen
Tetramethylethylendiamine (TEMED), $C_6H_{16}N_2$	AppliChem
Triton X-100	Sigma-Aldrich
Trizma base	Sigma-Aldrich
TRIzol reagent	Life Technologies
tRNA (*S. cerevisiae*)	Sigma-Aldrich
Tween 20	AppliChem
Water for molecular biology, DEPC-treated and sterile filtered	Sigma-Aldrich
β-Mercaptoethanol	Sigma-Aldrich

6.3 Enzymes

Material	Supplier
Antarctic Phosphatase	New England Biolabs
DNase I (RNase-free)	New England Biolabs
EcoRI	New England Biolabs
NheI	New England Biolabs
NotI	New England Biolabs
OneTaq DNA Polymerase	Promega
PacI	New England Biolabs
Proteinase K from *Engyodontium album*	Sigma-Aldrich
Q5 High-Fidelity DNA Polymerase	New England Biolabs
RNase A	Fermentas GmbH
T4 DNA Ligase	New England Biolabs
XhoI	New England Biolabs

6.4 Kits

Material	Supplier
Amaxa P3 Primary Cell 4D-NucleofectorTMX Kit L	Lonza
EndoFree Plasmid Maxi Kit	Qiagen
First Strand cDNA Synthesis Kit	Thermo Scientific
GeneJet Gel Extraction Kit	Thermo Scientific
GeneJET Plasmid Midiprep Kit	Thermo Scientific
GeneJET Plasmid Miniprep Kit	Thermo Scientific
MinElute PCR Purification Kit	Qiagen
Papain-based Neural Dissociation Kit	Milteney Biotec
Q5 Site-Directed Mutagenesis Kit	New England Biolabs
RNeasy Micro Kit	Qiagen
SuperScript VILO cDNA Synthesis Kit	Life Technologies
SuperSignal West Femto Maximum Sensitivity Substrate	Thermo Scientific

6.5 Standard DNA and protein weight markers

Material	Supplier
Color Prestained Protein Standard, broad range (11–245 kDa)	New England Biolabs
DNA ladder 1 kb	New England Biolabs
DNA ladder 100 bp	New England Biolabs

6.6 Buffers and solutions

All buffers were prepared and diluted with milliQ H_2O if not stated otherwise.

6.6.1 ChIP and FAIRE buffers

Paro Fix

50 mM HEPES/NaOH (pH 8.0)
1 mM EDTA
0.5 mM EGTA
100 mM NaCl
11% formaldehyde (add fresh before use)

DOC buffer

10 mM Tris (pH 8.0)
0.25 M LiCl
0.5% NP-40
0.5% DOC
1 mM EDTA

Elution buffer

1% SDS
0.1 M $NaHCO_3$

L1 buffer

0.05 M HEPES/KOH (pH 7.5)
140 mM NaCl
1 mM EDTA (pH 8.0)
10% glycerol
5% NP-40
0.25% Triton X-100

L2 buffer

0.2 M NaCl
1 mM EDTA (pH 8.0)
0.5 mM EGTA (pH 8.0)
10 mM Tris (pH 8.0)

L3 buffer

2 mM EDTA (pH 8.0)
5 mM EGTA (pH 8.0)
50 mM Tris (pH 8.0)
100 mM NaCl
0.1% Na-Deoxycholate
5 mg/ml N-Lauroylsarcosine sodium salt
1x protease inhibitor

6.6.2 Immunoprecipitation buffers

JS buffer

50 mM HEPES KOH (pH 7.5)
150 mM NaCl
5 mM EGTA (pH 8.0)
1.5 mM $MgCl_2$
1% glycerol
1% Triton X-100

NET buffer

50 mM Tris-HCl (pH 7.5)
150 mM NaCl
5 mM EGTA (pH 8.0)
0.1% Triton X-100

6.6.3 Buffers for SDS-PAGE and western blot

SDS stacking-buffer (SDS-PAGE)

0.1% SDS
0.5 M Tris
pH 6.8

For 1 l of SDS-stacking buffer 1.0 g SDS and 60.4 g Trizma base were dissolved in 800 ml H_2O. The pH was adjusted to 6.8 and volume subsequently filled up to 1 l with H_2O. The buffer was sterile filtered.

SDS resolving-buffer (SDS-PAGE)

0.1% SDS
1.5 M Tris
pH 8.8

For 1 l of SDS resolving-buffer 1.0 g SDS and 181.8 g Trizma base were dissolved in 800 ml H_2O. The pH was adjusted to 8.8 and volume subsequently filled up to 1 l with H_2O. The buffer was sterile filtered.

SDS-PAGE loading-buffer (6x)

0.5 M Tris/HCl (pH 6.8)
10% SDS
50% glycerol
20% 2-Mercaptoethanol
0.5% Bromphenol blue

The 6 x SDS-PAGE loading buffer was stored in 100 µl aliquots at -20 °C.

SDS running-buffer (10x, SDS-PAGE)

1% SDS
0.25 M Tris
1.92 M glycine

For 1 l of SDS resolving-buffer 10.0 g SDS, 30.3 g Trizma base and 144.1 g glycine were dissolved in 800 ml H_2O. The pH was adjusted to 8.45 and the volume subsequently filled up to 1 l with H_2O. The buffer was sterile filtered and diluted to 1x before usage.

Transfer buffer (10x, SDS-PAGE)

0.1% glycin
0.025 M Tris

For 1 l of transfer buffer 144.1 g Glycine and 30.3 g Trizma base were dissolved in 900 ml H_2O and the volume subsequently adjusted to 1 l with H_2O. The buffer was sterilize by autoclaving. Before usage the buffer was diluted to 1x including 10% v/v Mmethanol.

TBS-T buffer (10x)

0.25 M Tris
1.25 M NaCl
0.5 % HCl
0.1 % Tween

6.6.4 General buffers and solutions

PBS (10x)

1.4 M NaCl
27 mM KCl
83 mM Na_2HPO_4
15 mM KH_2PO_4
pH 6.8

For 1 l of 10x PBS, 80.0 g NaCl, 2.0 g KCl, 14.4 g Na_2HPO_4 x 2 H_2O and 2.4 g KH_2PO_4 were mixed in 800 ml H_2O. The pH was adjusted to 6.8 by adding HCl and the volume subsequently filled up to 1 l with H_2O. The buffer was sterile filtered and diluted to 1x before usage.

TAE buffer (50x)

2 M Tris acetate
50 mM EDTA

For 1 l of 50x TAE buffer 242.0 g Trizma base were dissolved in 500 ml H_2O. Subsequently 100 ml 0.5 M Na_2EDTA (pH 8.0) and 57.1 ml glacial acetic acid were added and the volume adjusted to 1 l with H_2O. The buffer was sterile filtered and diluted to 1x before usage.

TBE buffer (10x)

0.9 M Tris borate
20 mM EDTA

For 1 l of 50x TBE buffer 108.0 g Trizma base and 55.0 g boric acid were dissolved in 900 ml H_2O. Subsequently 40 ml 0.5 M Na_2EDTA (pH 8.0) were added and the volume adjusted to 1 l with H_2O. The buffer was sterile filtered and diluted to 1x before usage.

Triton extraction-buffer (TEB)

0.5% Triton X-100
2 mM phenylmethylsulfonyl fluoride (PMSF)
0.02% NaN_3
in PBS

RIPA buffer

50 mM Tris-HCl (pH 8.0)
150 mM NaCl
0.5% DOC
1% NP-40
0.1% SDS
5 mM EDTA
1% glycerol
2.5 mM $MgCl_2$
2 mM sodiumorthovanadate

Lysis buffer (gDNA isolation)

20 mM Tris pH8.0
4 mM EDTA
20 mM NaCl
1% SDS

TE buffer (pH 8.0)

10 mM Tris
1 mM EDTA

Borat buffer

150 mM boric acid/NaOH (pH 8.3)

TFB1 buffer (pH 5.8)

30 mM Potassium acetate
100 mM RbCl
10 mM $CaCl_2$ x 2 H_2O
50 mM $MnCl_2$
15% glycerol (v/v)

For 500 ml of TFB1 buffer 1.47 g potassium acetate, 6.05 g RbCl, 0.74 g $CaCl_2$ x 2 H_2O and 4.95 g $MnCl_2$ were dissolved in 400 ml H_2O. Subsequently, 15% glycerol (v/v) were added, the pH adjust to 5.8 with 0.2 M acetic acid and the volume adjusted to 500 ml with H_2O. The buffer was sterile filtered before usage.

TFB2 buffer (pH 6.5)

10 mM MOPS (PIPES)
75 mM $CaCl_2$ x 2 H_2O
10 mM RbCl
15% glycerol (v/v)

For 200 ml of TFB2 buffer 0.42 g MOPS, 2.20 g $CaCl_2$ x 2 H_2O and 0.24 g RbCl were dissolved in 120 ml H_2O. Subsequently 15% glycerol (v/v) were added, the pH adjust to 6.5 with KOH and the volume adjusted to 200 ml with H_2O. The buffer was sterile filtered before usage.

6.7 Cell culture material

6.7.1 Specific consumables

Material	Supplier
Aqua ad injectabilia	Braun
B-27 serum-free supplement (50x), liquid	Life Technologies
Boric acid, H_3BO_3	Riedel de Haen
BSA (Albumin from bovine serum)	Sigma-Aldrich
Dialyzed FBS	Sigma-Aldrich
DMSO	Sigma-Aldrich
Doxycycline hyclate (Dox)	Sigma-Aldrich
Dulbecco`s Modified Eagle Medium (DMEM)	Life Technologies

Dulbecco`s PBS (1x)	Life Technologies
F-12 nutrient mixture (Ham) 1x [+] L-Glutamine	Life Technologies
Fetal calf serum (FCS)	Life Technologies
G418 disulfate salt solution	Sigma-Aldrich
Gelatin	Sigma-Aldrich
Ham's F-12 nutrient mix	Life Technologies
Insulin	Sigma-Aldrich
Laminin from Engelbreth	Sigma-Aldrich
L-Arginine-0 (84 mg/mL)	Sigma-Aldrich
L-Arginine-10 (84 mg/mL)	Cambridge Isotope Laboratories
L-Glutamine, 200 mM solution	Life Technologies
LIF (Leukemia inhibitory factor)	own production
Lipofectamine 2000	Life Technologies
L-Lysine-0 (146 mg/mL)	Sigma-Aldrich
L-Lysine-8 (146 mg/mL)	Cambridge Isotope Laboratories
MEM non-essential amino acids solution, 100x	Life Technologies
Mitomycin C from *Streptomyces caespitosus*	Sigma-Aldrich
Neurobasal Medium (1x), liquid	Life Technologies
Opti-MEM I Reduced Serum Medium	Life Technologies
P3-Primary Cell Solution Box, small volume	Lonza
Poly-DL-ornithine hydrobromide (PORN)	Sigma-Aldrich
Progesteron	Sigma-Aldrich
Putrescine	Sigma-Aldrich
Retinoic acid	Sigma
SILAC DMEM (–Arg –Lys)	Life Technologies
Sodium selenite	Sigma-Aldrich
Transferrin	Sigma-Aldrich
Trypsin from bovine pancreas (powder)	Sigma-Aldrich
Trypsin-EDTA, 0.05%	Life Technologies

6.7.2 Media composition

ES cell culture medium	15% FCS
	10^3 U/ml LIF
	2 mM L-Glutamine
	1x Non-Essential Amino Acids
	5 µl/500 ml β-Mercaptoethanol
	in DMEM

CA culture medium	10% FCS 2 mM L-Glutamine 1x Non-Essential Amino Acids 5 µl/500 ml β-Mercaptoethanol (for differentiation add 5 µM retinoic acid) in DMEM
N2 medium	2 mM L-Glutamine 25 µg/ml Insuline 50 µg/ml Transferrin 20 nM Progesterone 100 nM Putrescine 30 nM sodium selenite 50 µg/ml BSA in DMEM:F-12 (1:1)
Complete medium	2% B27 supplement 2 mM L-Glutamine in Neurobasal medium
Neuro2a culture medium	10% FCS 2 mM L-Glutamine 1x Non-Essential Amino Acids in DMEM
Neuro2a differentiation medium	2% FCS 2 mM L-Glutamine 1x Non-Essential Amino Acids 20 µM retinoic acid in DMEM

Neuro2a SILAC medium	10% dialyzed FCS 2 mM L-Glutamine 1:2000 L-Arginine-(0/10) (84mg/mL) 1:2000 L-Lysine-(0/8) (146mg/mL) Light medium: Arg-0 + Lys-0 Heavy medium: Arg-10 + Lys-8 in SILAC DMEM (–Arg –Lys)

6.8 Antibodies

6.8.1 Primary antibodies

Antigen	Host	Supplier	Cat. Number	WB	IF	ChIP	IP
beta-actin	mouse	Santa Cruz	sc-47778	1:1,000	-	-	-
Flag-tag	mouse	Sigma	F1804	1:2,000	1:500	5 µg	4 µg
GFP	chicken	2BScientific	GFP-1020	-	1:500	-	-
H3K27ac	rabbit	Abcam	ab4729	-	-	2 µg	-
H3K27me3	rabbit	Active Motif	39155	-	-	2 µg	-
H3K4me1	rabbit	Abcam	ab8895	-	-	2 µg	-
HA-tag	rabbit	Abcam	ab9110	1:6,000	1:500	5 µg	4 µg
IgG	mouse	Santa Cruz	sc-2025	-	-	-	4 µg
IgG	rabbit	Santa Cruz	sc-2027	-	-	-	4 µg
KAP1	rabbit	Abcam	ab10483	1:500	-	5 µg	-
MBD3	rabbit	Abcam	ab16057	-	-	8 µg	-
MBD3	rabbit	Biomol	A302-528A	-	-	8 µg	-
NeuroD1	rabbit	Cell Signaling	#4373	-	-	1:50	-
PAX6	rabbit	Covance	PRB-278P-100	-	1:200	-	-
TBR1	rabbit	Abcam	ab31940	-	1:200	-	-
TBR2	rabbit	Abcam	ab23345	-	1:200	-	-
TBX3	goat	Santa Cruz	sc31657	-	-	8 µg	-
TUJ1	rabbit	Sigma	T2200	-	1:200	-	-

WB: dilution used for western blotting; IF: dilution used for immunocyto/histochemistry; ChIP: amount used for chromatin immunoprecipitation; IP: dilution used for immunoprecipitation.

Cell nuclei were visualized with Hoechst (200 µg/ml in H_2O) for immunocyto/histochemistry.

6.8.2 Secondary antibodies

All secondary antibodies were used according to the species in which the primary antibody was raised.

Antibody	Host	Supplier	Cat. Number	WB	IF
Alexa Fluor 488 goat anti-chicken IgG	goat	Invitrogen	A-11039	-	1:1,000
Alexa Fluor 488 goat anti-mouse IgG	goat	Invitrogen	A-11001	-	1:1,000
Alexa Fluor 568 goat anti-rabbit IgG	goat	Invitrogen	A-11011	-	1:1,000
anti-mouse IgG-HRP	goat	Santa Cruz	sc-2005	1:10,000	-
anti-rabbit IgG-HRP	goat	Santa Cruz	sc-2004	1:10,000	-

6.9 Oligonucleotides

6.9.1 Short hairpin RNA (shRNA)

shRNA	Sequence (5' – 3')
shZfp354c	ATGAGTAACTGTTGCATATAA
shNTC (scrambled)	CAACAAGATGAAGAGCACCA

6.9.2 Primers for PCR/cloning

Primer	Sequence (5' – 3')
A2Lox Loxin forward	ATACTTTCTCGGCAGGAGCA
A2Lox Loxin reverse	CTAGATCTCGAAGGATCTGGAG
ND1_3'_EcoRI	AATGAATTCCTAATCGTGAAAGATGGCATTAAG
ND1_5'_NotI	TAAGCGGCCGCATGACCAAATCATACAGCGAGAG
ND1_5'_XhoI	TAACTCGAGATGACCAAATCATACAGCGAGAG
Zfp354c_3'_EcoRI	TAAGAATTCTCAGAGGCTGCTGTTGAAGTTC
Zfp354c_5'_XhoI	TAACTCGAGATGGCCGTCGACCTGCTGGCCG
ZFP354c_KRABmut_F	GACTTCGAGATCTGGCCCG
ZFP354c_KRABmut_R	GGGCTCGGTGCCTCTAGC
ZFP354c_ZFCmut_F	CTGAAAGAACGGCTCTACAAGTG
ZFP354c_ZFCmut_R	GGGTTTCTGGCCGGGGTA

6.9.3 Primers for ChIP/FAIRE-qPCR

Primer	Sequence (5' – 3')
Apc2 F	AGGTTGCTGAGCCATGAACAGAG
Apc2 R	CAATCCTCCCAGGTGTTTGTGTC
Cmah F	CCACAACTGCAATGCAAACACAA
Cmah R	AAGTGCCAACACACCCTGTTGAA
Dll3 F	GGCCTCCCAGCTGTATGTAAATG
Dll3 R	GTTTGCCTCTCACACCTTCCTGA
Gapdh F	CTCTGCTCCTCCCTGTTCC
Gapdh R	TCCCTAGACCCGTACAGTGC
Hes6 F	AGGCAGCCTGTAGCCAATGAGAG
Hes6 R	CAAGAGAGAAGCCGGAGGTCACT
Hoxd4 F	CAGAGCCTTACTCGCCAGAGTCC
Hoxd4 R	GAGAAGCCTTTGCAAGTGACCAAA
IAP F	TGTGCCAGGCAGTAAACAAG
IAP R	ACCAATCACCACAGGTCACA
IAPEY3 F	GCCAAGTCACTGCCCATCC
IAPEY3 R	GGCGAAAGAAGAACAACGAACA
Intergenic F (ChIP normalization)	ATGCCCCTCAGCTATCACAC
Intergenic F (non-target)	GCTCCGGGTCCTATTCTTGT
Intergenic R (ChIP normalization)	GGACAGACATCTGCCAAGGT
Intergenic R (non-target)	TCTTGGTTTCCAGGAGATGC
LINE1 F	TTTGGGACACAATGAAAGCA
LINE1 R	CTGCCGTCTACTCCTCTTGG
MMERVK10C F	TATCGCCTCAGGGTTAATGC
MMERVK10C R	TGGATGCCACACAACTCATT
MMVL30 F	TGGGGGCTCGTCCGGGAT
MMVL30 R	ATTACCAAGCGACAGAACTTACC
Ncam1 Enhancer F	CCTCCAGCCCATAGTCCTAATGC
Ncam1 Enhancer R	AATGGGACTGATTTTCCCCCTTT
NeuroD1 F	TCACCCCTCCCCAGAACTTTCCT
NeuroD1 R	AATAGGCAGGTCACGTGGTTCCC
NeuroD4 Enhancer R	TGTCTCCTAGCAGAGGAGCCTGA
NeuroD4 Enhancer F	GAAAAGGAAGGCAACAGGCAGAT
Nhlh1 Enhancer F	TCCTCTTTGTGTCCTCCCATCTG
Nhlh1 Enhancer R	AAGGAGGCAGGAAGTAGGGTTCC
Nhlh2 F	GCCAGAGGAATTTTCCATGTTCA
Nhlh2 R	TGAAAGCAGCCCATGATAGAAGC
Numbl F	CTCCCTCCTGTTCCTGACAACCT
Numbl R	AAGTGGGCATTGGGTTAGACCAG
ORR1A0 F	GTGATGGTTTGTATATCCTTGG
ORR1A0 R	CTAGTGGAAGACTGACTTCCAGG
Pcsk2 F	CATAAGCATCACGTCTCCCTTGG

Pcsk2 R	CCACCTGCGATGGCTATTAGAGA
RLTR45 F	TGCTTTTCCGACATGGTAAT
RLTR45 R	AGTAACCCTGACCTGCTCCT
SINE B1 F	GGTGTGGTGGCGCACACC
SINE B1 R	CCTGGCTGTCCTGGAGCTC

6.9.4 Primers for RT-qPCR

Primer	Sequence (5' – 3')
Ctcf F	CACACACACAGGTACTCGTCCTCA
Ctcf R	CCACTGGTCACAAAGGCCATATC
Dll3 F	AATGGGGGCAGCTGTAGTGAA
Dll3 R	CACATCGAAGCCCGTAGAATCC
Hes6 F	GAAGCCCCTGGTGGAGAAGAA
Hes6 R	GTTCTCTAGCTTGGCCTGCACCT
Klf4 F	CGGGAAGGGAGAAGACACTGC
Klf4 R	GAGAGAGTTCCTCACGCCAACG
MapT F	CTGCTGTAGCCGCTTCGTTCTC
MapT R	GGAAATGACGAGAAGAAAGCCAAG
Myt1 F	CAGCCCAAAGTTCAAGACAAGTGA
Myt1 R	TCCTGATTTAGCTGCTTGATCTCC
Nanog F	AGCCTCCAGCAGATGCAAGAA
Nanog R	TTGCACTTCATCCTTTGGTTTTGA
Ncam1 F	AAGCACACAGAGCCCAACGAG
Ncam1 R	AAGGCAGCATGTCCTCGACAGT
NeuroD1 ectopic (HA-tag) F	ATGGGGTACCCATACGATGTTCC
NeuroD1 ectopic (HA-tag) R	TTTTCTGAGCCTCGAAGATGTCG
NeuroD1 endogenous F	AAGAAGTGCTAAGGCAACGCAAT
NeuroD1 endogenous R	AACCATGCAGCTTCATCAATTTTT
NeuroD1 total F	AGCTCCCACGTCTTCCACGTC
NeuroD1 total R	GGCTTTCAAAGAAGGGCTCCAG
NeuroD4 F	AAGAACTACTCGCGGGAGCTGAC
NeuroD4 R	TCCATCCAGGATTGTGTGTTGAC
Nhlh1 F	CCTCTCACTTGCCAGCTTGGAT
Nhlh1 R	TCAAAGTTCCATGGTCAAGATTCC
Nhlh2 F	GGCTGCTTTTAATATTTGCCAGGA
Nhlh2 R	CGGACTCAGCATCATTTTGGAG
Oct4 F	CGTGAAGTTGGAGAAGGTGGAAC
Oct4 R	TCTAGCTCCTTCTGCAGGGCTTT
Rpl19 F	CTCGTTGCCGGAAAAACA
Rpl19 R	TCATCCAGGTCACCTTCTCA
Sox2 F	CCATGGGCTCTGTGGTCAAGT
Sox2 R	GGAGTGGGAGGAAGAGGTAACCA

Syp F	CTATGGGCAGCAAGGCTACGG
Syp R	GGGGTCTTCGTGGGCTTCACT
Tbp F	GTTTCTGCGGTCGCGTCATTT
Tbp R	TGGGTTATCTTCACACACCATGAA
Tubb3 F	GGAGCGCATCAGCGTATACTACA
Tubb3 R	GGTTCCAAGTCCACCAGAATGG
vGlut2 F	GCTGGAAAATCCCTCGGACAGA
vGlut2 R	ATGGTCTCTCGGTTGTCCTGCTT

6.10 Plasmids

Plasmid	Source
12AB4JNC_Neurod1_pMK-RQ	GeneArt custom synthesis
p2loxEGFP	Kind gift of Dr. Michael Kyba
p2loxHA-Nanog	Cloned by Sandra Schick
p2loxHA-NeuroD1	Personally cloned
pCIDRE	Kind gift of Dr. Johan Holmberg
pCIDRE-NeuroD1	Personally cloned
pCIDRE-Zfp354c	Personally cloned
pCIDRE-Zfp354c-KRAB-mutant	Personally cloned
pGOv4_Zfp354c_rev	GeneOracle custom synthesis
pHAFlagAgo1	Addgene
pHAFlagNeuroD1	Personally cloned
pHAFlagZfp354c	Cloned by Dr. Angela Garding
pHAFlagZfp354c-KRAB-mutant	Personally cloned
pHAFlagZfp354c-ZFC-mutant	Personally cloned
pH1.shNTC	Cloned by Dr. Angela Garding
pH1.shZfp354c	Cloned by Dr. Angela Garding
pRNAT-H1.1/Hygro	GenScript

Personally cloned plasmids: I cloned these plasmids during my PhD studies using standard molecular biology methods. Detailed information about the cloning procedure is described under 7.2. Plasmid maps with relevant sequence information are available electronically (.gb and .xdna files) in the host laboratory at the IMB.

7 Methods

7.1 Nucleic acid methods

7.1.1 Preparation of genomic DNA (gDNA)

Collected cells (around $3x10^6$) were resuspended in 300 µl TE buffer, complemented with 300 µl gDNA lysis buffer containing 40 µg/ml Proteinase K and incubated at 55 °C o/n. The sample was vortexed after addition of 600 µl phenol and centrifuged for 10 min at RT with 16,000 x g. The upper aqueous phase was transferred to a new Eppendorf tube. After the addition of 600 µl chloroform:isoamyl alcohol (24:1) the sample was again centrifuged for 10 min at RT with 16,000 x g. The upper aqueous phase was transferred to a new Eppendorf tube and DNA was precipitated with 1.2 ml 100% ethanol (4 °C) containing 75 mM sodium acetate at 4 °C for 2 h with 16,000 x g. The pellet was washed with 250 µl 70% ethanol (4 °C) and centrifuged for 15 min with 16,000 x g at 4 °C. The supernatant was discarded and the pellet air dried for 20 min. The DNA pellet was resuspended in 60 µl TE buffer containing 20 µg/ml RNase A and incubated for 30 min at 37 °C. After quantification of DNA concentration (see 7.1.3) the gDNA was stored at -20 °C.

7.1.2 Extraction of RNA

The total RNA of cell populations was isolated for RNA-sequencing or used for the synthesis of cDNA for RT-qPCR. The RNA of *in vitro* cultured cells was isolated from pellets containing $6x10^5$ to $1x10^7$ cells using TRIzol reagent (Ambion) following manufacturer's instructions. The RNA of FAC-sorted cortical cells ($5x10^5$) was extracted using the RNeasy Micro Kit (Qiagen) following manufacturer's instructions which includes an on column DNase I treatment. RNA samples were stored in RNase free water at -80 °C.

7.1.3 Quantification of DNA and RNA in solution

The measurement of DNA and RNA quantities was carried out using a NanoDrop2000 UV-Vis Spectrophotometer (Thermo Scientific). The quantification principle is based on the absorption maximum of nucleic acids at 260 nm due to the aromatic ring structure of their purine and pyrimidine bases in the nucleotides and is measured as the optical density (OD) according to the law of Lambert-Beer. The absorption maximum of proteins is 280 nm due to the absorbance of aromatic amino acids. Therefore, the purity of nucleic acid solutions can be measured by the ratio of OD260/OD280. For pure DNA solutions a ratio of 1.8 is expected whereas a ratio of 2.0 is anticipated for pure RNA solutions. If nucleic acid solution are contaminated with residual phenol or proteins the OD260/OD280 ratio will decrease. Therefore, DNA samples with OD260/OD280 < 1.6 as well as RNA samples with an OD260/OD280 < 1.8 were discarded from further analysis in this thesis. The measurement of 1 µl nucleic acid sample was carried out after blanking the NanoDrop2000 with 1 µl of pure solvent.

7.1.4 Reverse transcription of complementary DNA (cDNA)

The RNA of *in vitro* cultured cells was reverse transcribed into complementary DNA (cDNA) with a random hexamer primer using a First Strand cDNA Synthesis Kit (Thermo Scientific) following manufacturer's

instructions, including a digestion of genomic DNA with DNase I. Each reaction was carried out on 1 µg RNA and the final cDNA sample was filled up to 100 µl with TE buffer.
The total RNA of FAC-sorted cortical cells was reverse transcribed using the SuperScript VILO cDNA Synthesis Kit (Life Technologies) according to manufacturer's instructions.

7.1.5 Agarose gel electrophoresis and analysis of DNA

DNA is negatively charged in neutral buffer solutions due to its phosphate backbone. Therefore, the nucleic acids molecules are able to migrate in an electrical field from the cathode to the positively charged anode through an agarose matrix and are separated by their size over time. The concentration of agarose in the gel matrix determines the size of the pores and hence migration speed of the molecules according to their size. In general, 1.2% agarose gels were prepared with 1x TAE buffer. The agarose was dissolved in TAE buffer by heating in a microwave and SYBR Safe DNA Gel Stain was added to the solution (1:20,000) after cooling to 60 °C. Gels were poured using the Sub-Cell GT Agarose Gel Electrophoresis System (Bio-Rad) and the desired amount of DNA was supplemented with 1x Orange G Loading Buffer before loading in the gel pockets. Either 8 µl of 1 kb or 100 bp DNA ladder were loaded to aid visualisation of DNA size. In general, electrophoresis was run at 100 V for around 40 min and images were taken using a ChemiDoc XRS (Bio-Rad) and ImageLab software (Bio-Rad).

7.1.6 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) was first used in 1986 by Mullis and colleagues as a method to enzymatically amplify a DNA fragment of choice *in vitro* [3]. Due to the use of heat stable DNA polymerases in combination with small, sequence specific oligonucleotides (15-25 bp "primers"), the DNA template can be exponentially amplified. In this thesis, PCR was used for four purposes: First, it was used to amplify gene sequences to be cloned into several target vectors. Second, the Loxin PCR according to Iacovino et al. [2] enabled the screening for positive transgenic A2lox mESCs clones at the level of genomic DNA. Third, quantitative PCR (qPCR) was used to analyse the DNA content received by ChIP- and FAIRE assays. Fourth, reverse transcription quantitative PCR (RT-qPCR) allowed to quantify the expression of certain genes in RNA samples from various conditions.

7.1.6.1 Oligonucleotide design

Oligonucleotide primers for qPCR were designed to amplify a product of 75 - 125 bp. Their size ranged from 21 bp to 24 bp to guarantee specificity, the melting temperature was set between 63 °C and 67 °C.
All primers for qPCR were designed with Primer3 software (http://frodo.wi.mit.edu/primer3/). The genomic target sequences for the design of qPCR primers were identified by using the *Mus musculus* genome assembly (NCBI37/mm9) provided by the National Center for Biotechnology Information (NCBI). The target sequence was derived at around 200 bp upstream of the target gene promoter using the USCS browser (http://genome.ucsc.edu). For enhancer regions, the genomic sequence inside this *cis*-regulatory element was directly used for primer design.

Primers for RT-qPCR were designed with the Universal Probe Library Assay Design Center from Roche Applied Systems (https://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp) using an intron spanning strategy. Therefore, an amplicon from genomic DNA contamination which contains large intron sequences (at least 1 kb) would hardly be amplified.
Primers for cloning purposes were design complementary to the endpoints of the desired sequence to be amplified. The specific palindromes of the restriction sites used in cloning were added to the 5' end (forward primer), or 3' end (reverse primer) of the primer and were additionally elongated by 3 nucleotides (TAA) to allow efficient cleavage close to the end of DNA sequences.

7.1.6.2 Loxin PCR

Iacovino and colleagues published transgenic mESCs (A2lox.Cre) which harbour a doxycycline inducible gene expression locus [2]. These cells enable in combination with the p2lox plasmid the insertion of a nucleotide sequence into this locus by recombinase-mediated cassette exchange (RMCE). The Loxin PCR was used to screen for transgenic mESC clones harbouring a correct insertion of the desired sequence into the genomic DNA of A2lox mESCs. Each 25 µl PCR reaction was set up as follows:

5.0 µl	DNA (20 ng/µl)
5.0 µl	5x OneTaq standard buffer
0.5 µl	10 mM dNTPs
1.0 µl	A2Lox Loxin forward primer (5 µM)
1.0 µl	A2Lox Loxin reverse primer (5 µM)
0.125 µl	OneTaq polymerase
12.375 µl	H_2O

The PCR reaction was performed with the Professional Trio Thermocycler system (Biometra) with the following protocol:

Initial denaturation	94 °C	180 sec	1x
Denaturation	94 °C	30 sec	
Annealing	57 °C	50 sec	35x
Elongation	68 °C	40 sec	
Final elongation	68 °C	300 sec	1x
Storage	4 °C		

7.1.6.3 Quantitative PCR analysis (qPCR)

The amount of DNA fragments in a sample was measured by quantitative PCR (qPCR) analysis [4, 5]. The quantity of DNA was detected by measuring the fluorescence of SYBR Green after each PCR cycle. This dye binds to double-stranded DNA and subsequently is able to emit fluorescence whereas in the unbound state this is not possible. Thus, the intensity of the fluorescence signal is proportional to the amount of

double-stranded DNA in the sample and allows the quantification of PCR products in real time: The intersection of the amplification curve with a defined threshold gives a so called CT (cycle threshold) value. This value defines the number of cycles required for the fluorescent signal to cross the threshold (defined by a calibration curve) and therefore exceeding background fluorescence levels. The ViiA 7 Real-Time PCR System (Applied Biosystems) was used for all measurements in a 384 well plate format. The qPCR analysis in this thesis was performed for ChIP and FAIRE samples as well as their respective inputs and each 10 µl qPCR reaction was set up in technical duplicates as follows:

0.3 µl	ChIP/FAIRE/Input DNA
2.5 µl	SYBR Green PCR master mix
0.5 µl	Forward primer (5 µM)
0.5 µl	Reverse primer (5 µM)
2.5 µl	TE buffer
3.7 µl	H_2O

The qPCR program was run as follows:

qPCR amplification	Initial denaturation	95 °C	600 sec	1x
	Denaturation	95 °C	15 sec	40x
	Annealing/Elongation	60 °C	60 sec	
Melting curve analysis	Denaturation	95 °C	15 sec	1x
	Annealing	60 °C	15 sec	
	Heating	up to 95 °C	0.05 °C/sec	
	Storage	4 °C		

The fluorescence of SYBR Green was measured at the end of each elongation phase. Additionally, a melting curve analysis was performed after the final qPCR cycle.
For the quantitative analysis, the CT values for each target region in ChIP and FAIRE samples were normalized to the corresponding CT values of the input sample to calculate the fold enrichment above input. For additionally calculating the fold enrichment above background, a second normalization to an intergenic control region (chr12:49000661 to 49000802) was performed.

7.1.6.4 Reverse transcription quantitative PCR (RT-qPCR)

Quantitative PCR was combined with reverse transcription of RNA into cDNA (see 7.1.4) to quantify expression of genes. The ViiA 7 Real-Time PCR System (Applied Biosystems) was used for all measurements with the same parameters as described above. Each 10 µl reaction consisted of the following components and was carried out in technical duplicates:

1.0 µl	cDNA
2.5 µl	SYBR Green PCR master mix
0.5 µl	forward primer (5 µM)
0.5 µl	reverse primer (5 µM)
2.5 µl	TE buffer
3.0 µl	H_2O

For the analysis of expression levels, all target gene CT values were normalized to the expression of a housekeeping gene (e.g. *Rpl19*) for calculation of relative mRNA levels in each sample.

7.2 Cloning of plasmids

7.2.1 Generation and transformation of chemically competent *E. coli*

The bacteria (*E.coli* DH5alpha) were streaked on a LB-plate (without antibiotics) and incubated at 37 °C o/n. Afterwards, 5 ml of LB-media (without antibiotics) were inoculated with a single bacterial colony and incubated at 37 °C o/n with 220 rpm. On the next day 200 ml of LB-media (without antibiotics) were inoculated with 3 ml of the o/n culture and again incubated 37 °C with 220 rpm until the OD_{600} was around 0.5 - 0.6. Afterwards, the bacteria were incubated on ice for 15 min and centrifuged for 10 min at 4 °C with 1,000 x g. The supernatant was discarded and the pellet resuspended with 80 ml TFB1 buffer (4 °C). The suspension as kept on ice for 5 min and centrifuged for 5 min at 4 °C with 1,000 x g. Again, the supernatant was discarded and the pellet resuspended with 2 ml ice-cold TFB 2 buffer. This suspension was kept on ice for 15 min, aliquoted (50 µl) in pre-cooled Eppendorf tubes (-20 °C) and immediately snap frozen in liquid nitrogen. The chemically competent bacteria were stored at -80 °C.
For the transformation with a purified plasmid one vial of competent bacteria was thawed on ice. A volume of 1 µl with 10 ng/µl plasmid DNA was added to the cells and mixed by tapping gently. The vial was incubated for 30 min on ice and heat shocked at 42 °C for 30 sec w/o mixing. Afterwards, the vial was removed from the heat block and quickly placed on ice. Pre-warmed LB medium (250 µl) was added and the cells were incubated at 37 °C for 1 h at 225 rpm. The mixture was plated onto a LB plate with an appropriate antibiotic and incubated inverted at 37 °C o/n.

7.2.2 Small scale plasmid-preparation (MiniPrep)

The GeneJET Plasmid Miniprep Kit (Thermo Scientific) was used for the preparation of small amounts of plasmids (1-5 µg). Therefore, a single bacterial colony containing the desired plasmid was inoculated into 5 ml of LB medium supplemented with an appropriate antibiotic (100 µg/ml Ampicillin or 30 µg/ml Kanamycin) and incubated at 37 °C o/n while shaking with 220 rpm. The next day, 1.5 ml of the culture were transferred to an Eppendorf tube and centrifuged at 5,000 x g for 5 min at RT. The pellet was resuspended in 250 µL of resuspension solution and further processed according to manufacturer's instructions. The final elution step was carried out in 30 µl elution buffer. One µg of the plasmid was subjected to restriction digest (see 7.2.4) and analysis on an agarose gel (see 7.1.5) to confirm correct insert and backbone size. In some

cases Sanger sequencing was carried out at GATC according to their conditions (see 7.2.7) to confirm sequence identity.

7.2.3 Medium scale plasmid-preparation (MidiPrep)

The GeneJET Plasmid Midiprep Kit (Thermo Scientific) was used for the preparation of medium amounts of plasmids (100 - 400 µg). For this purpose, a single bacterial colony containing (the sequence verified) plasmid of interest was inoculated into 3 ml of LB medium supplemented with an appropriate antibiotic (100 µg/ml Ampicillin or 30 µg/ml Kanamycin) and incubated at 37 °C for 6 h while shaking with 220 rpm to generate a pre-culture. The same pre-culture was transferred to 100 ml of LB media with an appropriate antibiotic and incubated o/n at 37 °C while shaking with 220 rpm. The next day, the entire culture was centrifuged at 4500 x g for 10 min at RT. The pellet was resuspended in 2 mL of resuspension solution and further processed according to manufacturer's instructions. The final elution step was carried out in 300 µl of elution buffer. In some cases Sanger sequencing was carried out at GATC according to their conditions (see 7.2.7) to confirm sequence identity.

7.2.4 Restriction digestion of DNA

Restriction endonucleases cleave double-stranded DNA at their enzyme specific recognition sites, therefore leaving either blunt ends or single stranded overhangs (sticky ends), which can be used for subsequent ligation reactions. A typical restriction digest reaction was set up as follows:

2 µl	10x Reaction buffer (NEB)
1 µl	NEB restriction enzyme (5-10 U)
1 µg	DNA template
x µl	H_2O (add up to 20 µl)

For a double digest 1 µl restriction enzyme each was used while decreasing the amount of H_2O accordingly. The reaction mix was incubated at appropriate temperature (usually 37 °C; specific enzyme needs were set according to manufacturer's recommendations) for 1 h and the restriction enzymes inactivated for 20 min at 80 °C. If an enzyme exhibited star-activity, the incubation time at 37 °C was decreased. The generated 5' ends were dephosphorylated using antarctic phosphatase. Therefore, 1/10 of 10x antarctic phosphatase reaction buffer was added to the restriction mix as well as 1 µl antarctic phosphatase (5 units). The solution was incubated for 15 min at 37 °C and heat inactivated afterwards for 5 min at 70 °C. The cleavage efficiency was examined by agarose gel electrophoresis (see 7.1.5).

7.2.5 Extraction of DNA from agarose gels

DNA derived by PCR or restriction digest was extracted from agarose gels (see 7.1.5) using the GeneJet Gel Extraction Kit. For this purpose, each DNA band was cut out from an agarose gel with a sterile scalpel on the Bio-Rad Chemidoc under UV light and transferred to an Eppendorf tube. The weight of the agarose gel piece was analysed and the DNA extracted according the manufacturer's instructions.

7.2.6 Ligation of DNA fragments

Two DNA fragments digested with compatible restriction enzymes can be fused together by using DNA ligases which enzymatically catalyse the formation of phosphodiester bonds between DNA fragments. For each ligation reaction 50 ng of appropriately restricted vector backbone was used. The needed amount of restricted insert was calculated as follows: Insert [ng] = (3 x 50 ng vector x length of insert [bp]) / length of vector [bp]. The pipetting scheme for a 20 µl ligation reaction was as follows:

50 ng	Digested vector backbone DNA
x ng	Digested insert DNA
2 µl	T4 10x ligation buffer
2 µl	T4 DNA ligase
x µl	H_2O (fill up to 20 µl)

T4 DNA ligase was added last and the reaction gently mixed. One additional reaction without the insert was used as religation control. The ligation reaction was incubated o/n at 16 °C. The transformation of chemically competent *E. coli* cells with 9 µl of the incubated ligation reaction was carried out as described in 7.2.1.

7.2.7 Sanger DNA sequencing

For sequencing, 20 µl of 30 - 50 ng/µl purified plasmid and 20 µl of a 10 µM sequencing primer were send to GATC Biotech, Konstanz. Sequencing results were received as .ab1 files and the nucleotide sequences were analysed using Finch TV software (geospiza). The alignment to the reference sequences was performed using SerialCloner (Serial Basics).

7.2.8 Cloning strategies

For all cloning steps the destination vector as well as the insert were digested, ligated, purified and validated as described above.

7.2.8.1 Custom synthesis of genes

The murine coding sequence (CDS) of *NeuroD1* (NM_010894.2) was custom synthesized at GeneArt and provided in the 12AB4JNC_Neurod1_pMK-RQ plasmid. The CDS is flanked by 5'-PacI/SbfI and NheI/AvrII-3' restriction sites.

The murine CDS of *Zfp354c* (NM_013922) was custom synthesized at GeneOracle. The codon usage was adapted at GeneArt to the codon bias of *Mus musculus* genes. In addition, regions of very high (> 80%) or very low (< 30%) GC content have been avoided if possible. During the optimization process the following cis-acting sequence motifs were avoided where applicable: internal TATA-boxes and ribosomal entry sites; AT-rich or GC-rich sequence stretches; RNA instability motifs; repeat sequences and RNA secondary structures; (cryptic) splice donor and acceptor sites in higher eukaryotes. The *Zfp354c* CDS was provided in the pGOv4_Zfp354c_rev plasmid and is flanked by 5'-PacI/SbfI and NheI/AvrII-3' restriction sites.

7.2.8.2 Cloning of p2loxHA-NeuroD1

The *NeuroD1* CDS was cloned into the p2loxHA-Nanog destination vector to generate an N-terminal HA-tagged NeuroD1 construct for the generation of A2lox.NeuroD1 mESCs according to Iacovino et al. [2]. The *NeuroD1* CDS was subcloned from 12AB4JNC_Neurod1_pMK-RQ into the p2loxHA-Nanog destination vector by replacing the *Nanog* CDS using a PacI/NheI double digest in NEB buffer 1.

7.2.8.3 Cloning of pCIDRE-NeuroD1

The *NeuroD1* CDS was cloned into the pCIDRE destination vector under the control of a CAG promoter to generate a plasmid for the overexpression of NeuroD1 along with dsRed (separated by an IRES element). Therefore, the *NeuroD1* CDS was amplified by PCR from 12AB4JNC_Neurod1_pMK-RQ using Q5 polymerase and the ND1_5'_XhoI and ND1_3'_EcoRI primers according to manufacturer's instructions. The pCIDRE destination vector as well as the PCR amplicon were double digested with XhoI/EcoRI in NEB buffer 4 for subcloning.

7.2.8.4 Cloning of pCIDRE-Zfp354c

The CDS of *Zfp354c* was cloned into the pCIDRE destination vector under the control of a CAG promoter to generate a plasmid for the overexpression of NeuroD1 along with dsRed (separated by an IRES element). The *Zfp354c* CDS was amplified by PCR from pGOv4_Zfp354c_rev using Q5 polymerase and the Zfp354c_5'_XhoI and Zfp354c_3'_EcoRI primers according to manufacturer's instructions. The pCIDRE destination vector as well as the PCR amplicon were double digested with XhoI/EcoRI in NEB buffer 4 for subcloning.

7.2.8.5 Cloning of pHAFlagNeuroD1

The *NeuroD1* CDS was cloned into the pHAFlagAgo1 vector to generate an N-terminal HA-Flag tagged NeuroD1 expressed from a CAG promoter. Therefore, the *NeuroD1* CDS was PCR amplified from 12AB4JNC_Neurod1_pMK-RQ using Q5 polymerase chemistry and the ND1_5'_NotI and ND1_3'_EcoRI primers according to manufacturer's instructions. The PCR product was subcloned into the pHAFlagAgo1 destination vector be replacing the *Ago1* CDS using a NotI/EcoRI double digest in NEB buffer 3.

7.2.9 Site directed mutagenesis for the generation of Zfp354c domain mutants

The deletions of the KRAB-domain as well as the zinc-finger cluster (ZFC) of *Zfp354c* were incorporated into the pHAFlagZfp354c and the pCIDRE-Zfp354c plasmids through the use of specifically designed forward and reverse primers in combination with the Q5 Site-Directed Mutagenesis Kit (NEB). The primers omitting the DNA sequence to be deleted were designed with NEBase changer and the optimal annealing temperature (68 °C) was calculated using the same tool. The site directed mutagenesis (exponential amplification, KLD reaction and transformation) was carried out according to manufacturer's instructions. For the generation of the KRAB-mutant (*Zfp354c* without KRAB domain) the ZFP354c_KRABmut_F and ZFP354c_KRABmut_R primers where used. The ZFP354c_ZFCmut_F and ZFP354c_ZFCmut_R primers were used to generate the ZFC mutant (*Zfp354c* without zinc-fingers).

7.3 Biochemical protein methods

7.3.1 Protein extraction from cells

Whole cell protein extracts were generated from cell pellets containing $1x10^6$ to $1x10^7$ cells. The pellet was resuspended in 100 - 200 µl RIPA buffer and incubated for 1 h at 4 °C while gently rotating. Afterwards, the solution was centrifuged for 10 min at 4 °C with 10,000 x g to remove cellular debris. The cleared supernatant was transferred to a new Eppendorf tube and the protein concentration was measured using the Bradford assay. Aliquots were stored at -80 °C for long term.

7.3.2 Histone extraction from cells

Acid extraction was used to extract highly basic proteins such as histones from nuclei. The pellet from $1x10^6$ to $1x10^7$ cells was resuspended in 1 ml Triton Extraction-buffer (TEB). The cells were lyzed for 10 min with gentle stirring at 4 °C. The solution was centrifuged at 2,000 x g for 10 min at 4 °C and the supernatant was discarded. Cells were washed in half the volume of TEB and centrifuged as before. The pellet was resuspended in 0.2 N HCl with ¼ of the original TEB volume. The acid extraction of histones was performed o/n at 4 °C. Afterwards, samples were centrifuged at 2,000 x g for 10 min at 4 °C and the supernatant was transferred to a new Eppendorf tube. The protein concentration was measured using the Bradford assay, aliquots were stored at -80 °C.

7.3.3 Bradford assay

The Bradford assay was used to quantify protein/histone extracts. Therefore, 1 ml of Bradford reagent (Bio-Rad) was supplemented with 4 µl of protein/histone extract and vortexed directly. After 10 min of incubation in the dark, the absorbance was measured at 600 nm (OD_{600}) with a Biophotometer plus (Eppendorf). A mixture of 1 ml Bradford reagent supplemented with 4 µl of the pure solvent was used for blanking the measurement.

7.3.4 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were separated according to their molecular weight by discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE) using Tris-Glycine gels [6]. The SDS-PAGE was performed using the Mini Protean Tetra Cell system (Bio-Rad). The separation as well as the stacking gel were casted as follows:

Separation gel:

Gel %	H_2O [ml]	30% Acrylamid [ml]	SDS resolving-buffer [ml]	10% APS [µl]	TEMED [µl]	Final volume [ml]
6	5.3	2.0	2.6	100	8	10
8	4.6	2.7	2.6	100	6	10
10	4.0	3.3	2.6	100	4	10
12	3.3	4.0	2.6	100	4	10

Stacking gel:

Gel %	H_2O [ml]	30% Acrylamid [µl]	SDS stacking-buffer [µl]	10% APS [µl]	TEMED [µl]	Final volume [ml]
5	2.7	670	540	40	4	4

For casting the gels, APS and TEMED were added to the mixture after all other components to start the polymerisation. The freshly prepared separating gel solution was poured into the gel chamber and isopropanol was added as a thin layer on top in order to guarantee a defined and smooth boundary of the gel. After polymerisation (approximately 30-45 min), the isopropanol was removed and the freshly prepared stacking gel solution was poured onto the polymerised separating gel. The desired comb was successively inserted. Gels were immediately used after polymerisation or stored in a wet chamber for up to 1 week at 4 °C if needed. Equal amounts of protein/histone extract were boiled in 1x SDS-PAGE loading-buffer for 6 min at 95 °C before loading on the gel. As a marker for protein size, 8 µl of Color Prestained Protein Standard (New England Biolabs) were used. Electrophoresis was performed in 1 x SDS running-buffer at 80 V for about 20 min to allow the samples to run slowly into the stacking gel. Afterwards, the voltage was increased to 100 - 130 V and electrophoresis was performed until the loading front reached the bottom of the separation gel.

7.3.5 Western blot (semi-dry)

The negatively charged SDS-protein complexes were transferred from the separation gel onto a Polyvinylidene fluoride (PVDF) membrane by western blotting using the Trans-Blot Turbo Transfer System (Bio-Rad). The PVDF membrane was activated in methanol for 30 sec, quickly rinsed in water and briefly stored in 1x transfer buffer. Four Whatman blotting papers were cut to 11.0 x 15.5 cm and soaked in 1x transfer buffer. The sandwich was built in the transfer cassette in the following order without air bubbles between the layers: Anode (+), 2 Whatman blotting papers, PVDF membrane, separation gel, 2 Whatman blotting papers, cathode (-). The transfer was performed using the StandardSD protocol for 45 min with constant 25 V.

7.3.6 Immunodetection of proteins of interest on western blot

The immobilised proteins on the PVDF membrane were visualised by immunodetection after the Semi-Dry transfer. The detection is based on the primary antibody directed against the target epitope, whereas the secondary antibody is coupled to a horse-radish peroxidase and targets the Fc part of the primary antibody. This coupled enzyme can catalyse a chemiluminescent reaction if the substrate is supplied and enables the visualisation of the target epitope on the membrane.
The membrane was first blocked for 1 h in 1x TBS-T containing 5% milk powder at RT while gently shaking. The primary antibody was diluted in 1x TBS-T containing 5% milk powder and the incubation was carried out at 4 °C o/n. The membrane was washed three times with 1x TBS-T for 10 min each and incubated afterwards for 1 h at RT with the secondary antibody dilluted in 1x TBS-T containing 5% milk powder. The membrane was again washed three times with 1x TBS-T for 10 min each. After washing, the membrane was developed using a Chemiluminescence Kit (SuperSignal West Femto Maximum Sensitivity Substrate, Thermo Scientific). The chemiluminescence was detected using the ChemiDoc XRS (Bio-Rad) and ImageLab software (Bio-Rad).

7.4 Cell culture

7.4.1 Culture of cells for maintenance and differentiation

In general, cells were cultured in an incubator with 88% relative air humidity, 37 °C and 7% CO_2. All cell culture media as well as Trypsin-EDTA were warmed at 37 °C in a water bath prior usage. Cells were counted using a TC 10 automated cell counter (Bio-Rad). Only sterile 1x PBS (Life Technologies) was used for all cell culture purposes.

7.4.1.1 Freezing and thawing of cells

All cells were frozen in their normal culture medium containing 10% DMSO for long term storage in liquid nitrogen (3 vials per 80% confluent 10 cm plate). For this purpose, cells were trypsinised and resupended in 1 ml of freezing medium per cryovial. These vials were immediately transferred to pre-cooled (4 °C) cryo boxes filled with isopropanol and stored at -80 °C. ES cells should ideally not be kept at this temperature longer than 48 h and needed to be transferred to -150 °C within this period. Longer storage at -80 °C will slowly lead to a loss of pluripotency and their differentiation potential will be severely impaired. All other cell types were transferred to -150 °C within one week.
Each cryovial was quickly thawed by resuspending the cells in 10 ml of pre-warmed appropriate culture medium and centrifuged for 5 min with 180 x g at RT. The cell pellet was resuspended in 8 ml of appropriate culture medium and distributed to a cell culture dish.

7.4.1.2 Neuronal differentiation of murine embryonic stem cells (mESCs)

The murine embryonic stem cells (mECs) were cultured and differentiated into glutamatergic neurons according to the work of Bibel and colleagues [7, 8].

To maintain pluripotency state, mESCs were cultured in 8 ml ES medium in 10 cm cell culture dishes on feeders (inactivated mouse embryonic fibroblasts (MEFs)). For the culture of mESCs without feeders the cell culture dishes were coated with 0.2% gelatin for 15 min at 37 °C to supply an adherent substrate. The cells were splitted every two days 1:5 to gradually remove feeders for further experiments and the ES medium was changed daily. For splitting, the ES medium was first removed and the culture dish was carefully washed with 10 ml 1x PBS. The dish was incubated with 1.2 ml Trypsin-EDTA at 37 °C until the cells detached from the surface. The trypsin was consecutively inactivated by addition of 8 ml ES medium. The cells were pelleted by centrifugation at 180 x g for 5 min at RT. The cell pellet was resuspended in fresh ES medium and transferred to new 10 cm dishes by splitting 1:5.

After five rounds of splitting on gelatin coated plates and cultured in ES medium, the mESCs were first cultured in CA medium. Only 25 ml pipettes have been used for culturing cellular aggregates (CAs) to ensure integrity of the formed CAs. Non-adhesive bacterial dishes (Greiner) were used to plate $3.5x10^6$ cells in 15 ml CA medium and the CA medium was changed after 2 days (2d). At d4 of CA stage as well as on d6, the medium was changed to CA medium supplemented with 5 µM retinoic acid (RA) to trigger the cell fate commitment towards neuronal lineage. PORN (poly-DL-ornithine) and laminin coated plates for neuronal differentiation of the derived progenitors were prepared at d6 and d7 of the CA stage. For this purpose, a 0.5 mg/ml PORN stock was prepared in borat buffer and 10 cm culture dishes were coated with 5 ml of a 1:5 dilution in aqua ad injectabilia at 37 °C o/n. Afterwards, the PORN-coated plates were gently washed three times with aqua ad injectabilia. Afterwards, 6 ml 1x PBS supplemented with 15 µl laminin were added per plate and incubated o/n at 37 °C. CAs were washed at d8 two times with 1x PBS and dissociated with 0.5 ml Trypsin (0.05% (w/v) in 0.05% (w/v) EDTA/PBS) per 2 CA plates during 3 min incubation at 37 °C while briefly shaking every 30 sec. Trypsin was inactivated by addition of 10 ml CA medium and the cells were vigorously resuspended to remove clumps. Centrifugation was carried out for 5 min with 180 x g, the pellet was resuspended in 10 ml N2 medium and filtered through a nylon filter drop by drop. The laminin/PBS solution was removed from the coated culture dishes and $9.5x10^6$ cells were plated per 10 cm dish in 6 ml N2 medium. The N2 medium was renewed after 2 h and also after 24 h as a modification of the original protocol from Bibel et al. [7]. The entire medium was changed to complete medium on d2 and renewed on d4 and d6 of neurogenesis. At d8, 4 ml complete medium were added on top of the residual medium. The differentiation was carried on until d10.

7.4.1.3 Induced terminal neuronal differentiation (iTN) of A2lox.NeuroD1 mESCs

In this thesis a protocol was developed to directly induce neuronal differentiation of mESCs without a cellular aggregate state as described in 7.4.1.2. For this purpose, A2lox.NeuroD1 mESCs were directly plated on PORN/Laminin-coated tissue culture dishes (see 7.4.1.2) and cultivated for 2 days in N2 media supplemented with doxycycline (e.g. 500 ng/ml). The medium was renewed after 24 h. At d3 of differentiation the medium was changed to complete medium and cells were cultured until d7. The complete medium was changed every two days during this process.

7.4.1.4 Neuro2a culture and differentiation with retinoic acid

The neuronal cell line Neuro2a was cultured in 10 cm dishes with 8 ml Neuro2a media and splitted every two days to keep the confluency below 80%. For differentiation, $1.2x10^6$ cells were seeded per 10 cm dish and neuritogenesis was induced with 8 ml differentiation media. The differentiation media was changed every two days.

7.4.1.5 NIH/3T3 culture

The NIH/3T3 fibroblast cell line was cultured in 10 cm dishes with 8 ml Neuro2a media and splitted every two days to keep the confluency below 80%. For experiments, $1.5x10^6$ cells were seeded per 10cm dish and cultured for 48 h.

7.4.1.6 Collection of cell pellets

Cell pellets were collected for the isolation of DNA, RNA or proteins. Therefore, cells were washed twice with cold 1x PBS, scraped in 1 ml cold 1x PBS and transferred to an Eppendorf tube. They were centrifuged for 3 min with 2,000 x g at 4 °C. The supernatant was removed, the pellet was snap frozen in liquid nitrogen and stored at -80 °C.

7.4.1.7 Inactivation of MEFs with Mitomycin C

Mouse embryonic fibroblasts (MEFs) were mitotically inactivated by Mitomycin C treatment, which leads to an inter-strand cross-linking of DNA. Freshly isolated MEFs were pre-cultured with CA medium in 10 cm dishes for 48 h and propagated in 15 cm tissue culture dishes until passage 5 (P5) by splitting according to density 1:3 or 1:5. The stock solution of Mitomycin C (0.5 mg/ml) was sterile filtered, protected from light and added to a confluent 15 cm dish to a final concentration of 10 µg/ml in CA medium. The treatment was carried out for 3 h in the cell culture incubator at 37 °C. Afterwards, the plates were washed twice with 15 ml 1x PBS and cultured for 2 additional days with 15 ml CA medium to check for successful inactivation of mitosis. These cells were called “feeders” from then onwards and stored in CA medium containing 10% DMSO at -150 °C. Feeders have been prepared with wildtype as well as DR4 MEFs.

7.4.2 Transient transfection of cells with Lipofectamine 2000

The transient transfection of Neuro2a or NIH/3T3 cells with Lipofectamine 2000 was performed directly after plating. In total 10 µg of the desired plasmid were diluted in 200 µl Opti-MEM, supplemented with 40 µl of Lipofectamine 2000 (1 µg/µl) and briefly mixed by pipetting. After incubating the transfection mix for 5 min at RT, the solution was added drop wise to the freshly seeded cells in the according culture medium. The medium was changed after 24 h.

7.4.3 Electroporation of A2lox.Cre mESCs for RMCE and clone selection

The Amaxa P3 Primary Cell 4D-Nucleofector TMX Kit was used for a highly efficient transfer of plasmid DNA into mESCs to generate transgenic inducible cell lines. The A2lox.Cre mESCs cells have been induced with 1 µg/ml Doxycycline 24 h prior to the electroporation in order to express the Cre recombinase to ensure

the recombinase mediated cassette exchange (RMCE) after transfection of the p2lox plasmid. The induced cells were trypsinised and $5x10^6$ cells were centrifuged in an Eppendorf tube for 5 min with 180 x g at RT. The transfection solution was prepared by adding 18 µl Supplement to 82 µl of Nucleofector solution at RT. After the addition of 4 µg of the desired p2lox plasmid to the transfection solution, the previously pelleted cells were resuspended in 100 µl of the prepared solution and mixed gently. The mix was transferred into a nucleocuvette without air bubbles and the cells were electroporated with the CG-104 program on the Amaxa Nucleofector 4D (Lonza). The cells were gently resuspended in 500 µl ES medium and plated on a 10 cm cell culture dish coated with 0.2% gelatin.

The medium was changed after 24 h to ES medium containing 2 µg/ml Puromycin (G418). The selection was carried out for 5 d while the Puromycin ES medium was replaced every day. The remaining cells were cultured in ES medium w/o Puromycin after this selection step until single clonal colonies were visible on the dish. Single clones were picked under sterile conditions and splitted by trypsinisation into two 48 well plates coated with 0.2% gelatin. After 2-4 days of incubation one plate was used for genomic DNA preparation (see 7.1.1) and Loxin PCR (see 7.1.6.2) whereas the second plate served for maintenance of the clones. Positive clones from the Loxin PCR screen were cultured for one week on feeders and stored as stocks at -150 °C.

7.4.4 Cellular assays

7.4.4.1 Immunocytochemistry of *in vitro* cultured cells

Cells were grown on coverslips and fixed with 4% paraformaldehyde in 1x PBS for 15 min at RT. After three washes with 1x PBS for 5 min the cells were permeabilized and simultaneously blocked with 10% goat serum and 5% FBS in 1x PBS supplemented with 0.2% Triton X-100 for 1 h at RT. Subsequently, the samples were incubated with primary antibodies o/n at 4 °C. After three washes with 1x PBS the coverslips were incubated with fluorochrome-labeled secondary antibody for 1 h at RT including a counterstain with Hoechst (200 µg/ml) for visualization of nuclei. The coverslips where washed three times with 1x PBS and mounted with Immu-Mount on microscope slides. Images where taken with a confocal laser-scanning microscope (Leica TCS SP5) and data were processed with ImageJ software.

7.4.4.2 Fluorescence-activated cell sorting (FACS) of *in vitro* cultured cells

The FACS (Fluorescence-activated Cell Sorting) technology provides a powerful tool to specifically select for a population of living single cells expressing fluorescencent protein(s) [9-11]. FACS of *in vitro* cultured cells was performed in collaboration with the Cytometry Core Facility of the IMB and therefore is only briefly mentioned here. In general, cells were transfected (see 7.4.2) with the desired plasmid co-expressing a red fluorescent protein (RFP) and cultured for 48 h under normal conditions. They were trypsinized, resupended in 500 µl cold 1x PBS and filtered through a 35 µm mesh to receive a single cell suspension. The single cell suspension was sorted with an Aria III SORP (Becton Dickinson). The first gate was set to distinguish cellular events from debris based on forward (FSC-A) and side scatter area (SSC-A), which are proportional to the size of a particle and intracellular granularity, respectively. This first population was additionally subgated to

separate single cells from doublets comparing the SSC-A with the side scatter width. These single cellular events were further gated based on RFP intensity (yellow-green laser: 561 nm) to separate transfected from non-transfected cells. From the original mixed population 50,000 RFP positive cells were collected in an Eppendorf tube in cold 1x PBS. The sorted cells were immediately processed for RNA extraction (see 7.1.2).

7.4.4.3 Immunoprecipitation (IP) with the anti-Flag M2 affinity gel

In total 1.2x10^6 Neuro2a cells were transiently transfected with 10 µg of a vector encoding a HA-Flag tagged protein with Lipofectamine 2000 (see 7.4.2). The Neuro2a culture medium was changed after 24 h to differentiation medium to push cells towards neuritogenesis. The cells were lysed in 600 µl JS lysis buffer (supplemented with protease inhibitors) for 30 min at 4 °C after 48 h of expressing the tagged proteins of interest. The lysate was sonicated for 5 rounds (30 sec on/ 30 sec off, high energy) with a Bioruptor Plus (Diagenode) at 4 °C and cleared by centrifugation with 16,000 x g at 4 °C for 15 min afterwards. The whole supernatant was subjected to IP after 10 % input have been taken (60 µl). For each IP reaction 40 µl of Anti-Flag M2 affinity gel were taken which have been washed 3x with 1 ml of JS buffer (centrifugation at 1,000 x g for 1 min at 4 °C) prior to usage. The cleared lysate was incubated with the antibody-coupled beads for 3 h at 4 °C while overhead rotating. Afterwards, the beads were wash two times with 1 ml NET buffer supplemented with protease inhibitors (centrifugation at 2,000 x g for 1 min at 4 °C). The beads were incubated with 60 µl 1x LDS sample buffer and boiled 10 min at 50 °C while shaking with 600 rpm for elution. The supernatant was saved after centrifugation with 4,000 x g for 2 min at RT, the elution was repeated with 60 µl 1x LDS sample buffer and the supernatants were finally combined. For WB analysis,10% input was supplemented with 1x LDS sample buffer and along with the IP elution fraction supplemented with 100 mM DTT, boiled at 90 °C for 5 min and directly used for SDS-PAGE (see 7.3.4).

7.4.4.4 SILAC-mass spectrometry (SILAC-MS)

Stable isotope labeling with amino acids in cell culture (SILAC) is a powerful, robust, yet relatively simple technique to measure proteomics data quantitatively [12]. This method enables to stably label cellular proteomes during propagation *in vitro* with either “heavy” or “light” (normal) isotope-containing amino acids and the mixture of these fractions allows the determination of protein abundances from the relative Mass Spectrometry (MS) signal intensities [12].

In this thesis Neuro2a cells were either labeled with Arg-0 + Lys-0 (light (L) medium) or Arg-10 + Lys-8 (heavy (H) medium) isotopes. For this purpose, Neuro2a cells were cultured with either H- or L-SILAC medium (see 6.7.2) for 6 passages. An incorporation test was performed by the Core Facility Proteomics (IMB, Mainz) and successfully labelled Neuro2a cells were frozen as stocks. For experiments, H- and L-labeled Neuro2a cells were cultured for one more passage and then transiently transfected (see 7.4.2) with respective HA-Flag tagged constructs to perform *in vitro* immunoprecipitation (IP, see 7.4.4.3). For each labelling condition a control pull down was additionally performed using normal IgG antibody of the same species as the primary antibody used to pull down the protein of interest (POI). After the first NET buffer wash, the beads for the SILAC pull down of H-POI and L-IgG (forward experiment) as well as H-IgG and L-POI (reverse experiment) were combined and the last wash was performed on the combined beads

of the different SILAC states. IP efficiency was checked by SDS-PAGE (see 7.3.4) followed by WB (see 7.3.5). In total 30 µL of the elution fraction in 1x LDS sample buffer supplemented with 100 mM DTT were boiled for 10 min at 70 °C in order to prepare the SILAC-IP material for MS analysis.
The MS analysis was performed by the Core Facility Proteomics (IMB, Mainz) and is therefore only briefly mentioned. The peptide fractions were analysed after Trypsine digest with a quadrupole Orbitrap (Q Exactive Plus, Thermo Scientific) equipped with a UHPLC system (EASY-nLC 1000, Thermo Scientific). Peptides were identified in the raw data files using MaxQuant (development version 1.5.2.8).

7.4.5 Chromatin assays

Two assays have been carried out in this thesis to investigate some parts of the epigenetic state of chromatin. First, chromatin immunoprecipitation (ChIP) was performed in order to elucidate whether a protein of interest (transcription factor or histone with certain post-translational modifications) is present at certain genomic loci [13]. Second, formaldehyde-assisted isolation of regulatory elements (FAIRE; [14, 15]) was carried out as a nucleosome-occupancy technique to measure chromatin compaction (to a certain degree) by robustly enriching for open chromatin regions which are nucleosome (protein) depleted.

7.4.5.1 Chromatin immunoprecipitation (ChIP)

Agarose beads (Millipore) were blocked in cold TE buffer supplemented with tRNA from yeast and BSA for 3 h at 4 °C to decrease unspecific binding of nucleic acids or proteins during the ChIP assay. The cells were cross-linked in medium containing 1% formaldehyde for 10 min at RT. The formaldehyde was neutralized with 0.125 M glycine for 5 min at 4 °C. Afterwards, cells were rinsed twice with 10 ml of cold 1x PBS and scraped off in 3 ml 1x PBS. They were pelleted by centrifugation with 600 x g for 7 min at 4 °C. The pellet was resuspended in 10 ml of L1 buffer and incubated at 4 °C for 10 min. This step was followed by centrifugation with 1,300 x g for 5 min at 4°C. Afterwards, the pellet was resuspended in 10 ml of L2 buffer and incubated for 10 min at RT followed by centrifugation with 1,300 x g for 5 min at 4 °C. The prepared nuclei were resuspended in L3 buffer (e.g. 900 µl per ES cell plate) containing protease inhibitors and incubated o/n at 4 °C. The solution was sonicated to a DNA fragment size around 300 bp with a Diagenode Bioruptor Plus (sonication 30 sec ON and 45 sec OFF per cycle; high energy) and afterwards cleared from cellular debris by centrifugation with 14,000 x g for 10 min at 4 °C. Chromatin was measured with a DQ 300 Fluorometer (Hoefer) using 490 µl TE, 500 µl Hoechst (200 ng/ml) and 10 µl sonicated solution. For each ChIP reaction 60 µg sonicated chromatin in 600 µl L3 buffer were precleared with 40 µl pre-blocked beads for 1 h at 4 °C. The chromatin solution was next incubated with overhead shaking o/n at 4°C with the appropriate antibody against the protein of interest after taking 1% input. Afterwards, the mixture was incubated with 40 µl pre-blocked beads for 3 h at 4 °C with overhead shaking. The beads were washed 2x with 1 ml L3 buffer and once with 1 ml DOC buffer for 5 min at RT, each time followed by 2 min centrifugation with 2,000 x g at 4 °C. The beads were transferred into a new Eppendorf tube with 1 ml cold TE buffer and centrifuged for 2 min with 2,000 x g at 4 °C, followed by removal of the complete supernatant. The bound chromatin fraction was eluted 2x for 20 min in 1% SDS/0.1 M $NaHCO_3$. The immunopurified chromatin as well as the input were further processed for reversal of crosslink. Treatment with RNase A (0.2 mg/ml) was

performed for 30 min at 37 °C for both fractions. Next, the IP fraction was complemented with 10 mM EDTA, 40 mM Tris (pH 6.5) and 50 µg/ml Proteinase K whereas the input was supplemented with 1% SDS, 100 mM NaCl and 200 µg/ml Proteinase K and incubated for 2.5 h at 55 °C followed by incubation at 65 °C o/n. The DNA was purified by phenol- and chloroform:isoamyl alcohol (24:1) extraction followed by ethanol precipitation. The purified DNA was recovered in 40 µl TE buffer and stored at -20 °C. Quantitative PCR analysis (ChIP-qPCR) was performed as described in 7.1.6.3 using SYBR Green chemistry (ABI) and ChIP or input DNA.

7.4.5.2 Formaldehyde-assisted isolation of regulatory elements (FAIRE)

For formaldehyde-assisted isolation of regulatory elements (FAIRE) the assayed cells were fixed, lysed and sonicated as described for ChIP (7.4.5.1). After clearing the cellular debris by centrifugation with 14,000 x g for 10 min at 4 °C, 10% input was stored at -20 °C. The DNA of the FAIRE material was isolated by adding an equal volume of phenol:chloroform:isoamyl alcohol (PCI, 25:24:1), vortexing, and centrifugation with 12,000 x g for 8 min at RT. The upper aqueous phase was transferred to a new Eppendorf tube followed by a second round of PCI purification. The upper aqueous phase was again collected and an equal volume of chloroform:isoamyl alcohol (24:1) was added, vortexed, and centrifuged with 12,000 x g for 8 min at RT. The upper aqueous phase was again transferred to a new Eppendorf tube, precipitated with cold ethanol and recovered in 40 µl TE buffer. The recovered material as well as the input were treated with RNase A (0.2 mg/ml) for 30 min at 37 °C. The input DNA was further supplemented with 1% SDS, 100 mM NaCl and 200 µg/ml Proteinase K and incubated for 2.5 h at 55 °C followed by overnight incubation at 65 °C while the FAIRE DNA was stored at -20 °C. The DNA of the FAIRE fraction as well as the input were further purified by phenol- and chloroform:isoamyl alcohol (24:1) extraction followed by ethanol precipitation. The purified DNA was recovered in 40 µl of TE buffer and stored at -20 °C. The quantitative PCR analysis (FAIRE-qPCR) was performed as described in 7.1.6.3 using SYBR Green chemistry (ABI) and FAIRE or input DNA.

7.5 Animals and *in vivo/ex vivo* experiments

Mice of C57BL/6N background were kept under standard housing conditions. All experiments were carried out according to local regulations.

7.5.1 *In utero* electroporation (IUE)

All *in utero* electroporations (IUEs) were performed by collaboration partners and therefore are only briefly mentioned here.

For the NeuroD1 project (Pataskar & Jung et al., 2016 [16]) the pCIDRE-NeuroD1 as well as the original pCIDRE vector were electroporated into the dorsal ventricular zone of E13.5 brains as previously described [17, 18]. Experiments were carried out at E15.5.

For the Zfp354c project the shZfp354c as well as shNTC vector were electroporated into the dorsal ventricular zone of E14.5 embryonic brains. Experiments were performed at E18.5.

7.5.2 Fluorescence-activated cell sorting (FACS) of *in vivo* derived cells

FACS of cortical cells was performed by collaboration partners as previously described [17] and is only briefly mentioned here. The electroporated cortices were dissociated at indicated time points using the Papain-based Neural Dissociation Kit (Milteney Biotec) after removal of meninges and ganglionic eminences. FACS was performed with a gating strategy for red (561 nm) or green (488 nm) fluorescence similar as described in 7.4.4.2.

The cells for ChIP experiments were cross-linked with 1% formaldehyde in 1x PBS for 10 min at RT after dissociation. The formaldehyde was neutralized with 0.125 M glycine and the cells washed two times with 1x PBS before continuing with FACS. Samples for RNA isolation were immediately processed for FACS after the dissociation and sorted into the RLT lysis buffer of the RNeasy Micro Kit. Afterwards, cells were snap frozen in liquid nitrogen and stored at -80 °C.

7.5.3 Chromatin immunoprecipitation *in vivo*

The *in vivo* ChIP of NeuroD1 (Cell Signaling) was performed on E14.5 embryonic cortices. For this purpose, time pregnant female mice were sacrificed by cervical dislocation. The E14.5 embryos were transferred to cold 1x PBS and the brains were extracted in cold 1x PBS. The meninges were removed from the dissected cortices, the cells resupended in 1x PBS containing 2 mM DSG (Sigma), and cross-linked for 45 min at RT. After four washing steps with 1x PBS, the ChIP assay was performed as described in 7.4.5.1.

The ChIP experiments on *in vivo* FAC-sorted cells were performed on 100,000 cells as described in 7.4.5.1. The fixation step with formaldehyde was omitted as the cells were already cross linked prior FACS (see 7.5.2).

7.5.4 Fixation and cryosectioning of embryonic mouse brains

Time pregnant female mice were sacrificed by cervical dislocation and embryos were isolated in cold 1x PBS. The heads were transferred to a new dish containing cold 1x PBS and embryonic brains were

carefully dissected and transferred to a 12 well plate. Isolated E15.5/18.5 embryonic brains were fixed for 24 h/overnight in 4% PFA in 1x PBS at 4 °C. Afterwards they were first cryoprotected in 10% sucrose (in PBS) for 2 h followed by 30% sucrose (in PBS) o/n. The embryonic brains were embedded in Tissue-Tek O.C.T. on dry ice for 30 min and stored at -20 °C. The embryonic brains were cryosectioned coronally with a CS3050S cryotome (Leica) in 12 µm sections, transferred to microscopy slides and stored at -20 °C.

7.5.5 Immunohistochemistry of brain cryosections

Cryosections were incubated with Blocking solution (2% BSA, 0.5% Triton X-100 in 1x PBS) for 1 h at RT. Sections were incubated with indicated primary antibodies o/n at 4 °C in blocking solution. After three washes with 1x PBS the sections were incubated with fluorochrome-labeled secondary antibody in blocking solution for 1 h at RT including a counterstain with Hoechst (200 µg/ml) to visualize nuclei. The sections where washed three times with 1x PBS, mounted with Shandon Immu-Mount (Thermo Scientific) and stored at 4 °C. Images where taken with a confocal laser-scanning microscope (Leica TCS SP5) and data were processed with ImageJ software.

7.6 Bioinformatic analyses

All bioinformatics analyses presented in this thesis were performed by collaborating partners and are therefore only briefly described here as mainly published in Pataskar & Jung et al. [16].

7.6.1 RNA-seq analysis

The RNA-seq poly(A) libraries were generated according to Illumina's instructions using oligo-dT primers by the Genomics Core Facility (IMB, Mainz). The RNA-sequencing output was in presented FASTQ format. After a quality check using FASTQC v2.6.14 [19], the output was aligned to the mouse genome (mm9) with UCSC's annotations using TopHat v2.0.8 [20]. Only uniquely mapped reads were retained for further analysis. SAMTOOLS v0.1.19 [21] was used to convert the BAM output to SAM format and to sort the BAM file. The read counts per gene were calculated using the HTSeq program, v0.5.4p1 [22]. The DESeq package [23] was used to generate normalized read counts and for differential gene expression analysis. DESeq called differentially expressed genes with FDR cutoff of 0.05 and abs(FC) > 1.5 were considered as significant differentially expressed genes. *In vivo* RNA-seq data from various embryonic tissues were normalized together by implementing RPKM normalization from Cufflinks package [24]. GO term analysis was performed using GenePattern GSEAPreranked [25] and ToppGene [26].
Gene lists derived from RNA-seq analysis following NeuroD1 induction in mESCs are supplied in 8.1 and 8.2.

7.6.2 ChIP-seq, FAIRE-seq, and motif analyses

Quality control of the ChIP/FAIRE-sequencing output was undertaken using FASTQC [19]. Bowtie v0.12.9 [27] with default parameters was used to align ChIP-seq and FAIRE-seq reads to the mouse mm9 genome with annotations from UCSC [28]. Each read was aligned to maximally one position in the genome.

SAMTOOLS v0.1.19 was used to convert the SAM file into BAM format and to sort and index the BAM file. The peaks were called using MACS2 v2.0.10.20120913 [29] using default parameters. Peaks reproducible across two biological replicates with margin of 100 base pairs were retained. WIG files were generated using the QuasR [30] package and visualized using the UCSC genome browser. The raw read counts that aligned to each genomic feature (promoters, peaks, and enhancers) were calculated using QuasR. Enrichment was calculated after merging biological replicates. The same protocol was used for FAIRE-seq analysis.

For the NeuroD1 project, enhancers were identified based on H3K27ac peaks in *in vitro* derived d1 terminal neurons that were enriched at least 2-fold over the input. The enhancer regions were defined as -1,000 to +1,000 bases from the H3K27ac peak summit. Enhancers were associated with genes by nearest gene approach excluding promoter regions (-800 to +200 bp from TSS) using annotatePeaks.pl method from HOMER v4.7 [31]. Enhancer classification into genomic location subtypes was based on UCSC mm9 genome annotation [28]. HOMER v4.7 was used to annotate the peaks. Motif analysis of the promoters was performed by submitting the gene list to the HOMER findMotifs.pl program, and the motifs were defined from -800 to +200 nucleotides from the TSS. For motif analysis at the enhancers, the input regions were given to the findMotifs.pl program in BED format. De novo motif was built by sequence analysis using rGADEM [32] on enriched (E > 1) NeuroD1 peaks. Gene lists associated with NeuroD1-enriched genomic loci are provided in 8.3 and 8.4.

For the Zfp354c project, a similar ChIP-seq and motif analysis was performed. However, for the Zfp354c ChIP-seq data only the peaks above Enrichment score of 0.58 (log2(1.5)) were considered as significantly enriched peaks and retained for analysis. Additionally, the mapping parameters were adapted for repetitive sequences [33].

7.6.3 Bayesian modeling

For the NeuroD1 project, ChIP-seq data for 8 histone marks (HMs) and 50 transcription factors (TFs) in mESC (see 8.5) were downloaded from the NCBI GEO database after selection through quality check using FASTQC [34]. After ChIP-seq processing and analysis (as described above), a matrix of the raw read counts that aligned with promoters as well as enhancers was generated and further normalized for library size using Trimmed Mean of M-values (TMM) normalization in the edgeR v3.0 package [35]. The NaiveBayes module from WEKA v3.6 [36] was used for classification. For promoter analysis, targets were defined by NeuroD1-bound promoters which are upregulated upon its induction, while a control dataset was generated from randomly selected equal number of non-target promoters of induced genes. Similarly, for enhancer analysis, target enhancer sites associated with induced genes were taken as target set while the control set was generated by compiling equal number of randomly selected non-bound enhancer sites associated with induced genes. Calculations of classification accuracy and area under the curve were performed using ROCR v1.0.5 [37]. The information gain quotient for the selected features was calculated using infoGainAttributeEval method as provided in WEKA. The set of the most relevant features (transcription factors and histone marks) were selected using Wrapper approach with Naive Bayes as a classification algorithm and Best First as a search method.

7.6.4 Data deposition

All the next-generation sequencing datasets used for the NeuroD1 project have been submitted to GEO and will be publicly available under accession number GSE65072.

7.7 References Material and Methods

1. Mohn, F., et al., *Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors.* Mol Cell, 2008. **30**(6): p. 755-66.
2. Iacovino, M., et al., *Inducible cassette exchange: a rapid and efficient system enabling conditional gene expression in embryonic stem and primary cells.* Stem Cells, 2011. **29**(10): p. 1580-8.
3. Mullis, K., et al., *Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction.* Cold Spring Harb Symp Quant Biol, 1986. **51 Pt 1**: p. 263-73.
4. Arya, M., et al., *Basic principles of real-time quantitative PCR.* Expert Rev Mol Diagn, 2005. **5**(2): p. 209-19.
5. Bustin, S.A., *Real-time, fluorescence-based quantitative PCR: a snapshot of current procedures and preferences.* Expert Rev Mol Diagn, 2005. **5**(4): p. 493-8.
6. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
7. Bibel, M., et al., *Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells.* Nat Protoc, 2007. **2**(5): p. 1034-43.
8. Bibel, M., et al., *Differentiation of mouse embryonic stem cells into a defined neuronal lineage.* Nat Neurosci, 2004. **7**(9): p. 1003-9.
9. Ibrahim, S.F. and G. van den Engh, *Flow cytometry and cell sorting.* Adv Biochem Eng Biotechnol, 2007. **106**: p. 19-39.
10. Kumar, N. and N. Borth, *Flow-cytometry and cell sorting: an efficient approach to investigate productivity and cell physiology in mammalian cell factories.* Methods, 2012. **56**(3): p. 366-74.
11. Pierzchalski, A., A. Mittag, and A. Tarnok, *Introduction A: recent advances in cytometry instrumentation, probes, and methods--review.* Methods Cell Biol, 2011. **102**: p. 1-21.
12. Ong, S.E. and M. Mann, *A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC).* Nat Protoc, 2006. **1**(6): p. 2650-60.
13. Orlando, V., H. Strutt, and R. Paro, *Analysis of chromatin structure by in vivo formaldehyde cross-linking.* Methods, 1997. **11**(2): p. 205-14.
14. Giresi, P.G., et al., *FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) isolates active regulatory elements from human chromatin.* Genome Res, 2007. **17**(6): p. 877-85.
15. Simon, J.M., et al., *Using formaldehyde-assisted isolation of regulatory elements (FAIRE) to isolate active regulatory DNA.* Nat Protoc, 2012. **7**(2): p. 256-67.
16. Pataskar, A., et al., *NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program.* EMBO J, 2016. **35**(1): p. 24-45.
17. Aprea, J., et al., *Transcriptome sequencing during mouse brain development identifies long non-coding RNAs functionally involved in neurogenic commitment.* EMBO J, 2013. **32**(24): p. 3145-60.
18. Artegiani, B., C. Lange, and F. Calegari, *Expansion of embryonic and adult neural stem cells by in utero electroporation or viral stereotaxic injection.* J Vis Exp, 2012(68).
19. Andrews, S., *FastQC: a quality control tool for high throughput sequence data.* Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc, 2010.
20. Trapnell, C., L. Pachter, and S.L. Salzberg, *TopHat: discovering splice junctions with RNA-Seq.* Bioinformatics, 2009. **25**(9): p. 1105-11.
21. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.
22. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.
23. Oshlack, A., M.D. Robinson, and M.D. Young, *From RNA-seq reads to differential expression results.* Genome Biol, 2010. **11**(12): p. 220.
24. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.

25. Subramanian, A., et al., *GSEA-P: a desktop application for Gene Set Enrichment Analysis.* Bioinformatics, 2007. **23**(23): p. 3251-3.

26. Chen, J., et al., *ToppGene Suite for gene list enrichment analysis and candidate gene prioritization.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W305-11.

27. Langmead, B., *Aligning short sequencing reads with Bowtie.* Curr Protoc Bioinformatics, 2010. **Chapter 11**: p. Unit 11 7.

28. Karolchik, D., et al., *The UCSC Genome Browser database: 2014 update.* Nucleic Acids Res, 2014. **42**(Database issue): p. D764-70.

29. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.

30. Gaidatzis, D., et al., *QuasR: quantification and annotation of short reads in R.* Bioinformatics, 2015. **31**(7): p. 1130-2.

31. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-89.

32. Li, L., *GADEM: a genetic algorithm guided formation of spaced dyads coupled with an EM algorithm for motif discovery.* J Comput Biol, 2009. **16**(2): p. 317-29.

33. Treangen, T.J. and S.L. Salzberg, *Repetitive DNA and next-generation sequencing: computational challenges and solutions.* Nat Rev Genet, 2011. **13**(1): p. 36-46.

34. Edgar, R., M. Domrachev, and A.E. Lash, *Gene Expression Omnibus: NCBI gene expression and hybridization array data repository.* Nucleic Acids Res, 2002. **30**(1): p. 207-10.

35. Robinson, M.D., D.J. McCarthy, and G.K. Smyth, *edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.* Bioinformatics, 2010. **26**(1): p. 139-40.

36. Frank, E., et al., *Data mining in bioinformatics using Weka.* Bioinformatics, 2004. **20**(15): p. 2479-81.

37. Sing, T., et al., *ROCR: visualizing classifier performance in R.* Bioinformatics, 2005. **21**(20): p. 3940-1.

8 Appendix

8.1 URG in A2lox.NeuroD1 mESCs upon 48 h doxycycline induction

0610040B10Rik, 1110051M20Rik, 1110067D22Rik, 1300014I06Rik, 1700025G04Rik, 1700037C18Rik, 1700052K11Rik, 1810041L15Rik, 2010011I20Rik, 2010107G23Rik, 2010109K11Rik, 2310021P13Rik, 2310028H24Rik, 2410066E13Rik, 2410137F16Rik, 2510009E07Rik, 2610008E11Rik, 2610017I09Rik, 2610034B18Rik, 2610034M16Rik, 2610100L16Rik, 2610109H07Rik, 2610301G19Rik, 2610306M01Rik, 2700081O15Rik, 2700089E24Rik, 2810030E01Rik, 2810405F15Rik, 2810405K02Rik, 2810408M09Rik, 2900052N01Rik, 2900092D14Rik, 3110003A17Rik, 3110035E14Rik, 3110052M02Rik, 3632451O06Rik, 3830406C13Rik, 4732456N10Rik, 4930402H24Rik, 4930412O13Rik, 4930506M07Rik, 4930578N16Rik, 4930588N13Rik, 4931428F04Rik, 4932418E24Rik, 4933426M11Rik, 4933432B09Rik, 4933433P14Rik, 4933436C20Rik, 5330437I02Rik, 5430417L22Rik, 5730409E04Rik, 5930412G12Rik, 6030446N20Rik, 6230427J02Rik, 6330403A02Rik, 6330403K07Rik, 6330406I15Rik, 6330407J23Rik, 6330439K17Rik, 6330527O06Rik, 6430548M08Rik, 6430571L13Rik, 6430704M03Rik, 6530402F18Rik, 6530418L21Rik, 8430408G22Rik, 8430427H17Rik, 9030425E11Rik, 9130023H24Rik, 9130206I24Rik, 9330159F19Rik, 9330182L06Rik, 9430021M05Rik, 9830001H06Rik, 9930013L23Rik, A430110N23Rik, A730011L01Rik, A930004D18Rik, A930038C07Rik, Aacs, Abat, Abca5, Abca9, Abcd2, Abhd14b, Abhd15, Abtb2, Acaa2, Acadvl, Acap3, Acbd3, Accn1, Accn2, Accn4, Accs, Ace, Acer2, Ache, Acly, Acpl2, Acss3, Acta2, Actc1, Actg2, Actl6b, Adam10, Adam12, Adamts12, Adamts16, Adamts20, Adamts4, Adamts6, Adamts7, Adamtsl1, Adamtsl2, Adarb1, Adcy8, Adcyap1, Adcyap1r1, Add1, Add2, Adra1b, Adra2c, Adrbk2, Adssl1, Afap1, Aff1, Aff2, Aff3, Agap1, Agap2, Agap3, Agphd1, Ahrr, AI429214, AI593442, AI646023, Airn, Ajap1, Akap13, Akap2, Akap6, Akna, Akr1b8, Akr1e1, Akt3, Alad, Alcam, Alk, Alpk1, Amz2, Angpt1, Angptl2, Ank2, Ankrd13b, Ankrd33b, Ankrd42, Ankrd43, Ankrd44, Ankrd50, Ankrd54, Ankrd6, Anks6, Ano1, Ano4, Ano5, Antxr2, Ap1m1, Ap1s2, Ap3b2, Ap3m2, Apba2, Apbb1, Apc2, Apcdd1, Aph1b, Aplnr, Apool, App, Arhgap24, Arhgap28, Arhgap31, Arhgap33, Arhgap36, Arhgap42, Arhgef10l, Arhgef9, Arid1b, Arid3a, Arid3b, Arl15, Arl16, Arl2bp, Armcx1, Armcx2, Arnt2, Arpp21, Arsi, Arx, Asap1, Asap2, Asb4, Astn1, Astn2, Asxl3, Atad2b, Atat1, Atg10, Atl1, Atoh8, Atp1a2, Atp2b2, Atp8b4, Atxn1, Atxn7l3b, AU040320, Auts2, AV249152, AW011738, Axin2, B230312A22Rik, B3galt1, B3galt2, B3gat1, B630019K06Rik, B9d2, Baalc, Bace2, Bach2, Bai3, Barhl2, Basp1, Baz2b, Bbs1, BC005764, BC013712, BC018242, BC046404, BC068157, BC089491, Bche, Bcl2, Bcl2l1, Bcl7a, Bcl9, Bco2, Bcorl1, Bcr, Bean1, Begain, Bex2, Bgn, Bhlhe22, Bhlhe23, Bhmt, Bicc1, Bin1, Bmi1, Bmp1, Bmp5, Boc, Bpgm, Brd9, Brsk1, Brsk2, Bsn, Btbd17, Btg1, Btg2, Btla, C1s, C2cd4c, C530008M17Rik, C630004H02Rik, Cables1, Cachd1, Cacna1b, Cacna1c, Cacna1d, Cacna1e, Cacna1g, Cacna1h, Cacna1i, Cacna2d1, Cacna2d2, Cacna2d3, Cadm2, Cadm3, Cadps, Calcoco1, Calcrl, Calm1, Camk1g, Camk2a, Camk2n1, Camk2n2, Camkk1, Camsap1l1, Cand2, Cap2, Car10, Car3, Car9, Carhsp1, Carns1, Cartpt, Cask, Caskin1, Casp3, Casp6, Casq2, Casr, Cass4, Cbfa2t2, Cbfa2t3, Cblb, Cbln1, Cbx2, Cbx4, Cbx8, Ccdc109b, Ccdc148, Ccdc159, Ccdc19, Ccdc28b, Ccdc33, Ccdc46, Ccdc8, Ccdc85a, Ccdc96, Ccl25, Ccnd1, Ccnd2, Ccng2, Ccni, Ccnj, Ccnjl, Ccno, Cd200, Cd24a, Cd40, Cdc14a, Cdc42ep4, Cdh10, Cdh11, Cdh13, Cdh2, Cdh20, Cdh22, Cdh6, Cdh9, Cdhr5, Cdk14,

Cdk5r1, Cdk5r2, Cdkl1, Cdkl5, Cdkn1b, Cdkn1c, Cdkn2a, Cdkn2aip, Cdkn2d, Cdon, Cecr6, Celf2, Celf3, Celf6, Celsr2, Celsr3, Cend1, Cep170, Cep97, Cerk, Cerkl, Ch25h, Chat, Chd3, Chd7, Chga, Chgb, Chl1, Chmp1b, Chn1, Chn2, Chodl, Chrd, Chrdl1, Chrna3, Chrna4, Chrna5, Chrna7, Chrnb2, Chrnb4, Chrng, Chst1, Chst15, Chst2, Chsy3, Cib2, Cited2, Clasp1, Cldn5, Clip2, Clip3, Clstn2, Clvs1, Cmpk2, Cmtm3, Cnih2, Cnksr2, Cnksr3, Cnp, Cnr1, Cnrip1, Cntfr, Cntln, Cntn1, Cntn2, Cntn3, Cntn4, Cntnap1, Cntnap5a, Cog1, Col11a1, Col11a2, Col14a1, Col15a1, Col28a1, Col3a1, Col4a1, Col4a2, Col4a5, Col4a6, Col6a1, Commd3, Copa, Copz2, Coro2b, Cotl1, Cp, Cpe, Cpeb2, Cpeb4, Cplx1, Cplx2, Cpne2, Cpne6, Crabp1, Crb2, Creg2, Crem, Crhr1, Crispld2, Crmp1, Crtc3, Cryba1, Cryba2, Csgalnact1, Csk, Csrnp3, Csrp1, Ctbp1, Cthrc1, Ctnna2, Ctnnd2, Ctso, Cttnbp2, Ctxn1, Cuedc1, Cul7, Cxcr4, Cxcr7, Cxx1b, Cxx1c, Cxxc4, Cxxc5, Cygb, Cyp1b1, Cyp26b1, Cyp2j9, Cyp4f15, Cyp7b1, Cyth1, D030025P21Rik, D0H4S114, D430019H16Rik, D430041D05Rik, D430050G20, D630003M21Rik, Daam2, Dab2, Dach1, Dact1, Darc, Dbh, Dbn1, Dbx1, Dcaf5, Dcc, Dchs1, Dclk1, Dcx, Ddah2, Ddx25, Ddx26b, Ddx4, Dennd1a, Dennd2a, Dexi, Dfna5, Dgki, Dgkk, Diras1, Disp2, Dkk3, Dlg3, Dlg4, Dlgap2, Dlk2, Dll1, Dll3, Dll4, Dmrt3, Dmrta2, Dnahc11, Dnahc5, Dnaja4, Dnajc12, Dner, Dnlz, Dnm3, Dnm3os, Dnmt3a, Doc2b, Dock10, Dock4, Dock5, Dock7, Dock8, Dok1, Dok5, Dopey2, Dpf2, Dpyd, Dpysl2, Dpysl3, Dpysl4, Dpysl5, Drd2, Dscam, Dtx4, Dusp10, Dusp15, Dusp18, Dusp22, Dusp26, Dusp4, Dusp6, Dusp8, Dvl3, Dync1i2, Dync2h1, Dynlt1a, Dyrk1b, Dyrk2, E130203B14Rik, E130309F12Rik, E130317F20Rik, E2f6, Ebf1, Ebf2, Ebf3, Ecscr, Edil3, Edn3, Ednra, Ednrb, Eef2k, Efcab6, Efna2, Efna5, Efnb1, Efnb2, Egfem1, Egflam, Egln1, Ehd3, Eid1, Eif1b, Eif2c1, Eif2c4, Eif4e3, Elavl2, Elavl3, Elavl4, Elfn1, Elovl2, Eme2, Emid2, Emilin1, Emilin2, Eml1, En1, En2, Enc1, Enho, Enox2, Entpd1, Epas1, Epb4.1, Epha3, Epha4, Epha5, Epha7, Ephb1, Ephb6, Epm2a, Erbb2ip, Erbb4, Esrrg, Etv1, Evi5l, Evl, Exoc7, Eya1, Eya2, F930015N05Rik, Fabp5, Fabp7, Fads3, Fam101b, Fam102b, Fam109a, Fam110a, Fam115a, Fam117b, Fam123a, Fam123c, Fam124b, Fam125b, Fam131b, Fam13a, Fam149a, Fam151b, Fam155a, Fam163a, Fam171a1, Fam171a2, Fam171b, Fam172a, Fam174b, Fam181b, Fam188b, Fam189a1, Fam196b, Fam19a4, Fam19a5, Fam33a, Fam53b, Fam55d, Fam57a, Fam57b, Fam59b, Fam5b, Fam63b, Fam65b, Fam70a, Fam78b, Farp1, Fat3, Fat4, Fbll1, Fbn2, Fbp1, Fbxl16, Fbxl22, Fbxl7, Fbxo10, Fbxo16, Fbxo21, Fbxo32, Fes, Fez1, Fezf2, Fgd3, Fgd4, Fgd5, Fgf10, Fgf12, Fgf13, Fgf14, Fgf15, Fgf3, Fgf9, Fgfbp3, Fhad1, Fhod3, Fign, Filip1, Filip1l, Fkbp14, Fli1, Flrt1, Flrt3, Flt1, Fmnl2, Fmo2, Fn3krp, Fndc5, Foxa1, Foxb1, Foxc1, Foxd4, Foxl2, Foxl2os, Foxp2, Foxp4, Frat1, Frem1, Frem2, Frk, Frmd3, Frmd4a, Frmd4b, Frmd5, Fry, Frzb, Fscn1, Fsd1l, Fst, Fstl5, Ftsjd1, Fxyd1, Fyn, Fzd1, Fzd10, Fzd3, Gab2, Gabbr2, Gabrb2, Gadd45g, Gal, Gal3st4, Galntl2, Galntl4, Gamt, Gap43, Gas1, Gata4, Gatsl2, Gbe1, Gbx2, Gck, Gdap1, Gdap1l1, Gdf11, Gdi1, Gdpd1, Gdpd2, Gdpd5, Gem, Gfod2, Gfra1, Gfra2, Gfra3, Ggt7, Ghsr, Gimap6, Glis3, Glra1, Glrx, Gltpd1, Gm10421, Gm10584, Gm11346, Gm11818, Gm13476, Gm14207, Gm1568, Gm16039, Gm266, Gm5567, Gm606, Gm672, Gm6724, Gm7455, Gmip, Gnao1, Gnas, Gnat3, Gnaz, Gnb4, Gnb5, Gng11, Gng2, Gng4, Gng8, Gpc2, Gpc6, Gpcpd1, Gpm6a, Gpm6b, Gpr123, Gpr137c, Gpr139, Gpr146, Gpr153, Gpr156, Gpr173, Gpr30, Gpr85, Gprin1, Gpsm1, Gpx3, Grasp, Greb1, Greb1l, Grem2, Gria1, Gria2, Gria4, Grid2ip, Grik2, Grik4, Grin2c, Grin3a, Grip1, Grip2, Grm4, Grm6, Grm7, Grm8, Gsbs, Gsdmd, Gse1, Gsg1l, Gsk3b, Gstm7, Gtf2ird2, Gulp1, H1f0, H2afy2, H3f3a, Habp4, Hacl1, Haghl, Hand2, Hcn4, Hdac11, Hdac7, Hdac9, Hdgfrp3,

Hdhd3, Hectd2, Hecw1, Hecw2, Heg1, Helt, Hes1, Hes5, Hes6, Hey1, Hey2, Heyl, Hif1a, Hipk2, Hivep1, Hmcn1, Hmga2, Hmgb3, Hmgn2, Hmgn3, Hmgn5, Hmx3, Hn1, Hnrnph3, Hoga1, Homer2, Hoxa1, Hoxa2, Hoxa3, Hoxb1, Hoxb2, Hoxb3, Hoxb4, Hoxb5, Hoxb8, Hoxd13, Hp1bp3, Hpca, Hprt, Hpse, Hrh3, Hs3st1, Hs3st2, Hs3st3a1, Hs3st3b1, Hs3st5, Hsd11b2, Hsd17b7, Hspa12a, Hspa12b, Hspa4l, Htr1a, Htr1d, Htr3a, Htra3, Hyi, Ick, Id1, Id2, Id3, Id4, Idh1, Ier5, Ier5l, Iffo2, Ifit1, Ift122, Ift74, Igdcc3, Igdcc4, Igf2bp3, Igfbp4, Igfbp5, Igfbpl1, Iglon5, Igsf11, Igsf21, Ikzf2, Il11, Il18, Il1rap, Il2rg, Ildr2, Ina, Inpp4a, Inpp5e, Insc, Insm1, Insrr, Invs, Ip6k2, Ipcef1, Iqcc, Iqsec2, Irgq, Irs1, Irs2, Irs3, Irx1, Irx2, Irx3, Irx4, Irx5, Isl1, Islr2, Ism1, Ispd, Itga4, Itgbl1, Itih5, Itm2a, Itpkb, Jag1, Jakmip2, Jakmip3, Jam3, Jhdm1d, Jph3, Jun, Kalrn, Katnal1, Katnal2, Kbtbd4, Kcna2, Kcna5, Kcna6, Kcnab1, Kcnb2, Kcnc1, Kcnd3, Kcne1l, Kcnf1, Kcnh1, Kcnh4, Kcnh5, Kcnip2, Kcnj5, Kcnj6, Kcnk10, Kcnk13, Kcnk2, Kcnk9, Kcnmb1, Kcnmb2, Kcnmb4, Kcnn3, Kcnq2, Kcnq4, Kcnt1, Kctd1, Kctd10, Kctd7, Kdr, Khdrbs1, Kif19a, Kif1b, Kif21a, Kif21b, Kif26a, Kif26b, Kif3a, Kif3c, Kif5a, Kif5c, Kirrel3, Kitl, Klc1, Klf14, Klf7, Klhdc2, Klhdc8a, Klhl14, Klhl29, Klhl34, Krt12, Ky, Kynu, L1cam, L3mbtl3, Lama2, Larp7, Lats2, Lbh, Lbx1, Lcor, Ldoc1l, Lef1, Leprel1, Letm2, Letmd1, Lgi3, Lgi4, Lgr5, Lhfpl4, Lhpp, Lhx1, Lhx2, Lhx3, Lhx4, Lhx5, Lhx9, Limch1, Limd2, Lingo1, Lix1, Lix1l, Lmo1, Lmo2, Lmo3, Lmod2, Lonp2, Lonrf2, Lox, Loxl1, Loxl3, Lpar6, Lphn3, Lpo, Lrfn2, Lrp1, Lrp12, Lrp3, Lrrc14, Lrrc16b, Lrrc36, Lrrc4, Lrrc46, Lrrc49, Lrrc4b, Lrrc7, Lrrn1, Lrrn2, Lrrn3, Lrrtm1, Lrrtm2, Lrrtm3, Lrrtm4, Lrtm2, Lsamp, Lsp1, Ltbp2, Lxn, Ly6h, Lynx1, Lyrm2, Lzts1, Mab21l1, Mab21l2, Madd, Maf, Mafb, Maged1, Maged2, Magee1, Mamdc2, Maml2, Maml3, Man1c1, Maob, Map2k6, Map3k13, Map4k3, Mapk10, Mapk7, Mapk8ip1, Mapkapk2, Mapre2, Marcks, Masp1, Mast4, Mbd5, Mbp, Mctp1, Mdga1, Mecp2, Mef2c, Megf10, Megf6, Megf9, Meis1, Meis2, Meis3, Mest, Met, Mex3a, Mex3b, Mfap2, Mfap4, Mfhas1, Mfng, Mgat5b, Miat, Mical2, Micalcl, Micall1, Midn, Mir3078, Mkrn3, Mll5, Mllt11, Mllt3, Mmd2, Mmp15, Mmp24, Mn1, Mns1, Mospd1, Mpdz, Mpp2, Mpped1, Mpped2, Mrc2, Mrgpre, Msi1, Msra, Msrb3, Mtap1a, Mtap1b, Mtap2, Mtap6, Mtap7d1, Mterfd3, Mthfd2l, Mthfsd, Mtvr2, Muc4, Mum1l1, Mxd4, Mxra7, Mybl1, Mybpc1, Mycbp2, Mycl1, Myl9, Mylk, Myo16, Myo1b, Myo5b, Myt1, Myt1l, N28178, Naalad2, Nacad, Nalcn, Nanos2, Nap1l5, Napb, Narf, Nat1, Nat14, Nav1, Nav3, Nbea, Ncald, Ncam1, Ncam2, Ncan, Ncf2, Nckap5, Nckap5l, Ncoa2, Ncoa5, Ncoa6, Ncoa7, Ncs1, Ndn, Ndrg4, Ndst1, Ndst4, Nebl, Nedd9, Nefm, Nelf, Nell2, Neo1, Neto2, Neurod1, Neurod2, Neurod4, Neurog1, Nf1, Nfasc, Nfatc1, Nfia, Nfil3, Ngfr, Ngly1, Nhlh1, Nhlh2, Nhs, Nhsl1, Nhsl2, Nicn1, Nin, Nkain4, Nkd1, Nkx1-2, Nkx2-5, Nkx6-1, Nlgn1, Nme5, Nnat, Nog, Nol4, Nos1, Nos1ap, Notch1, Nova1, Nova2, Npas3, Npr2, Npy, Nr2f1, Nr2f2, Nr4a2, Nr6a1, Nrarp, Nrbp1, Nrcam, Nrg1, Nrip1, Nrip3, Nrl, Nrn1, Nrp1, Nrxn1, Nrxn2, Nrxn3, Nsg1, Nsg2, Nt5m, Ntf3, Ntm, Ntn3, Ntn4, Ntng2, Ntrk1, Ntrk2, Ntrk3, Nuak1, Nuak2, Nudt7, Numbl, Nxn, Nxph1, Nxph4, Nynrin, Oat, Obscn, Odz1, Odz2, Ogfr, Olfm1, Olfml2a, Olfml2b, Olfml3, Olfr618, Olig1, Olig2, Olig3, Onecut1, Onecut2, Opcml, Ophn1, Opn3, Oprl1, Optc, Orai2, Osbp2, Osbpl5, Otud7a, P2rx3, Pacs1, Pacs2, Pafah1b2, Pag1, Paip1, Pak3, Pak7, Palm, Palmd, Pappa2, Papss1, Pard3, Pard3b, Parm1, Parp11, Parp3, Parp6, Parp8, Pax2, Pax3, Pax7, Pbx1, Pbx3, Pcbp4, Pcdh17, Pcdh18, Pcdh19, Pcdh7, Pcdh8, Pcdh9, Pcdha1, Pcdha10, Pcdha11, Pcdha2, Pcdha3, Pcdha5, Pcdha6, Pcdha9, Pcdhb10, Pcdhb11, Pcdhb12, Pcdhb13, Pcdhb14, Pcdhb15, Pcdhb16, Pcdhb17, Pcdhb18, Pcdhb19, Pcdhb2, Pcdhb20, Pcdhb21, Pcdhb22, Pcdhb3, Pcdhb4, Pcdhb5, Pcdhb6, Pcdhb7, Pcdhb8, Pcdhga10, Pcdhga2, Pcdhga3, Pcdhga4, Pcdhga5, Pcdhga6, Pcdhga7, Pcdhga9,

Pcdhgb2, Pcdhgb4, Pcdhgb5, Pcdhgb6, Pcdhgc3, Pcdhgc4, Pclo, Pcmtd2, Pcsk2, Pcsk5, Pde10a, Pde1a, Pde1c, Pde2a, Pde3a, Pde4b, Pde5a, Pde9a, Pdgfd, Pdgfra, Pdgfrb, Pdk2, Pdlim4, Pdlim5, Pdrg1, Pdzrn3, Pdzrn4, Pea15a, Peci, Peg10, Peg12, Peg3, Peli1, Peli2, Per2, Pex5l, Pfkfb4, Pfn2, Pgbd5, Pgcp, Pgm2l1, Pgm5, Pgpep1, Phactr1, Phactr2, Phc2, Phf14, Phf21b, Phf6, Phlda1, Phldb1, Phldb2, Phox2b, Phyhip, Phyhipl, Pid1, Pigp, Pik3c2b, Pik3r1, Pik3r3, Pim2, Pip4k2a, Pirt, Pitx2, Pja1, Pja2, Pkd2l1, Pkdcc, Pkia, Pkib, Pknox2, Plagl1, Plat, Plcb1, Plcd1, Plcg1, Plch1, Plcl1, Plcl2, Plcxd3, Pld5, Plekha1, Plekha6, Plekha8, Plekhg1, Plekhh3, Plekho1, Plekho2, Plk3, Plscr3, Plxna2, Plxna4, Plxnc1, Pm20d2, Pmp22, Pnmal2, Pnrc1, Pnrc2, Poc1b, Podxl2, Pogk, Pomc, Pou3f1, Pou3f2, Pou3f3, Pou4f1, Pou6f2, Ppfia2, Ppm1e, Ppp1r14a, Ppp1r3c, Ppp1r3d, Ppp1r9b, Ppp2r2b, Ppp2r3a, Ppp2r5b, Ppp3ca, Prdm12, Prdm6, Prdm8, Prdx2, Prex1, Prickle1, Prickle2, Prkacb, Prkar1a, Prkar1b, Prkar2b, Prkcb, Prkce, Prkch, Prkd1, Prkg1, Prkg2, Prkra, Prmt2, Prmt8, Prokr2, Prox1, Proz, Prph, Prrx1, Prrxl1, Prtg, Prune2, Psd, Psd3, Psme2, Ptbp2, Ptchd2, Ptger1, Ptgfr, Ptgir, Pth1r, Pth2r, Ptk2, Ptn, Ptpn13, Ptpn5, Ptpn9, Ptprd, Ptprn, Ptprn2, Ptpro, Ptprt, Pts, Pygl, Pygo1, Qsox1, R3hcc1, R3hdml, Rab11fip5, Rab13, Rab28, Rab36, Rab38, Rab39, Rab3a, Rab3c, Rab3d, Rab6b, Rabl2, Rabl5, Radil, Rai14, Rai2, Ralgds, Rapgef5, Rarb, Rasa3, Rasal2, Rasd1, Rasgef1b, Rasgef1c, Rasgrp1, Rasl10b, Rasl11a, Rasl11b, Rasl12, Rassf2, Rassf4, Rb1, Rbfox1, Rbfox2, Rbfox3, Rbm18, Rbm20, Rbm24, Rbm46, Rbms3, Rcan2, Rcbtb2, Rcor2, Rdh5, Refbp2, Reln, Resp18, Ret, Rftn1, Rftn2, Rfx3, Rfx4, Rgag4, Rgl1, Rgma, Rgmb, Rgs13, Rgs16, Rgs2, Rgs3, Rgs4, Rgs7, Rgs8, Rgs9, Rhbdl3, Rhobtb1, Rhobtb3, Rhog, Rhoj, Rhpn1, Rilpl2, Rimbp2, Rit1, Rln1, Rnasel, Rnd2, Rnd3, Rnf112, Rnf122, Rnf144a, Rnf146, Rnf150, Rnf152, Rnf165, Rnf182, Rnf219, Robo1, Robo2, Robo3, Ror1, Rorb, Rorc, Rpgrip1l, Rprm, Rprml, Rragd, Rtkn, Rtkn2, Rtn1, Rtn2, Rtn4rl1, Rtn4rl2, Rufy3, Rundc2a, Rundc3a, Runx1, Runx1t1, Rusc1, Rxra, Rxrg, Ryr3, S100a16, S1pr3, Sall2, Sall3, Samd12, Samd14, Samd5, Sarm1, Sat2, Satb1, Satb2, Sbk1, Scarf2, Sccpdh, Scd2, Scg2, Scg3, Scg5, Scn1a, Scn3a, Scn8a, Scn9a, Scrn1, Scrt1, Scrt2, Sct, Sctr, Scube1, Scube2, Sdc3, Sdcbp, Sdk1, Sdk2, Sema3a, Sema3d, Sema4c, Sema5a, Sema5b, Sema6d, Sema7a, Senp7, Sepn1, Sepp1, Sepw1, Serf1, Serinc2, Serpinb8, Serping1, Sesn1, Sesn3, Setbp1, Sez6, Sez6l, Sez6l2, Sfi1, Sgcb, Sgcd, Sgce, Sgip1, Sgk1, Sgpp2, Sh2b2, Sh3bgrl, Sh3bp2, Sh3bp4, Sh3kbp1, Sh3pxd2a, Sh3pxd2b, Shank2, Shc2, Shc3, Shd, Shf, Shisa4, Shisa9, Shoc2, Shox2, Sim1, Six1, Six2, Six3, Six5, Six6, Skor1, Slc10a4, Slc12a5, Slc14a1, Slc15a2, Slc17a6, Slc17a7, Slc18a2, Slc18a3, Slc1a2, Slc22a15, Slc22a23, Slc24a2, Slc24a3, Slc25a14, Slc25a24, Slc25a29, Slc27a6, Slc2a10, Slc2a13, Slc2a6, Slc35d2, Slc35f1, Slc36a4, Slc44a5, Slc4a4, Slc5a3, Slc6a4, Slc7a10, Slc8a1, Slco1a5, Slco3a1, Slco5a1, Slit1, Slit2, Slitrk1, Slitrk2, Slitrk5, Slitrk6, Smad1, Smad3, Smad5, Smarca1, Smarca2, Smarce1, Smoc2, Smpd3, Smpdl3a, Smug1, Snai1, Snai2, Snap25, Snap91, Snca, Sncaip, Snn, Snrk, Snx1, Snx33, Sobp, Socs3, Sorbs2, Sorcs1, Sorcs2, Sorcs3, Sos2, Sox1, Sox11, Sox12, Sox13, Sox18, Sox3, Sox4, Sox5, Sox6, Sox9, Sp4, Sp8, Spag4, Spag9, Sparcl1, Spata18, Spata6, Spats2l, Spg20, Sphkap, Spic, Spin4, Spns2, Spock1, Spred3, Spsb4, Srgap1, Srgap3, Srrm3, Srrm4, Ssbp2, Ssbp3, Sst, Sstr2, St18, St3gal1, St6galnac5, St7, St8sia2, St8sia3, Stambp, Stambpl1, Stat1, Stat3, Stc2, Stk32b, Stk32c, Stk33, Stk36, Stk39, Stmn1, Stmn1-rs1, Stmn3, Stmn4, Stra6, Stx1a, Stx1b, Stxbp1, Stxbp3a, Stxbp4, Stxbp5l, Stxbp6, Sufu, Suv39h1, Svep1, Sycp2, Syne2, Synm, Synpo, Synpr, Syp, Syt1, Syt11, Syt13, Syt16, Syt2, Syt6, Tada1, Tagln3, Tanc2, Tax1bp3, Tbata, Tbc1d16,

Tbc1d19, Tbc1d30, Tbc1d4, Tbx15, Tbx4, Tbx5, Tceal1, Tceal8, Tcf12, Tcf4, Tcfap2a, Tcfap2d, Tcfap4, Tcfcp2, Tcp11l1, Tcta, Tdrd1, Tdrd7, Tdrkh, Tead2, Tet3, Tfdp2, Tff3, Tgfb1i1, Tgfb3, Tgfbi, Tgfbr1, Tgfbr2, Tgfbr3, Tgm2, Thbs2, Thbs3, Thra, Thrb, Thsd4, Thsd7b, Tigd5, Timp3, Timp4, Tle1, Tll1, Tll2, Tlx1, Tlx2, Tmc3, Tmc7, Tmcc2, Tmem108, Tmem114, Tmem117, Tmem119, Tmem132b, Tmem132c, Tmem132e, Tmem145, Tmem151b, Tmem159, Tmem164, Tmem169, Tmem173, Tmem176a, Tmem176b, Tmem178, Tmem179, Tmem198, Tmem2, Tmem229b, Tmem80, Tmem86b, Tmem8b, Tmem90b, Tmod1, Tmsb4x, Tmtc2, Tnc, Tnfrsf21, Tnfsf11, Tnik, Tnr, Tnrc18, Tns1, Top1, Top2b, Tox, Tox2, Tox3, Tpm4, Tppp, Tppp3, Tram1l1, Trappc2, Trdmt1, Trdn, Trhde, Trib2, Tril, Trim21, Trim35, Trim36, Trim62, Trim68, Trim9, Trio, Trp53bp1, Trp53i11, Trp53inp2, Trp53rk, Trpc4ap, Tsc22d3, Tshz1, Tshz2, Tshz3, Tspan11, Tspan12, Tspan18, Tspan5, Tspyl4, Tspyl5, Ttbk1, Ttc12, Ttc21b, Ttc28, Ttc3, Ttc39c, Ttc8, Ttll7, Ttn, Ttyh3, Tuba1a, Tubb2b, Tubb3, Tubb4, Tulp4, Twist1, Txndc16, Txnl4b, Txnrd3, Ubash3b, Ube2d1, Ube2e3, Ube2g1, Ube2h, Ube2l6, Ube2ql1, Ube2r2, Ubqln2, Ubtd2, Uchl1, Ucma, Ugdh, Uggt2, Ulk2, Ulk4, Umodl1, Unc13a, Unc5c, Unc79, Unc80, Uncx, Uros, Ush2a, Usp11, Usp12, Usp29, Usp3, Usp46, Usp47, Usp6nl, Vamp4, Vash1, Vash2, Vat1l, Vav2, Vav3, Vax2, Vcam1, Vegfc, Vezf1, Vgll4, Vim, Vip, Vipr1, Vit, Vopp1, Vps53, Vps8, Vstm2b, Vsx1, Wbp1, Wbscr17, Wdfy3, Wdr13, Wdr17, Wdr41, Wdr47, Wdr6, Wdr60, Whsc1, Wnt10a, Wnt5a, Wnt8a, Wscd1, Wscd2, Xkr7, Xpr1, Yaf2, Ypel1, Ypel4, Ywhaq, Zadh2, Zbed3, Zbp1, Zbtb10, Zbtb12, Zbtb16, Zbtb4, Zbtb42, Zbtb46, Zbtb5, Zc3h12c, Zc3hav1l, Zc4h2, Zcchc16, Zcchc18, Zcchc24, Zeb1, Zeb2, Zfand3, Zfand5, Zfhx4, Zfp113, Zfp128, Zfp184, Zfp238, Zfp282, Zfp286, Zfp287, Zfp316, Zfp319, Zfp334, Zfp358, Zfp362, Zfp36l1, Zfp385b, Zfp386, Zfp41, Zfp422, Zfp423, Zfp438, Zfp467, Zfp518b, Zfp521, Zfp536, Zfp592, Zfp60, Zfp606, Zfp608, Zfp618, Zfp629, Zfp637, Zfp641, Zfp647, Zfp658, Zfp663, Zfp672, Zfp68, Zfp697, Zfp703, Zfp711, Zfp768, Zfp790, Zfp804a, Zfp811, Zfp821, Zfp9, Zfp90, Zfpm1, Zfpm2, Zhx1, Zhx2, Zhx3, Zic1, Zic4, Zim1, Zkscan2, Zkscan4, Zkscan6, Zmiz1, Zmynd11, Znrf2, Zpld1, Zscan2, Zscan21, Zswim4, Zswim5, Zswim6, Zyx.

8.2 DRG in A2lox.NeuroD1 mESCs upon 48 h Doxycycline induction

0610010O12Rik, 0610031J06Rik, 1110007C09Rik, 1190005I06Rik, 1300010F03Rik, 1500011K16Rik, 1600002K03Rik, 1600029D21Rik, 1700001L05Rik, 1700007K13Rik, 1700013B16Rik, 1700016M24Rik, 1700019D03Rik, 1700019G17Rik, 1700019N12Rik, 1700024P16Rik, 1700028K03Rik, 1700029I01Rik, 1700029P11Rik, 1700061G19Rik, 1810014F10Rik, 1810019J16Rik, 1810032O08Rik, 1810063B07Rik, 2010204K13Rik, 2200001I15Rik, 2200002D01Rik, 2210010C17Rik, 2210403K04Rik, 2210411K11Rik, 2310014L17Rik, 2310030G06Rik, 2310043J07Rik, 2310046K01Rik, 2310057J16Rik, 2410137M14Rik, 2610018G03Rik, 2610305D13Rik, 2610528J11Rik, 2810008D09Rik, 2810474O19Rik, 3110048L19Rik, 4632434I11Rik, 4732471J01Rik, 4831440E17Rik, 4930427A07Rik, 4930433N12Rik, 4930447C04Rik, 4930511M06Rik, 4930519F09Rik, 4930572J05Rik, 4933402E13Rik, 4933427D06Rik, 6330416G13Rik, 6720401G13Rik, 8430410A17Rik, 8430419L09Rik, 9030409G11Rik, 9030617O03Rik, 9130008F23Rik, 9130017N09Rik, 9330151L19Rik, 9330176C04Rik, 9530053A07Rik, 9530077C05Rik, A130022J15Rik, A130040M12Rik, A230050P20Rik, A230072C01Rik, A2ld1, A330069E16Rik, A4galt, A630089N07Rik, A830007P12Rik, A830080D01Rik, AA792892, Aass, Abca8b, Abcb1a, Abcb1b, Abcc1, Abcc2, Abcc4,

Appendix

Abcd4, Abhd11, Abhd14a, Abhd3, Ablim2, Abt1, Acacb, Acan, Acbd4, Acbd7, Accsl, Acnat1, Acot9, Acp5, Acp6, Acrbp, Acsl3, Acsl5, Actn3, Acvr1b, Adam23, Adam8, Adamts15, Adamts19, Adamts8, Adamtsl4, Adamtsl5, Adap1, Adap2, Adat3, Adcy3, Adm, Adm2, Adpgk, Adrb2, Aes, Agpat2, Agpat9, Agtrap, Ahnak, Ahr, AI414108, AI428936, AI450353, AI467606, AI507597, AI661453, AI747448, Aifm2, Aig1, Aim1, Aim1l, Aire, Ak7, Akap1, Akr1c14, Akr1c18, Aldh3a1, Alg1, Alg6, Alox12e, Alox5, Als2cl, Amhr2, Amn, Ampd3, Amz1, Angptl4, Ankle1, Ankle2, Ankrd45, Ankrd56, Ankrd9, Anln, Ano9, Anpep, Anxa1, Anxa11, Anxa3, Anxa7, Aoah, Ap1m2, Ap1s3, Apob48r, Apoc1, Apoe, Apof, Aqp3, Arap2, Arc, Arhgap22, Arhgap27, Arhgap30, Arhgap8, Arhgap9, Arhgef16, Arhgef19, Arhgef3, Arhgef4, Arhgef5, Arl6ip5, Arntl2, Arrdc1, Arrdc2, Artn, Asah1, Asap3, Aspg, Ass1, Atad5, Atf3, Atf7ip2, Atg9b, Atl2, Atox1, Atp10a, Atp10d, Atp11b, Atp12a, Atp1b1, Atp2c2, Atp6v0a4, Atp6v0b, Atp6v0c, AU015836, AU018091, AU023871, Aurkc, Avpi1, AW555464, B230206F22Rik, B2m, B3gnt7, B4galnt2, B4galnt3, B4galt5, B4galt6, BC018465, BC021767, BC021891, BC024386, BC024479, BC025920, BC030867, BC032203, BC064078, BC066135, Bcam, Bcap29, Bcat1, Bcl11b, Bcl2l14, Bcl3, Bdh2, Bhlhe40, Bhlhe41, Bik, Birc3, Blvrb, Bmp4, Bmp7, Bmp8b, Bnipl, Brca1, Brca2, Bscl2, Bspry, Bst2, Btbd19, Btnl7, C130074G19Rik, C1ql1, C2, C230081A13Rik, C3ar1, C530028O21Rik, C77080, C80913, Cabp1, Cacng6, Calca, Calcoco2, Calcr, Calr, Calr3, Calr4, Camk1d, Capg, Capsl, Car12, Car13, Car14, Car4, Cars, Casc4, Casp8, Cblc, Cbr3, Cbx6, Ccdc120, Ccdc125, Ccdc18, Ccdc60, Ccdc64b, Ccdc68, Ccdc79, Ccdc84, Ccdc88b, Ccnb1ip1, Ccne1, Ccnf, Cd164l2, Cd22, Cd37, Cd55, Cd6, Cd68, Cd74, Cd79a, Cd79b, Cd81, Cd84, Cd9, Cd97, Cdc42bpg, Cdc42ep5, Cdca2, Cdcp1, Cdh1, Cdh3, Cdhr1, Cdk6, Cds1, Cdyl2, Cep55, Cftr, Cgn, Cgref1, Chac1, Chchd10, Chchd4, Chd5, Chek2, Chka, Chmp4c, Chrna1, Chrna9, Chrnb1, Cisd3, Clca4, Clca6, Clcn3, Clcn7, Clcnkb, Cldn1, Cldn3, Cldn4, Cldn6, Cldn7, Cldn9, Clgn, Clic5, Cln3, Cln5, Cln6, Cln8, Clstn3, Cltb, Clu, Cmah, Cmtm6, Cnga3, Cnksr1, Cnnm2, Cnpy3, Cntnap3, Cobl, Col18a1, Col8a2, Col9a3, Comtd1, Cox6b2, Cpeb1, Cpm, Cpn1, Cpne5, Cpox, Cpsf4l, Cpt1a, Cpz, Crb3, Creb3l1, Creb3l3, Creb3l4, Crebl2, Creld1, Creld2, Crispld1, Crlf1, Crtap, Csn3, Cstb, Ctgf, Cth, Ctsb, Ctsc, Ctsd, Cxcl10, Cxcl16, Cxcl5, Cxcr2, Cyb561, Cyb5r1, Cyba, Cyp1a1, Cyp2j13, Cyp2j7-ps, Cyp2s1, Cyp2u1, Cyp3a13, Cyr61, Cysltr1, Cysltr2, D10Bwg1379e, D14Ertd668e, D19Ertd652e, D1Pas1, D630023F18Rik, D630039A03Rik, D830030K20Rik, D930048N14Rik, Dbf4, Dcbld1, Dcdc2c, Dcn, Ddx43, Dedd2, Def6, Defb42, Dennd1c, Dennd2c, Dennd3, Depdc7, Deptor, Dgke, Dgkh, Dhcr24, Dhh, Dhrs13, Dhrs7b, Diap1, Dleu7, Dmc1, Dmrt1, Dmrtc2, Dnahc6, Dnajc22, Dnajc4, Dnase2a, Dock6, Dok2, Dpep2, Dpep3, Dpm3, Dpp7, Dppa3, Dppa5a, Dpys, Dqx1, Dsc2, Dscaml1, Dsg2, Dsp, Dusp14, Dusp9, E330020D12Rik, Ebi3, Ebp, Echdc2, Efha2, Egfl6, Egfl7, Egfr, Egr1, Eid3, Eif2ak3, Eif4ebp1, Elf3, Ell2, Ell3, Elmo3, Eln, Elovl7, Emb, Emp1, Endod1, Eng, Eno3, Enoph1, Enpp2, Enpp3, Enpp4, Enpp5, Entpd2, Epb4.9, Epcam, Epdr1, Epha1, Epha2, Ephb3, Ephx1, Ephx2, Ephx3, Eppk1, Eps8l2, Epx, Eras, Erbb3, Erlin1, Ermp1, Ern1, Ero1l, Ero1lb, Esco2, Espn, Esrp1, Esrp2, Esrrb, Esyt3, Etl4, Etnk1, Etv4, EU599041, Extl1, Ezr, F11r, F3, F5, F830016B08Rik, Faah, Fabp3, Fam110b, Fam110c, Fam116b, Fam122b, Fam129a, Fam131c, Fam132b, Fam134b, Fam167a, Fam169b, Fam176a, Fam198b, Fam38a, Fam3c, Fam46a, Fam69b, Fam81a, Fam82a1, Fam83b, Fam83f, Fam83h, Fas, Fat1, Fat2, Fbxo15, Fbxo27, Fbxo34, Fbxo43, Fbxo48, Fbxo7, Fcgbp, Fcgr2b, Fcgrt, Fdxr, Fer1l4, Fermt1, Fetub, Fgd1, Fgf17, Fgf4, Fgfr3, Fgg, Fkbp10, Fkbp11, Fkbp9, Flnb, Fmn1, Fmnl1, Fmr1nb, Fndc1,

Folr1, Foxj1, Foxr1, Foxr2, Foxred2, Frmd7, Frrs1, Fstl3, Ftl1, Ftl2, Fuca2, Fut1, Fut9, Fxyd2, Fxyd6, Fzd5, Fzd7, Fzd8, Fzd9, G630016D24Rik, Gaa, Gabra2, Gabra4, Gabre, Gabrp, Gad1, Gadd45b, Gal3st1, Galnt12, Galnt3, Galnt6, Gas5, Gas6, Gbp1, Gcnt2, Gcnt3, Gdf15, Gfod1, Gfpt2, Ggnbp1, Ggt1, Ggta1, Ghr, Gipc2, Gjb1, Gjb3, Gjb5, Gla, Glb1, Gldc, Gli1, Glp1r, Gls2, Glt1d1, Glyctk, Gm1006, Gm10069, Gm10324, Gm10416, Gm106, Gm1082, Gm10863, Gm11127, Gm12794, Gm128, Gm129, Gm13051, Gm13139, Gm13152, Gm13154, Gm13157, Gm13238, Gm13242, Gm13251, Gm14164, Gm15070, Gm16516, Gm2016, Gm2022, Gm2a, Gm3002, Gm3336, Gm4349, Gm4884, Gm4934, Gm4975, Gm4984, Gm5039, Gm5114, Gm5134, Gm5458, Gm5592, Gm5662, Gm5887, Gm6300, Gm6402, Gm6484, Gm6578, Gm6792, Gm7102, Gm7325, Gm8300, Gm8994, Gm949, Gm98, Gmfg, Gnai1, Gnat2, Gnb1l, Gnmt, Got1l1, Gpa33, Gpnmb, Gpr114, Gpr124, Gpr133, Gpr135, Gpr160, Gpr19, Gpr4, Gpr56, Gpr68, Gpr75, Gprc5a, Gpsm3, Gpx2, Gpx6, Gramd1c, Grb7, Grhl1, Grhl2, Grhl3, Grik3, Grk4, Grn, Gstm3, Gstt2, Gtpbp6, Gtse1, Gtsf1, Gtsf1l, Gulo, Gusb, Gyltl1b, H2afj, H2-Bl, H2-DMb1, H2-Eb1, H2-Gs10, H2-M2, H2-M3, H2-Q1, H2-Q10, H2-Q2, H2-Q7, Hbegf, Hck, Hcn2, Hcrtr2, Helb, Hexa, Hexb, Hiat1, Hip1r, Hist1h2bc, Hmha1, Hmmr, Hnf1a, Hook2, Hormad1, Hormad2, Hoxa10, Hpn, Hr, Hrc, Hsd17b11, Hsd17b14, Hsf2bp, Hsp90b1, Hspb1, Hspb2, Hspb8, Hspg2, Htra1, Hvcn1, Hydin, Icosl, Idua, Ier3, Ifi30, Ifitm1, Ifitm2, Ifitm3, Ifitm5, Ifitm7, Igfals, Igfbp2, Igfbp3, Igsf9, Ikzf1, Il13ra1, Il16, Il17re, Il23a, Il27ra, Il28ra, Il34, Il6ra, Il7, Ildr1, Inadl, Incenp, Inhba, Inhbb, Inpp5d, Iqcd, Iqgap3, Irak2, Irf6, Itga3, Itga5, Itga6, Itga9, Itgb4, Itgb5, Itgb7, Itpka, Itpr3, Itpripl1, Izumo1, Jag2, Jam2, Jarid2, Jdp2, Jmjd7, Jph1, Junb, Kank2, Kank3, Kbtbd11, Kcna7, Kcnab3, Kcnh6, Kcnj10, Kcnj4, Kcnk1, Kcnk5, Kcnk6, Kcnn4, Kcnq1ot1, Kcnq3, Kcnu1, Kctd12b, Kctd16, Kctd19, Kdelr3, Kdsr, Kifc1, Kifc3, Kifc5b, Kirrel2, Kit, Klf15, Klf2, Klf5, Klf9, Klhl30, Klhl4, Klk10, Klk13, Klk8, Klk9, Klrg2, Kndc1, Kremen2, Krt17, Krt18, Krt19, Krt42, Krt7, Krt8, Krtcap3, L1td1, Lace1, Lad1, Lama3, Lama5, Lamb3, Lamc1, Lamc2, Laptm4b, Laptm5, Lax1, Lcat, Lck, Ldb3, Ldhb, Ldhd, Ldlrap1, Lefty1, Lemd3, Leprel4, Lgals12, Lgals3, Lgals9, Lgi2, Lgmn, Lif, Lima1, Lingo3, Lipg, Liph, Lipi, Llgl2, Lmo7, Lmtk2, LOC100303645, LOC100316870, Lonrf3, Loxl4, Lpar2, Lpar5, Lpcat2, Lpcat4, Lpl, Lrp10, Lrp5, Lrrc15, Lrrc2, Lrrc33, Lrrc34, Lrrc8c, Lrrc8d, Lrrc8e, Lrriq4, Lrrn4cl, Lsr, Ltb, Ltb4r1, Ltbp4, Ltk, Ly6g6c, Ly6g6e, Ly75, Lypd6, Lypd6b, M6pr, Macrod1, Madcam1, Mael, Magi2, Mal2, Mall, Man2b1, Manba, Map3k8, Mapk13, Mapk1ip1, Mapkapk3, Mars, Marveld1, Marveld2, Marveld3, Masp2, Matn1, Matn3, Matn4, Mbnl3, Mboat1, Mboat7, Mcam, Mcl1, Mcoln3, Mctp2, Mdm4, Mef2b, Mep1b, Mfap3l, Mfge8, Mfi2, Mfsd5, Mfsd7c, Mgat2, Mgl2, Mgll, Mia1, Mif4gd, Miip, Mip, Mir1949, Mir715, Mllt6, Mlxipl, Mme, Mmp17, Mmp23, Mmrn2, Mobp, Morc1, Mov10, Moxd1, Mpv17l, Mpv17l2, Mpzl2, Mreg, Mrps18b, Ms4a6d, Msc, Msln, Mt1, Mt2, Mtap7, Mtdh, Mterfd2, Mthfd2, Mto1, Myadm, Mybl2, Mybpc2, Myh14, Myh6, Myl3, Myl7, Mylpf, Myo1d, Myo1f, Myo1h, Myo5c, Myof, Myom2, Myst4, N4bp3, Naaa, Naaladl1, Nanog, Napsa, Nceh1, Nckap1l, Ncln, Ndrg2, Neb, Neil2, Neu1, Nfatc2, Nfatc2ip, Nfib, Nfic, Nfkbia, Nfkbid, Nfu1, Nhedc2, Nid2, Ninj1, Nipal2, Nkx2-9, Nlrp1a, Nlrp4a, Nlrp9b, Nodal, Nomo1, Nos3, Notch4, Notum, Npc1l1, Nphs1, Nphs1as, Npnt, Npr1, Npr3, Nptx2, Npy1r, Nr0b1, Nr3c1, Nr4a1, Nr5a2, Nrp2, Nt5e, Ntn1, Nudt13, Nudt19, Nupr1, Oas1a, Oas3, Obfc2a, Ocel1, Ocln, Olfr613, Olfr701, Olfr856-ps1, Olfr985, Olfr986, Ooep, Optn, Orai1, Orc1, ORF61, Osbpl10, Osm, Ostm1, P2rx2, P2rx4, P2rx7, P2ry2, Pacsin1, Padi1, Padi3, Pak6, Palm3, Paqr6, Paqr7, Pard6b, Park2, Parvb, Parvg, Patl2, Pcnxl2, Pcolce2, Pcsk4, Pcsk6, Pcsk9, Pdgfa,

Pdgfc, Pdgfrl, Pdia6, Pdk1, Pdk4, Pdlim1, Pdlim7, Pdpn, Pdyn, Pdzd7, Pecam1, Penk, Perp, Pfkp, Pga5, Pgc, Pgr, Phf19, Phlda2, Phtf2, Pif1, Pigf, Pigo, Pigt, Pigyl, Pim3, Piwil4, Pkmyt1, Pkp3, Pla1a, Pla2g10, Pla2g12a, Pla2g15, Pla2g1b, Pla2g4b, Pla2g4f, Pla2g5, Plac8, Plaur, Plbd1, Plcd3, Pld6, Plec, Plek2, Plekha4, Plekha7, Plekhb1, Plekhf2, Plekhg3, Plekhg5, Plekhg6, Plekhm1, Plk5, Pllp, Plp2, Pls1, Pltp, Plvap, Plxdc1, Plxnb3, Pmaip1, Pmel, Pmepa1, Pml, Pmm1, Pnldc1, Pnpla7, Podn, Podnl1, Pof1b, Polk, Pom121, Ppap2c, Ppfibp2, Ppl, Ppm1h, Ppm1j, Ppm1n, Ppp1r13l, Pqlc1, Pqlc2, Pramef12, Prdm1, Prdm14, Prickle3, Prlr, Prom2, Prr15l, Prrg4, Prss16, Prss23, Prss35, Prss41, Prss42, Prss50, Prss8, Prx, Psat1, Psd4, Psg28, Psph, Ptch2, Ptger4, Ptges, Ptgfrn, Ptgs2, Ptp4a3, Ptpn7, Ptprcap, Ptprk, Ptprv, Pttg1, Pvr, Pvrl2, Pvrl4, Pvt1, Pycard, Rab25, Rad54b, Rasa4, Rasal1, Rasd2, Rasef, Rassf6, Rbl2, Rbm47, Rdh11, Rdh12, Rdh18-ps, Rdh9, Rdm1, Reep4, Reep6, Relb, Reps2, Rest, Retsat, Rfx2, Rfx6, Rgs11, Rgs14, Rgs9bp, Rhbdf2, Rhbg, Rhd, Rhox9, Rinl, Ripk4, Rln3, Rmnd5b, Rnase4, Rnf125, Rnf128, Rnft1, Robo4, Rpl10l, Rpl39l, Rpn1, Rpp25, Rps27, Rrbp1, S100a1, S100a10, S100a6, S1pr2, Sat1, Scamp3, Scara5, Scarb1, Scn1b, Scn5a, Scpep1, Scube3, Sdc4, Sdcbp2, Sdf2l1, Sec1, Sec14l4, Sec14l5, Sec16a, Sema3b, Sema3c, Sema4a, Sema4b, Sema4d, Serinc4, Serpina7, Serpinb6a, Serpinb6c, Serpinb9b, Serpine1, Serpine2, Sertad1, Sesn2, Setd4, Sfn, Sfrp4, Sft2d1, Sft2d2, Sgk2, Sgms1, Sgms2, Sh2d4a, Sh3d19, Sh3gl2, Sh3rf2, Sh3tc1, Shb, Shisa2, Shisa3, Shroom2, Sigirr, Siglec5, Sil1, Sirt4, Six4, Ska1, Slc11a1, Slc12a4, Slc12a8, Slc12a9, Slc15a1, Slc16a1, Slc16a10, Slc16a2, Slc16a3, Slc16a8, Slc16a9, Slc17a9, Slc19a2, Slc19a3, Slc1a1, Slc1a3, Slc1a4, Slc1a5, Slc20a1, Slc20a2, Slc22a18, Slc22a7, Slc23a1, Slc23a2, Slc25a19, Slc25a36, Slc25a40, Slc25a42, Slc25a43, Slc26a2, Slc27a2, Slc28a1, Slc29a1, Slc29a2, Slc2a3, Slc2a8, Slc30a2, Slc30a3, Slc30a6, Slc35f2, Slc35f5, Slc38a5, Slc39a4, Slc39a8, Slc3a2, Slc40a1, Slc44a1, Slc44a3, Slc44a4, Slc45a4, Slc46a3, Slc47a1, Slc47a2, Slc4a1, Slc4a11, Slc4a5, Slc5a11, Slc5a5, Slc5a6, Slc6a1, Slc6a14, Slc6a18, Slc6a8, Slc6a9, Slc7a1, Slc7a3, Slc7a5, Slc7a7, Slc9a1, Slco4a1, Slfn10-ps, Slfn5, Slfn8, Slfn9, Slitrk3, Slitrk4, Smad7, Smagp, Smox, Smpdl3b, Snai3, Sned1, Snhg12, Snhg3, Snhg7, Snhg8, Snora17, Snora23, Snora28, Snora43, Snora44, Snora52, Snora61, Snora65, Snora70, Snora75, Snora81, Snord100, Snord12, Snord16a, Snord19, Snord2, Snord22, Snord33, Snord34, Snord47, Snord49a, Snord57, Snord88a, Snord96a, Snord99, Soat1, Soat2, Sord, Sorl1, Sox15, Sp6, Spats1, Specc1, Spib, Spint2, Spn, Spna1, Spnb1, Spnb4, Spns1, Spon1, Spp1, Sptlc2, Sqle, Srcrb4d, Srd5a3, Srpr, Srxn1, Ssr2, St14, St6gal2, St6galnac2, St8sia6, Stac2, Stag3, Stard8, Stat4, Stat6, Steap1, Steap2, Steap3, Stk30, Ston2, Stra8, Sult6b1, Sun1, Susd1, Svil, Swap70, Sycp3, Syngr1, Syngr2, Syngr3, Synj2, Synpo2l, Sypl, Syt12, Syt15, Syt3, Syt5, Tacstd2, Taf5l, Tagln2, Tap1, Tapbp, Tapbpl, Tarm1, Tbc1d8b, Tbl2, Tc2n, Tcea3, Tceal5, Tcfap2c, Tcfcp2l1, Tcfeb, Tcirg1, Tcl1, Tdgf1, Tdrd12, Tdrd5, Tead4, Tek, Tekt1, Terf1, Tert, Tet1, Tet2, Tex14, Tex15, Tex19.1, Tex2, Tex264, Tfrc, Tgm1, Tgm3, Tgm4, Thpo, Thy1, Timp1, Tinagl1, Tjp3, Tlcd1, Tle6, Tlr2, Tm4sf5, Tm7sf3, Tm7sf4, Tmc4, Tmc6, Tmc8, Tmcc3, Tmco4, Tmed5, Tmed9, Tmem116, Tmem120b, Tmem125, Tmem144, Tmem151a, Tmem154, Tmem180, Tmem184a, Tmem184b, Tmem19, Tmem191c, Tmem194b, Tmem20, Tmem30b, Tmem37, Tmem38b, Tmem39a, Tmem40, Tmem45a, Tmem47, Tmem48, Tmem5, Tmem54, Tmem59l, Tmem63a, Tmem63c, Tmem79, Tmem81, Tmem82, Tmem92-ps, Tmem97, Tmie, Tmppe, Tmprss13, Tmprss2, Tmprss5, Tmtc3, Tnfrsf10b, Tnfrsf11a, Tnfrsf12a, Tnk1, Tnnt1, Tns3, Tns4, Tor2a, Tor3a, Tprn,

Tpst2, Trank1, Treml2, Trh, Trib3, Trim47, Trim52, Trim54, Trim6, Trim63, Trim67, Triml1, Triml2, Trp73, Trpm1, Trpm4, Trps1, Trpv4, Tsix, Tspan1, Tspan17, Tspan4, Tspan7, Tspo, Tst, Tsx, Ttc13, Ttc39b, Ttc9, Ttll6, Ttpa, Tuba4a, Tuft1, Uap1l1, Ubash3a, Ube2t, Ugt8a, Ulk1, Unc13d, Unc5b, Unc5cl, Uqcrq, Usp2, Usp26, Usp28, Usp43, Usp50, Usp53, Utf1, Vgll3, Vill, Vkorc1l1, Vmn2r29, Vmn2r55, Vmn2r-ps129, Vrtn, Vsig10, Vsig2, Vwf, Was, Wdr31, Wdr64, Wdr86, Wfdc2, Whamm, Wnt3, Wnt4, Wnt7a, Wnt7b, Wnt9a, Wsb2, X99384, Xist, Ybx2, Ydjc, Yif1b, Yipf4, Yipf5, Yy2, Zan, Zbtb32, Zbtb7a, Zc3h12a, Zdhhc12, Zdhhc18, Zdhhc9, Zfp185, Zfp277, Zfp296, Zfp353, Zfp36, Zfp42, Zfp459, Zfp473, Zfp488, Zfp493, Zfp52, Zfp534, Zfp595, Zfp600, Zfp704, Zfp71-rs1, Zfp72, Zfp750, Zfp819, Zfp874a, Zfp936, Zfp946, Zfp951, Zfyve28, Zmynd15, Zp3, Zscan4a, Zscan4c, Zscan4d, Zscan4f, Zscan5b, Zswim1, Zwilch.

8.3 Promoter URT of NeuroD1

5330437I02Rik, 6330406I15Rik, Adamts16, Aff3, Ankrd6, Apc2, Arsi, BC018242, C630004H02Rik, Casp3, Cass4, Chrna5, Chrng, Ctso, Daam2, Dll3, Dll4, Dusp4, Dyrk2, Efcab6, Eif4e3, Eml1, Fam110a, Fgd4, Fxyd1, Gadd45g, Glis3, Gm11346, Gpr173, Hdac11, Hdac9, Hes6, Iffo2, Ifit1, Insm1, Irx4, Itgbl1, Kcna2, Klhl34, Kynu, Lgi4, Mfap4, Mmd2, Nat1, Ndst4, Neurod1, Nhlh1, Nog, Notch1, Ntf3, Numbl, Opcml, Pacs2, Pax2, Pcbp4, Pcsk2, Phyhip, Pou3f2, Ppp1r3d, Prdm8, Rbm20, Rbm24, Rnf182, Rorb, Runx1t1, Sbk1, Sct, Sema5b, Six5, Slc18a3, Slc1a2, Snai2, Snap91, Sox12, Srrm4, Stmn1, Tagln3, Tcfap4, Tcp11l1, Tet3, Thsd4, Tmem159, Ube2l6, Ush2a, Usp46, Zbp1, Zbtb12, Zfp238.

8.4 Enhancer URT of NeuroD1

1700025G04Rik, 1810041L15Rik, 2610100L16Rik, 4932418E24Rik, 4933436C20Rik, Adcyap1, Adcyap1r1, Akap2, App, Asb4, Basp1, Bhlhe23, Btg1, Btg2, Camk2n1, Casq2, Cbfa2t2, Cdh10, Cdh6, Cdk5r2, Celf2, Chd7, Chrna4, Chrnb4, Chst15, Clvs1, Cmpk2, Cnr1, Cntn2, Cplx2, Cxxc5, Dact1, Dcc, Dll1, Ebf3, Efnb2, Enc1, Fam174b, Fam19a4, Frmd4b, Gadd45g, Gatsl2, Grip2, Gse1, Hes6, Hpse, Hs3st3a1, Hs3st3b1, Id1, Igdcc4, Irx3, Jph3, Kalrn, Kcnd3, Kif21a, Klhl14, Lbx1, Lzts1, Maf, Map3k13, Mapre2, Marcks, Ncam1, Ncf2, Nckap5, Ncoa5, Ndrg4, Neurod4, Nhlh1, Nkx6-1, Nova1, Nova2, Nrg1, Nudt7, Nxph1, Pard3, Pja1, Plch1, Plekha1, Pmp22, Pou3f2, Ppp3ca, Prdm8, Psd3, Ptbp2, Rcor2, Rgmb, Rgs8, Robo1, Sall3, Scd2, Sgcd, Six3, Socs3, Sox11, Sox4, Spsb4, Srrm4, Sstr2, Syt13, Tgfbi, Tle1, Tox3, Trp53i11, Ulk4, Vash1, Wscd1, Zfand5, Zfp438, Zmiz1, Zpld1.

8.5 GEO accession numbers for publicly available RNA-seq and ChIP-seq datasets used in Pataskar & Jung et al. (2016)

Sequenced sample	Experiment	GEO ID	Sequenced sample	Experiment	GEO ID
AFF4	ChIP-seq	GSM749810	**TAF1**	ChIP-seq	GSE31270
BRD4	ChIP-seq	GSM823382	**TAF3**	ChIP-seq	GSM77494
CBP	ChIP-seq	GSM1246866	**TBP**	ChIP-seq	GSM774944
CBX7	ChIP-seq	GSM820726	**TBX3**	ChIP-seq	GSM476146
CDK7	ChIP-seq	GSE60027	**TET1**	ChIP-seq	GSM706672
CDK8	ChIP-seq	GSE44288	**TETc**	ChIP-seq	GSM611193
CDK9	ChIP-seq	GSM1082347	**TETn**	ChIP-seq	GSM611195
CDX2	ChIP-seq	GSM364858	**TFE3**	ChIP-seq	GSM979714
CTCF	ChIP-seq	GSE30203	**UTF1**	ChIP-seq	GSM970524
CTR9	ChIP-seq	GSE43231	**WDR5**	ChIP-seq	GSM566279
DPY30	ChIP-seq	GSM651192	**ZFX**	ChIP-seq	GSM288352
E2F1	ChIP-seq	GSM288349	**H2AZ**	ChIP-seq	GSM984544
ELL2	ChIP-seq	GSM749814	**H3K4me1**	ChIP-seq	GSM747542
ELL3	ChIP-seq	GSM935891	**H3K4me2**	ChIP-seq	GSM632035
ESRRB	ChIP-seq	GSM288355	**H3K4me3**	ChIP-seq	GSM970526
KAP1	ChIP-seq	GSM1406445	**H3K27ac**	ChIP-seq	GSM851278
KDM5B	ChIP-seq	GSE31968	**H3K27me3**	ChIP-seq	GSE33252
KLF4	ChIP-seq	GSM288354	**H3K36me3**	ChIP-seq	GSM801982
KLF5	ChIP-seq	GSM1208218	**H3K9me3**	ChIP-seq	GSM307621
MBD3	ChIP-seq	GSE31690	**VZ E14.5**	RNA-seq	GSE38805
MCAF1	ChIP-seq	GSM656526	**SVZ E14.5**	RNA-seq	GSE38805
MED1	ChIP-seq	GSM560347	**CP E14.5**	RNA-seq	GSE38805
MED12	ChIP-seq	GSM1267725	**Heart E11**	RNA-seq	GSE49847
NANOG	ChIP-seq	GSM1082342	**MEF E13.5**	RNA-seq	GSM723775
NFYA	ChIP-seq	GSE25533	**Lung E14.5**	RNA-seq	GSM929710
NIPBL	ChIP-seq	GSM560349	**Pancreas E13.5**	RNA-seq	GSE40823
Oct4	ChIP-seq	GSM566277	**CD4+ T-lymphocytes H3K27ac**	ChIP-seq	GSM1301745
P300	ChIP-seq	GSE41545	**Liver H3K27ac**	ChIP-seq	GSM851275
PHF19	ChIP-seq	GSM1020086	**Fibroblast H3K27ac**	ChIP-seq	GSM801538
RBBP	ChIP-seq	GSM566278	**NIH3T3 H3K27ac**	ChIP-seq	GSM1246689
RYBP	ChIP-seq	GSM1041375	**CD4+ T-lymphocytes H3K27me3**	ChIP-seq	GSE53831
RNAPol	ChIP-seq	GSM723019	**Liver H3K27me3**	ChIP-seq	GSM1112814
SA1	ChIP-seq	GSE36561	**Fibroblast H3K27me3**	ChIP-seq	GSM640759
SIN3A	ChIP-seq	GSM611196	**NIH3T3 H3K27me3**	ChIP-seq	GSM1246690
SMC1	ChIP-seq	GSM766455	**Input NIH-3T3**	ChIP-seq	GSM1246672
SMC3	ChIP-seq	GSE22562	**Input CD4 + T-lymphocytes**	ChIP-seq	GSE53831
SOX2	ChIP-seq	GSM1082341	**Input Liver**	ChIP-seq	GSM647039
SP1	ChIP-seq	GSM1267913	**Input Fibroblast**	ChIP-seq	GSM593399
SUZ12	ChIP-seq	GSE39513	**EMT timecourse**	RNA-seq	GSE54133

8.6 RNA-seq expression of marker genes for neuronal subtypes in A2lox.NeuroD1 mESCs

RNA-seq expression data derived in A2lox.NeuroD1 mESCs after 48 h +/- Dox induction.

Neuron subtype	Marker gene	Gene ID	Name	Minus DOX	Plus DOX	fold change	log2 fold change	p-value	adjusted p-value
glutamatergic neurons	*vGlut2*	NM_080853	solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 6 (Slc17a6)	0,464125953	133,9625381	288,6340168	8,173097528	1,11569E-07	8,91139E-07
	Gria1	NM_14799	glutamate receptor, ionotropic, AMPA 1	2,466068517	17,47788027	7,087345768	2,825245435	0,001743336	0,006469132
	Gria2	NM_14800	glutamate receptor, ionotropic, AMPA2 (alpha 2)	2,611507269	57,27756739	21,932762	4,455015597	7,10943E-11	8,41851E-10
	Grik2	NM_14806	glutamate receptor, ionotropic, kainate 2 (beta 2)	9,254709362	93,24136423	10,07501809	3,332710524	2,62568E-16	5,2753E-15
	Grik5	NM_14809	glutamate receptor, ionotropic, kainate 5 (gamma 2)	524,6801294	723,5351475	1,379002381	0,463624948	0,003602714	0,012321067
	Grin1	NM_14810	glutamate receptor, ionotropic, NMDA1 (zeta 1)	76,8395306	63,14009593	0,821713712	-0,283292253	0,314098951	0,517730898
	Grin2c	NM_14813	glutamate receptor, ionotropic, NMDA2C (epsilon 3)	17,48063775	94,09540814	5,382836113	2,428366502	7,61863E-12	1,00863E-10
	Grm3	NM_108069	glutamate receptor, metabotropic 3	0,464125953	3,469003829	7,474272461	2,901933156	0,265782468	0,455459824
	Grm5	NM_108071	glutamate receptor, metabotropic 5	0,536845329	0,95318363	1,775527473	0,828247683	1	1
GABAergic neurons	*Gad2*	NM_008078	glutamic acid decarboxylase 2	24,16874952	17,25644506	0,713998258	-0,48600754	0,416050225	0,634310062
	Gat3	NM_144512	solute carrier family 6a13 (Slc6a13)	2,93019447	2,427246116	0,828356664	-0,271676015	0,999564202	1
	Calb1	NM_009788	calbindin 1	2,538787893	2,515820198	0,990953283	-0,013111049	1	1
	Calb2	NM_007586	calbindin 2, calretinin	21,80321007	11,88107397	0,544923153	-0,875875306	0,151735576	0,296376111
	Gabra2	NM_14395	gamma-aminobutyric acid (GABA) A receptor, subunit alpha 2	19,33714156	3,857021562	0,199461826	-2,325815433	0,003728865	0,012697865
	Gabrb2	NM_14401	gamma-aminobutyric acid (GABA) A receptor, subunit beta 2	1,465097235	22,28808546	15,21270051	3,927204374	1,82021E-05	0,000102808
	Gabrb3	NM_14402	gamma-aminobutyric acid (GABA) A receptor, subunit beta 3	322,6421497	268,1550525	0,831122198	-0,266867487	0,085868298	0,187970558
	Gabrg3	NM_14407	gamma-aminobutyric acid (GABA) A receptor, subunit gamma 3	18,07310247	3,424716788	0,189492468	-2,399787591	0,192402173	0,356691231
other neuronal subtypes	*Adra2b*	NM_009633	adrenergic receptor	19,62801906	21,5457714	1,097704834	0,134490174	0,817712923	0,965936055
	Tph1	NM_009414	tryptophan hydroxylase 1	6,078547068	8,500644225	1,398466464	0,483845658	0,647955858	0,851291819
	Th	NM_009377	tyrosine hydroxylase	4,686169209	7,370312431	1,572779834	0,653316729	0,529746785	0,745773501
	Chat	NM_009891	choline acetyltransferase	3,002913846	55,62635674	18,52412676	4,21133363	1,47847E-06	9,92321E-06
	Mnx1	NM_019944	motor neuron and pancreas homeobox 1	3,931165752	11,22733399	2,855980821	1,513986291	0,164155375	0,315224528

8.7 Genes of the NeuroZincOme

2610008E11Rik, 3110052M02Rik, 5730507C01Rik, 5730601F06Rik, A030009H04Rik, A430033K04Rik, Adnp, Agap1, Agap2, Agap3, Anubl1, Araf, Asap3, Asxl3, Atmin, Atrx, Atxn7l3, AW146154, Baz2b, Bcl11a, Bcl11b, Birc2, Birc7, Bmi1, Bsn, C030039L03Rik, Cbfa2t2, Cbfa2t3, Chd3, Chd5, Chfr, Chn1, Chn2, Cul9, Cxxc4, D3Ertd254e, Ddx59, Deaf1, Dgkb, Dgke, Dhx57, Dnajc24, Dnlz, Drp2, Dtna, Dtx1, Dtx3, Dtx4, Dzip1, Dzip3, Ebf1, Ebf2, Ebf3, Ebf4, Ehmt1, Etohi1, Fam164a, Fbxl19, Fbxo11, Fezf1, Fezf2, Fhl3, Flywch1, Foxp2, Foxp3, Foxp4, Fus, G2e3, Git1, Gli2, Gm13251, Gm14326, Gm14420, Hdac2, Hivep3, Ikzf4, Ing1, Ing2, Insm1, Insm2, Irf2bp1, Jazf1, Kat5, Kdm4b, Kdm5b, Klf12, Klf8, Ksr2, Limk1, Lnx1, Lonrf2, Magee1, Magel2, Matr3, Mbtd1, Mex3a, Mex3b, Mex3c, Mex3d, Mkrn3, Mll5, Morc1, Mtf2, Mtmr4, Myrip, Myst4, Myt1, Myt1l, Mzf1, Nanos1, Nanos2, Nbr1, Neurl1a, Nr1i3, Nr2e1, Nr2f1, Nr4a2, Nufip1, Otud7a, P4htm, Park2, Parp1, Pcgf2, Pclo, Pdcd2, Pdzrn4, Pgr, Phc1, Phc2, Phf14, Phf21a, Phf21b, Phf6, Pias2, Plekhm3, Pnma3, Prdm10, Prdm12, Prdm8, Prkcb, Prkcz, Prkrir, Pygo1, Rai1, Rasgrp1, Rbm22, Rbm4b, Rbm6, Rc3h2, Rims1, Rims2, Rnf103, Rnf112, Rnf113a1, Rnf113a2, Rnf114, Rnf115, Rnf14, Rnf146, Rnf152, Rnf157, Rnf165, Rnf166, Rnf180, Rnf182, Rnf2, Rnf20, Rnf208, Rnf220, Rnf24, Rnf32, Rnf34, Rnf44, Rnft2, Rorb, Rspry1, Rufy2, Runx1t1, Rxrg, Sall1, Sall2, Sall3, Sap30, Sap30l, Scrt2, Sh3rf1, Sh3rf3, Slu7, Slx1b, Sobp, Sp4, Sp5, Sp9, Srek1ip1, St18, Stac, Taf1b, Tax1bp1, Tet1, Thap1, Thra, Traf3, Trafd1, Trim13, Trim2, Trim24, Trim3, Trim32, Trim33, Trim36, Trim37, Trim43c, Trim46, Trim67, Trim8, Trim9, Trip4, Tshz1, Tshz2, Tshz3, Ttc3, Uhrf2, Unc13a, Unc13c, Unk, Usp22, Usp27x, Usp3, Usp33, Usp49, Usp51, Vav2, Vav3, Vezf1, Wbp4, Whsc1l1, Wiz, Wrnip1, Xpa, Xrn2, Yaf2, Zbed4, Zbtb12, Zbtb16, Zbtb26, Zbtb33, Zbtb39, Zbtb41, Zbtb44, Zbtb45, Zbtb49, Zbtb5, Zc3h12a, Zc3h15, Zc3h3, Zc4h2, Zcchc11, Zcchc12, Zcchc18, Zdhhc2, Zfand5, Zfhx2, Zfhx3, Zfhx4, Zfml, Zfp1, Zfp105, Zfp113, Zfp119b, Zfp12, Zfp120, Zfp128, Zfp14, Zfp148, Zfp157, Zfp160, Zfp161, Zfp167, Zfp182, Zfp184, Zfp189, Zfp191, Zfp2, Zfp202, Zfp238, Zfp239, Zfp263, Zfp27, Zfp273, Zfp277, Zfp28, Zfp280d, Zfp281, Zfp282, Zfp286, Zfp287, Zfp292, Zfp3, Zfp317, Zfp334, Zfp341, Zfp352, Zfp354b, ***Zfp354c****, Zfp362, Zfp37, Zfp386, Zfp389, Zfp397, Zfp40, Zfp408, Zfp41, Zfp418, Zfp423, Zfp426, Zfp428, Zfp442, Zfp444, Zfp445, Zfp451, Zfp454, Zfp455, Zfp458, Zfp462, Zfp493, Zfp51, Zfp512, Zfp513, Zfp518a, Zfp518b, Zfp532, Zfp536, Zfp558, Zfp563, Zfp57, Zfp575, Zfp579, Zfp580, Zfp583, Zfp59, Zfp597, Zfp60, Zfp606, Zfp607, Zfp608, Zfp609, Zfp612, Zfp617, Zfp618, Zfp619, Zfp62, Zfp639, Zfp641, Zfp644, Zfp647, Zfp654, Zfp658, Zfp663, Zfp68, Zfp689, Zfp691, Zfp697, Zfp709, Zfp711, Zfp712, Zfp715, Zfp72, Zfp738, Zfp740, Zfp746, Zfp763, Zfp770, Zfp773, Zfp775, Zfp78, Zfp788, Zfp790, Zfp791, Zfp799, Zfp804a, Zfp804b, Zfp81, Zfp810, Zfp811, Zfp82, Zfp820, Zfp821, Zfp825, Zfp846, Zfp866, Zfp867, Zfp868, Zfp879, Zfp882, Zfp90, Zfp93, Zfp930, Zfp931, Zfp932, Zfp933, Zfp935, Zfp936, Zfp940, Zfp942, Zfp943, Zfp945, Zfp946, Zfp951, Zfp952, Zfp953, Zfp954, Zfp955b, Zfp958, Zfp960, Zfr, Zfr2, Zfyve28, Zglp1, Zic1, Zic2, Zic3, Zic4, Zic5, Zik1, Zkscan2, Zkscan4, Zmat1, Zmat4, Zmiz1, Zmiz2, Zmym1, Zmym2, Zmym3, Zmym5, Zmym6, Zmynd11, Zmynd8, Znhit1, Znrf1, Zscan12, Zscan18, Zscan21, Zswim5, Zswim6, Zxdb.*

9 Acknowledgements

removed due to data protection

10 CV

removed due to data protection