UNIVERSITÄT
DES
SAARLANDES

Exploring cell type diversity by comprehensive epigenome and transcriptome analysis

Dissertation

zur Erlangung des Grades des Doktors der Naturwissenschaften

der Naturwissenschaftlich-Technischen Fakultät der Universität des Saarlandes

von

Anna Welle

Saarbrücken

2019

Tag des Kolloquiums: 05.03.2020

Dekan:	Prof. Dr. Guido Kickelbick
Berichterstatter:	Prof. Dr. Jörn Walter,
	Prof. Dr. Frank Kirchhoff,
	Prof. Dr. Erik Boddeke
Vorsitz:	Prof. Dr. Alexandra Kiemer
Akad. Mitarbeiter:	Dr. Eva Steinmetz

Danksagung

Eine wissenschaftliche Arbeit ist nie das Werk einer einzelnen Person, deshalb möchte ich mich an dieser Stelle bei all denjenigen bedanken, die mich während der Anfertigung dieser Doktorarbeit unterstützt und motiviert haben. Allen voran danke ich meinem Doktorvater Prof. Dr. Jörn Walter, der mich stets unterstützte und dadurch diese Arbeit ermöglichte. Ich danke ihm für die Freiheiten, die er mir gegeben hat, eigene Ideen und Interessen zu verfolgen. Aber auch für die zahlreichen konstruktiven Gespräche und Ratschläge, die mich zu neuen Erkenntnissen führten oder vor Sackgassen bewahrten.

Prof. Dr. Frank Kirchhoff danke ich für die freundliche Übernahme des Zweitgutachtens. Ebenso bedanke ich mich bei Prof. Dr. Frank Kirchhoff für die hervorragende Zusammenarbeit und die stete Diskussionsbereitschaft.

Prof. Dr. Simon Fillatreau danke ich für die erfolgreiche Zusammenarbeit, die von wertvollen Diskussionen und gegenseitiger Unterstützung geprägt war. Bei Dr. Andreia Lino bedanke ich mich für ihren Einsatz bei der Isolierung zahlreicher B-Zell Proben. Bei Dr. Carmen Kasakow und Dr. Laura Stopper bedanke ich mich für die Unterstützung bei der Isolierung der Astrozyten und der Durchführung der Immunofluoreszenz Färbungen. Dr. Elmar Krause von der FACS facility Homburg möchte ich für die Einführung in den Umgang mit dem FACS Gerät danken.
Besonderen Dank auch an die gesamte Arbeitsgruppe für die freundschaftliche Arbeitsatmosphäre und stete Hilfsbereitschaft, die wesentlich zum Gelingen dieser Arbeit beigetragen haben. Ein ganz großer Dank gilt Dr. Gilles Gasparoni, der mir die Grundlagen in R beibrachte und es mir somit ermöglichte den Großteil der Daten selbstständig auszuwerten. Des Weiteren waren die vielen Gespräche rund um das Thema Datenanalyse mit Dr. Karl Nordström und Abdulrahman Salhab äußerst hilfreich und erhellend. Beate Schmidt, Christina LoPorto und Jasmin Kirch danke ich für ihr Engagement im Labor. Meinen ehemaligen und den derzeitigen Mit-Doktoranden danke ich für den regen Erfahrungsaustausch aber auch ganz besonders für die lustige und schöne Zeit.

Außerdem möchte ich Gideon Gießelmann für das Korrekturlesen meiner Doktorarbeit danken.

Ein spezieller und großer Dank gilt meinen Eltern und meinem Ehemann, die mich immer unterstützen und motivieren.

Für meinen Sohn David,

der mittendrin statt nur dabei war.

Abstract

The mammalian organism comprises hundreds of different, and highly specialized cell types. They share identical DNA sequences and yet differ in their phenotypes and functionalities. This cellular diversity is governed by regulatory mechanisms and is imprinted in the epigenome of each cell.

This work focused on the epigenomic profiles of astroglia subpopulations from adult murine brain, as well as the methylomes of a multitude of B-cell subpopulations and integrated gene expression profiles to unravel regulatory mechanisms in the establishment of cell diversity.

The comprehensive analysis of transcriptomes, methylomes and open chromatin sites of astroglia populations from the cortex and the cerebellum revealed shared epigenetic programs but also highlighted strong differences in chromatin organization and local epigenetic signatures connected to specific regionally expressed transcription factor networks.

The molecular characterization of Lag3-expressing and Lag3-non-expressing plasma cells confirmed the immuno-regulatory function of Lag3-expressing plasma cells and outlined unique transcriptional and epigenetic signatures. Moreover, DNA methylation signatures shed light on the cell ontogeny of Lag3-expressing plasma cells.

In summary, this work showcases approaches for the characterization and interpretation of epigenetic signatures to enhance our understanding of epigenetic gene regulation.

Zusammenfassung

Ein Säugetierorganismus besteht aus hunderten von unterschiedlichen und hochspezialisierten Zelltypen. Diese besitzen dieselbe DNA Sequenz, unterscheiden sich jedoch in ihrem Phänotyp und in ihrer Funktionalität. Diese zelluläre Diversität wird durch regulatorische Mechanismen bestimmt und ist im Epigenom jeder Zelle eingeprägt.

Die vorliegende Arbeit befasst sich mit epigenomischen Profilen von astroglialen Subpopulationen aus dem adulten Mausgehirn, sowie einer Vielzahl von B-Zell Subpopulationen zusammen mit den jeweiligen Genexpressionsprofilen.

Die umfassende Analyse der Transkriptome, Methylome und offenen Chromatinstellen von astroglialen Zellen aus der Großhirnrinde und dem Kleinhirn enthüllte gemeinsame epigenetische Programme aber auch Unterschiede in der Chromatinorganisation und lokalen epigenetischen Signaturen, die mit Regionen-spezifischen Transkriptionsfaktor Netzwerken verbunden waren.

Lag3-exprimierende und Lag3-nicht-exprimierende Plasmazellen wurden auf molekularer Ebene charakterisiert. Dies bestätigte zum einen die immuno-regulatorische Funktion der Lag3-exprimierenden Plasmazellen und enthüllte deren einzigartige Signaturen. Des Weiteren beleuchten die DNA-Methylierungsprofile die Zellontogenese der Lag3-exprimierenden Plasmazellen.

Zusammengefasst demonstriert diese Arbeit Möglichkeiten und Herangehensweisen zur Charakterisierung und Interpretierung von epigenetischen Profilen, um unser Verständnis der epigenetischen Genregulierung zu erweitern.

Contents

Chapter 1.

Introduction

1.1 Cell diversity arises from epigenetic diversity

In the 19th century scientists became aware that cells are the structural units of all organisms and established the cell theory (Remak 1852; Schleiden, Schwann, and Smith 1847). This awareness led to investigation of the vast diversity of cell morphology and cell function and in turn the term *cell type* arose. However, until today there is no uniform definition of what a cell type is (Clevers et al. 2017). It is generally known that multiple cellular characteristics like different morphologies, cyto-anatomies, molecular and biochemical properties should be considered when defining a cell type (Armañanzas and Ascoli 2015; Maclean and Hall 1987; Poulin et al. 2016; Vickaryous and Hall 2006). Additionally, cells from the same cell type can undergo dynamic changes in their molecular features determined by extracellular signals, aging, or disease. In this case it is more appropriate to define cell states rather than distinct cell types (Tasic, Levi, and Menon 2017). Current estimations suggest 411 distinct cell types in the human body (Vickaryous and Hall 2006). Intriguingly, this multitude of distinct cell types originates from a single totipotent cell, the zygote. Hence, all cells in an organism share the same DNA sequence which contains the fundamental information for cells to synthesize proteins. The synthesized proteins and the amount of the respective proteins will in turn form cell identity features like morphology and function of the respective cell. Since the genomic information is the same among each cell in a multicellular organism, mechanisms that regulate cell-specific gene expression enable the establishment of cell diversity. The bridge between the genome and gene expression is the *epigenetics* and is defined as "the study of any potentially stable and, ideally, heritable change in gene expression or cellular phenotype that occurs without

changes in Watson-Crick base-pairing of DNA" (Goldberg, Allis, and Bernstein 2007). Epigenetic mechanisms involve modifications of DNA bases, modifications of histone residues, nucleosome positioning, chromatin compartmentalization, and posttranscriptionally regulation of messenger RNA (mRNA) by micro RNAs (miRNA). The entity of this regulatory machinery is described as the *epigenome* of a cell. The epigenome of each cell orchestrates the continuous spectrum of transcripts in differentiating cells and cell-specific gene expression in terminally differentiated cells. In 1957, before the emergence of sophisticated techniques to study epigenetics, Conrad Waddington proposed the concept of the epigenetic landscape. His metaphor depicts the progression of a cell from an undifferentiated state to the state of terminally differentiation undergoing distinct intermediate fates (Waddington 1957). In his analogy a marble rolling down the landscape represents the differentiation of a totipotent zygote (**Figure 1-1**). As the marble rolls down the landscape the valleys bifurcate into new valleys symbolising commitment to alternative cell lineages. The downward paths become more limited indicating the progressed differentiation of the cell.

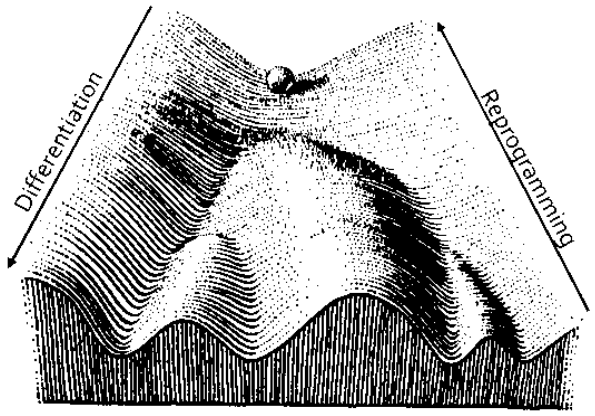


Figure 1-1 Epigenetic landscape*. The figure represents the cell fate commitment during differentiation. Each path will lead to distinct cell types determined by molecular factors symbolized in the topology of the landscape. Figure adapted from Waddington 1957.*

Intense research efforts and methodological advances in the epigenetic field enhanced the understanding of the complex processes in establishing the epigenetic landscape. This knowledge can be used to target cell fate determining pathways, which would correspond to a position change in the epigenetic landscape ("reprograming"). For instance, differentiated fibroblasts can be reprogrammed into induced pluripotent stem (iPS) cells by overexpression of four transcription factors (Oct4, Sox2, Klf4, cMyc) involved in developmental signaling pathways (Takahashi and Yamanaka 2006; Xiaosong Liu et al. 2008). Human iPS cells bear a great potential in the regenerative medicine and became a popular model for pluripotency and diseases. The generation of iPS cells is only one of many examples to highlight the relevance and advancement that epigenetic studies entail. The following passage provides a more detailed background on epigenetic mechanisms in the regulation of cell-specific gene expression and outlines the methodology to generate epigenetic data.

1.1.1 Epigenetic gene regulation

The initial step in gene expression is the transcription of DNA into mRNA by the RNA polymerase II (RNAPII). This requires a tight regulation by the epigenetic machinery which enables or disables access of RNAPII to the DNA. In fact, 80.4% of the genome is estimated to be associated with regulatory functions, whereas only 2.94% of the human genome encodes proteins (The ENCODE Project Consortium 2012). RNA interference, organization of chromatin, and the modification of DNA play central roles in dynamic and specific gene regulation.

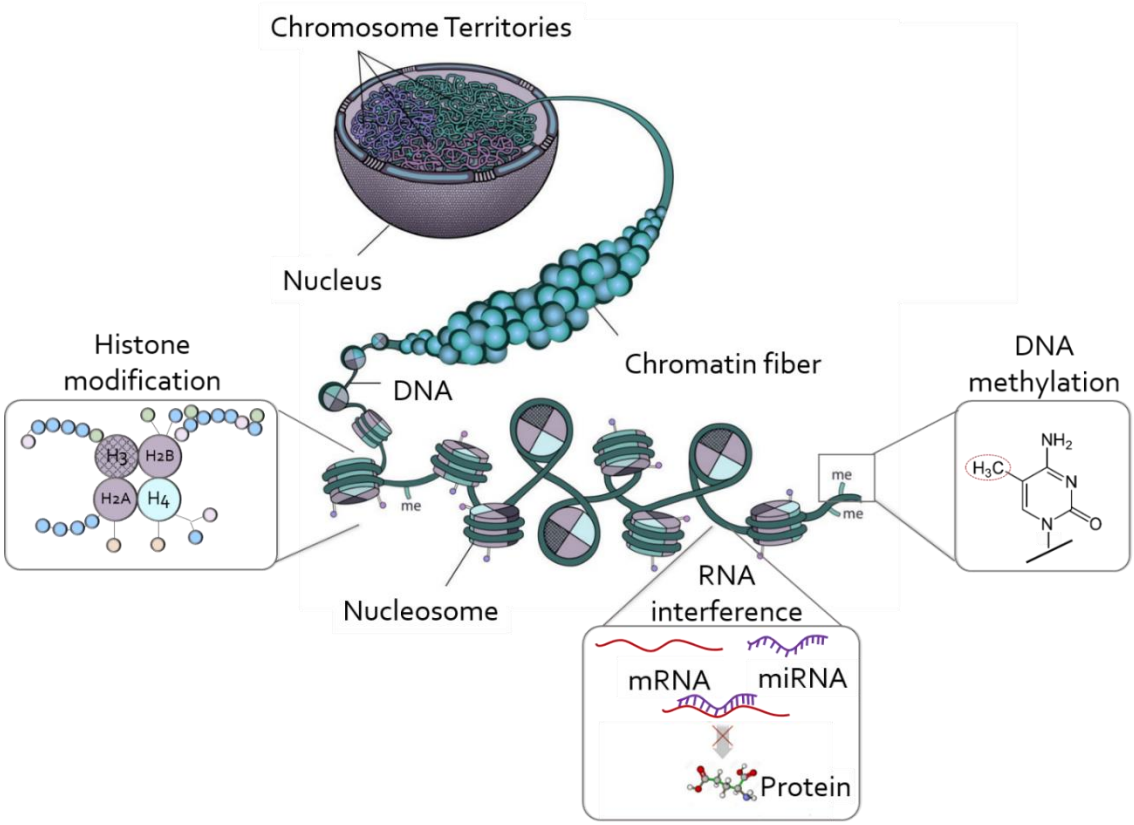


Figure 1-2 Overview of chromatin organization and epigenetic mechanisms. *DNA is wrapped around a core histone complex which is further arranged into a chromatin fiber. The chromatin fibers form higher order conformation of the chromatin resulting in a highly compacted form of DNA. Modifications of histone residues, DNA methylation, higher order organization of the chromatin, and RNA interference represent the major components of epigenetic regulation of gene expression. Figure adapted from Rosa and Shaw 2013 and American Society of Hematology. "ASH agenda for hematology research."*

Chromatin organization

Historically, the term *chromatin* was introduced by Emil Heitz in 1928 based on the observation of distinct DNA properties in chromosomal staining (Jost, Bertulat, and Cardoso 2012). Densely compacted DNA appeared darker and was referred to as *heterochromatin*, whereas the lighter compartments of the DNA was termed *euchromatin*. This categorization is nowadays used to describe functionally distinct compartments of the genome. Euchromatin represents protein-accessible DNA and is

transcriptionally active, while heterochromatin is condensed, has low gene density and is in general transcriptionally inactive.

Besides the regulatory aspect, chromatin organization is necessary for packaging 2 meters of DNA into a nucleus of 5-10 µm. For this purpose, 146 base pairs of DNA is wrapped 1.7 turns around a histone octamer, which consists of two molecules of each of the histone proteins H3, H4, H2A, and H2B. This complex forms the basic unit of the chromatin and is referred to as the nucleosome (**Figure 1-2**).

This structure constitutes the first level of DNA-compaction into the nucleus. Of note, the histones H3, H4, H2A and H2B build the core histone molecule in the S-phase of the cell cycle. However, many other replication-independent histone variants are assembled into the chromatin throughout the cell cycle (Kamakaka and Biggins 2005; Lyons et al. 2016). The incorporation of histone variants increases the diversity of nucleosomes and therefore increases the complexity and plasticity of chromatin. Besides the core histones, the linker histone H1 binds to the nucleosome at the entry and the exit sites of the DNA, keeps the integrity of nucleosome assembly, and enables compaction and stabilization of the chromatin structure. Furthermore, H1 is required in spacing nucleosomes and in the formation of a chromatin fiber. The fiber is further arranged into loops that are assembled into Topological Associating Domains (TADs), ultimately forming the chromosome territories (**Figure 1-2**) (Cremer and Cremer 2010; Nora et al. 2012; Dixon et al. 2012; Gibcus and Dekker 2013). TADs are characterized by high-frequency chromatin interactions in spatial neighborhoods which are similar from one cell type to another and are conserved between human and mouse (Phillips-Cremins et al. 2013; Dixon et al. 2012; Nora et al. 2012). In contrast to TAD formation, dynamic local DNA configuration within the TADs ("sub-TAD") is cell type-specific (Berlivet et al. 2013; Phillips-Cremins et al. 2013).

The cell type-specific sub-TAD organization and the factors responsible for such chromatin compartmentalization resulting in unique transcriptional outputs are subjects of intensive research. Recently developed chromatin conformation capture technologies in combination with next generation sequencing (NGS) have provided important insights into the role of insulators in 3D genome organization (Dixon et al. 2012; Le Gall, Valeri, and Nollmann 2015). In the mammalian genome the protein CTCF and the protein complex cohesin were found to bind TAD boundaries and affecting the

sub-TAD architecture (Phillips-Cremins et al. 2013; Sofueva et al. 2013; Zuin et al. 2014). This led to the current model of loop-extruding TAD formation, in which cohesin progressively forms DNA loops and stalls when it reaches CTCF-bound TAD boundaries. In contrast to the cohesin-CTCF interaction, DNA contacts within the loop are arranged by the interaction of cohesin with cell type-specific transcription factors (**Figure 1-3**). For instance, in murine embryonic stem cells pluripotency factors like Oct4, Sox2 and Nanog were found to bind to the same DNA regions as cohesin (Fudenberg et al. 2016). Emerging evidence, provided by chromatin conformation capture studies, links cell type-specific chromatin structure with gene expression by demonstrating the physical contact of distal regulatory elements like enhancers and promoters (**Figure 1-3**).

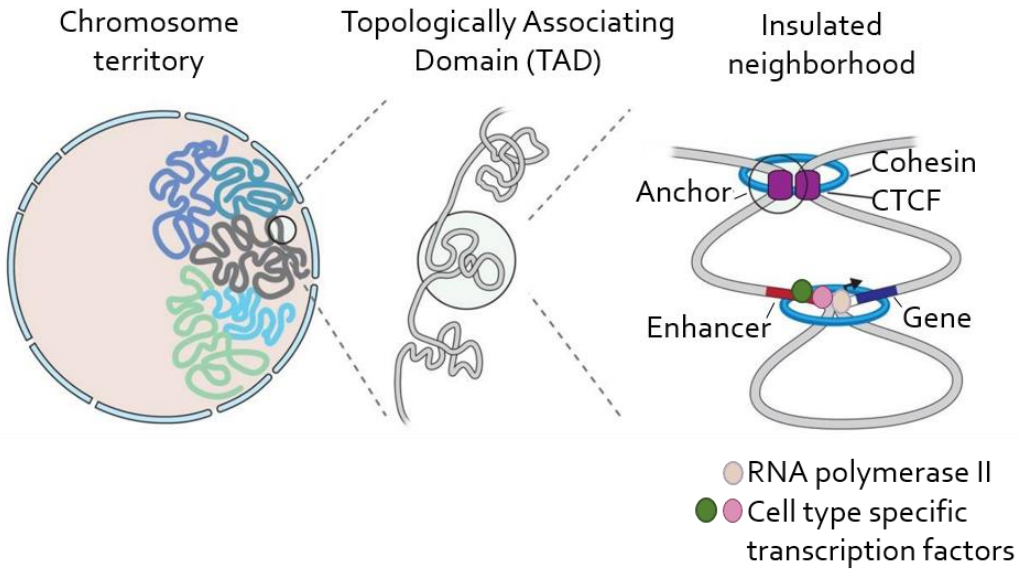


Figure 1-3 Schematic representation of chromatin organization in mammals*. Regulatory elements like promoters and enhancers are physically connected by the binding of cell type-specific transcription factors to enhancers and recruitment of RNA polymerase II. Cohesin complex together with CTCF stabilize the long range connection. (Hnisz, Day, and Young 2016)*

Regulatory elements

The diversity and the identity of cells is shaped by the combinatorial interaction of *cis*-regulatory elements (CREs), present in the non-coding parts of the genome, with *trans*-acting factors like transcription factors (TFs) and miRNAs. CREs are DNA sequences classified into distinct functional groups. CREs include binding sites for RNAPII at the transcription start site (TSS) or at close proximity to the TSS, called promoters. The binding of RNAPII to the promoter initiates the transcription of a gene into mRNA, however this binding alone is not sufficient to regulate dynamic and specific gene expression in mammalian cells. It requires the interaction of further CREs such as enhancers, silencers, and insulators (**Figure 1-3**). Among the CREs, enhancers have a major contribution to the initiation and regulation of gene expression. These DNA segments, typically located in intergenic and intronic regions, recruit TFs through a short (6 to 12 nt) TF-specific DNA sequence (TF binding site, TFBS). Importantly, TFs bind to enhancers in a spatiotemporal and combinatorial manner. Hence, the function of an enhancer deviates from cell type to cell type depending on the context-dependent occupancy by multiple TFs (Spitz and Furlong 2012). Moreover, a single enhancer can interact with multiple promoters and multiple enhancers can physically contact a single gene, leading to a complex regulatory network. Since enhancers and promoters are genomic features present in each cell, the usage of these elements has to be dynamically and specifically controlled in each cell. This is achieved by selectively facilitate or impede access of TFs to these loci through epigenetic modification. The following sections outline these mechanisms.

Histone modification

The amino-terminal residues of core histones are flexible and extrude from the central protein domains. The amino acids of these "tails" are targets for post-translational modification by histone modifying enzymes. These modifications include e.g. acetylation of lysines, methylation of lysines and arginines, ubiquitination of lysines, and phosphorylation of serines and threonines. In particular, acetylation and methylation of H_3 lysine (K) residues function positively or negatively in gene expression regulation. Acetyl groups are transferred onto lysines by the histone acetyltransferase (HAT) enzyme family. The negatively charged acetyl group reduces the interaction with the negatively charged DNA, which presumably results in a more accessible DNA, and thus correlates with active gene expression. The major targets of acetylation are H_3K_9, H_3K_{14}, H_3K_{18}, H_3K_{23}, and H_3K_{27} localized around the TSS of a gene or at enhancers. The removal of acetyl groups is catalyzed by histone deacetylases (HDACs). In contrast to lysine acetylation, lysine methylation is associated activation and repression of transcription depending on the lysine residue and degree of methylation. Methyl groups are transferred by histone methyltransferases (HMTs) in a successive manner resulting in lysine monomethylation (me1), dimethylation (me2), and trimethylation (me3). Trimethylation of lysine 4 of H_3 ($H_3K_4me_3$) occurs at promoters of genes and is generally associated with active transcription of the respective gene. Trimethylation of lysine 27 and/or lysine 9, on the other hand, confers a repressive chromatin state (Mikkelsen et al. 2007). The methyl groups are removed by histone demethylases (HDMs). In the case of acetylation, regulatory effects are exerted through influence on chromatin structure. However, the other modifications are involved in recruitment of effector molecules. For example, readers of lysine methylation recognize the degreed of methylation and contain protein domains called Chromo-, Tudor-, and PHD-domains. Whereas lysine acetylation is recognized by proteins containing a Bromo-domain. The binding of these effectors is assumed to remodel the chromatin and therefore participate in gene expression. The diverse modifications occur in a combinatorial manner which in turn leads to an interplay of the readers and writers of these modifications. This fundamental process was termed "histone code" and is studied extensively to reveal transcriptionally active or silent

chromatin (Jenuwein and Allis 2001). The complexity of the histone code is increased further by a "cross-talk" among the modifications. For example, H3 serine 10 phosphorylation promotes the acetylation of H3 lysine 14 by the acetyltransferase Gcn5 (Lo et al. 2000).

The usage of specific antibodies against histone marks enables to identify the genome wide location of respective modifications using an approach called chromatin immunoprecipitation coupled with NGS (ChIP-Seq) (Barski et al. 2007; Johnson et al. 2007; Robertson et al. 2007). The proteins associated with DNA are crosslinked to DNA and labeled with an antibody. The chromatin is subsequently fragmented, and the antibody-bound fragments are separated from the rest. By sequencing those antibody-bound fragments and aligning the sequencing reads to the reference genome it is possible to identify the presence of a histone modification in a tissue and/or a cell type. Similarly, TF binding sites can be determined using TF-specific antibodies. By combining the histone mark profiles with the entire gene expression of a cell type (transcriptome) it is possible to reveal cell type-specific gene regulation through the presence and location of histone modifications. For example, the analysis of 5 human cell lines showed a rather similar chromatin signatures at promoters and largely invariant CTCF binding sites. Interestingly, the cell type-specific expression pattern correlated with cell type-specific histone modification patterns at enhancer regions, highlighting the great relevance of enhancer activity for cell diversity (Heintzman et al. 2009).

Nucleosome positioning

The access to the chromatin for regulatory factors and transcriptional machinery plays a crucial role in various regulatory effects. As described above, lysine acetylation reduces the interaction of histones with DNA resulting in an accessible state of the DNA. Together with the modification of histones, the placement of nucleosomes around regulatory regions results in repression or activation of a gene. Specifically, promoters and TFBS have been shown to be nucleosome-depleted in order to facilitate the DNA binding and in turn activate gene expression. In contrast, the rest of the

genome tends to be occupied by nucleosomal arrays. The nucleosome positioning pattern is determined by the interplay of several factors, including the DNA sequence, the RNAPII transcription machinery, transcription factors, and ATP-dependent nucleosome remodelers (Yuan et al. 2005; Lee et al. 2007; Schones et al. 2008; Valouev et al. 2011). Remodeling complexes of the imitation switch (ISWI) and SWI/SNF classes rearrange nucleosomal position or initiate histone eviction in an ATP-dependent fashion, enabling the transition from closed to open chromatin (Dann et al. 2017; Erdel and Rippe 2011). These remodelers contain protein subunits that allow to recognize distinct histone modifications and to direct their activity to certain genomic loci. Besides the remodeling complexes, some transcription factors are able to bind their DNA recognition sites within closed chromatin. These *pioneer factors* initiate the remodeling of the nucleosomes which in turn leads to an open chromatin state and enables the binding of further TFs to activate gene expression (Magnani, Eeckhoute, and Lupien 2011). The first described pioneer factor activity was observed in mouse liver cells. Factors of the GATA family occupied liver-specific enhancers even before differentiation and activation of the target gene (Bossard and Zaret 1998; Gualdi et al. 1996). Pioneer factors therefore play an important role in establishing lineage-specific chromatin state and in turn the cell type-specific transcriptional programs.

DNA methylation

Together with histone modifications and chromatin organization, the addition of a methyl group to the 5′ carbon of cytosine is a key mechanism in gene regulation and cell type specification. Historically, the function of 5-methylcytosine (5-mC) was first described in 1975. It was proposed that 5-mC is an inheritable mark involved in gene silencing (Holliday and Pugh 1975; Riggs 1975). Later, it was revealed that 5-mC occurs predominantly symmetrically in the context of the dinucleotide CpG in mammals. Of note, the frequency of CpG sites in the genome is lower than the expected from the GC content of the DNA. This underrepresentation is caused by the high deamination rate of 5-mC which leads to mutations. Over a long time period this mutagenicity caused the depletion of CpGs in the genome (A. P. Bird 1980). Moreover, CpG sites are not equally distributed across the genome but are frequently concentrated in CpG dense

regions called *CpG islands* (CGIs) (Gardiner-Garden and Frommer 1987). On average, CGIs span regions of 1000 bp, are highly conserved between mice and humans, and coincide with approximately 70% of all annotated gene promoters (Saxonov, Berg, and Brutlag 2006). Methylation of cytosines directly impairs binding of transcription factors to their binding site or indirectly by recruiting methyl-binding proteins that occupy transcription factor binding sites. This mechanism is therefore associated with stable gene silencing. Interestingly, CGIs associated with promoters are usually unmethylated, whereas the majority of CpGs in a non-CGI context are methylated (A. Bird et al. 1985). This high methylation of the genome results from methylation of transposable elements, which ensures their stable inactivation (Schulz, Steinhoff, and Florl 2006). Moreover, DNA methylation plays a major role in X-chromosome inactivation and genomic imprinting (Hellman and Chess 2007; Reik and Walter 2001). Emerging evidence shows that DNA methylation also affects alternative splicing of precursor mRNA by demarcate exons and flanking introns (Shayevitch et al. 2018; Maor, Yearim, and Ast 2015).

Besides the methylation of cytosines in the CpG context, non-CpG methylation (mCpH[1], mainly mCpA) can be detected in embryonic stem cells and neuronal cells (Ramsahoye et al. 2000; Lister et al. 2009). Closer inspection of gene expression and genome wide mCpH distribution in mouse brain neurons from the cortex and embryonic (ES) stem cells revealed that mCpH is enriched in the gene body and depleted in protein binding sites and enhancers. Intriguingly, in ES cells intergenic CpH methylation positively correlates with gene expression whereas in neurons mCpH levels and gene expression correlate inversely (Lister et al. 2009, 2013).

DNA methyltransferases

The main enzymes establishing and maintaining DNA methylation patterns are DNA-methyltransferases (DNMTs). The *de-novo* methyltransferases DNMT3A[2] and DNMT3B[2] catalyze the transfer of the methyl group from S-andenosyl methionine

[1] H = A, C, or T
[2] Equivalent for mouse are Dnmt3a, Dnmt3b, Dnmt1, Dnmt3l.

(SAM) onto the 5' cytosine-carbon and thus establish new DNA methylation patterns during development/cell differentiation. The upregulation of Dnmt3a but not Dnmt3b in developing mouse brain coincides with the increase of mCpH which suggests methylation activity of Dnmt3a also on non-CpGs. In contrast, the maintenance methyltransferase DNMT1[2] is responsible for propagating the cellular methylation patterns after DNA replication by the recognition of hemimethylated DNA (Hermann, Goyal, and Jeltsch 2004; Okano et al. 1999). The third member of the DNA-Methyltransferases is DNMT3L, that lacks a catalytical domain but interacts with DNMT3A and DNMT3B and enhances their activity (Suetake et al. 2004).

DNA demethylation

In general, global DNA methylation level of somatic cells is high and static. However, during germ cell and pre-implantation development the methylation profiles are highly dynamic. A rapid loss of 5-mC was observed in the paternal pronucleus in the zygote before DNA replication commences (Mayer et al. 2000; Oswald et al. 2000). Similarly, primordial germ cells (PGCs) undergo a genome-wide rearrangement of the epigenome with rapid erasure of DNA methylation (Hajkova et al. 2002). During development, both zygote and PGCs acquire *de-novo* methylation. These DNA methylation kinetics suggest DNA demethylation mechanisms. Indeed, evidence for two mechanisms emerged during the last years, namely passive or active DNA demethylation. The passive demethylation characterizes the loss of 5-mC after successive round of replication and incomplete maintenance machinery activity. On the other hand, the active demethylation involves the consecutively enzymatic oxidation of the 5-mC methyl group to 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC), and 5-carboxycytosine (5-caC) (**Figure 1-4**). This Fe^{2+} and α-ketoglutarate dependent oxidation is catalysed by the members of the ten-eleven translocation (TET) family TET1[3], TET2[3], and TET3[3] (Tahiliani et al. 2009; Ito et al. 2010, 2011). Subsequently 5-fC and 5-caC can be excised by the thymine DNA glycosylase TDG and replaced by an unmodified cytosine (He et al. 2011; Maiti and Drohat 2011). Besides the role as

[3] Equivalent for mouse are Tet1, Tet2, and Tet3.

intermediate in the active demethylation, 5-hmC may play a role in epigenetic gene regulation by the interaction with DNA associated proteins or indirectly by impairment of the interaction of 5-mC with 5-mC binding proteins (Valinluck et al. 2004). Interestingly, the detected 5-hmC level is the highest in central nervous system, followed by medium levels in kidney, nasal epithelium, bladder, heart, muscle, lung, and low levels in liver, spleen, and testes (Globisch et al. 2010). The functional relevance of 5-hmC in different cell types is not yet fully understood and in the focus of current investigations. The same holds true for the distinct expression levels and distinct transcript variants of TETs in different tissues/ cell types.

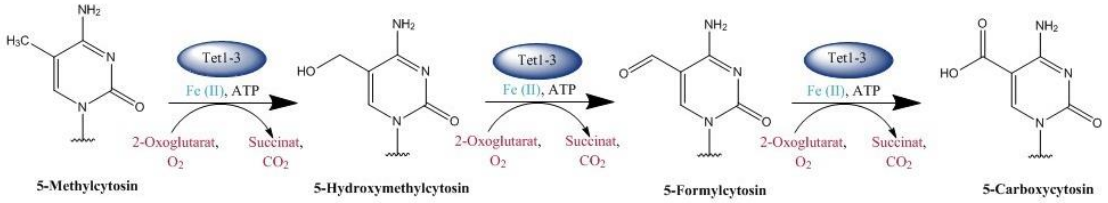


Figure 1-4 Cytosine modifications in mammals. *5-Methylcytosine can be oxidized to 5-hydroxycytosine and further to 5-formylcytosine and 5-carboxylcytosine by the 2-oxoglutarate- and iron(II) dependent enzymes Tet1, Tet2 and Tet3.*

Crosstalk of DNA methylation and histone modifications

DNA methylation and other epigenetic mechanisms like histone modifications cooperate with each other to regulate gene expression. It has been demonstrated that DNMTs interact with histone modifiers that mediate repressive modifications to establish a repressive state of the genomic region (Fuks et al. 2003, 2000; Geiman et al. 2004). In contrast, active histone marks like H3K4me3 impairs the binding of DNMTs and thus prevent DNA methylation at the respective locus (Yingying Zhang et al. 2010; Ooi et al. 2007). Important mediators of the crosstalk between DNA methylation and histone modifications are methyl-CpG binding proteins like MeCP2. MeCP2 conveys the crosstalk by binding to methylated DNA and recruitment of histone deacetylases (Jones et al. 1998; Nan et al. 1998)

Non-coding RNA

Non-coding RNAs (ncRNAs) function in the transcriptional and posttranscriptional regulation of gene expression and, unlike mRNA, are not translated into proteins. Epigenetic related ncRNA include miRNA, short interfering RNAs (siRNA), Piwi-interacting RNAs (piRNA), and long non-coding RNA (lncRNA). These ncRNAs have been shown to play a role in heterochromatin formation, histone modification, DNA methylation, and in gene silencing through interaction with mRNA. ncRNAs can be classified according to their length in short ncRNAs (<30 nts; miRNAs, siRNAs, piRNAs) and long ncRNAs (>200 nts; lncRNAs). miRNAs bind to mRNAs with partially complementary sequences and inhibit the translation of bound mRNA, induce mRNA degradation, or destabilize their target mRNA. siRNAs, found in plants, fungi and non-mammals, similarly interfere with mRNA processing by binding their target mRNA fully complementary. The third class of short non-coding RNAs, piRNAs, are mainly involved in suppression of transposable elements activity in germline cells (Morris and Mattick 2014). The majority of non-coding RNAs comprises long non-coding RNAs. lncRNAs span tens or hundreds of kilobases and harbor classical hallmarks of bona fide genes like conserved promoters, epigenetic modifications and regulation by trans acting factors (Mattick 2009). They are implicated with various processes including chromatin remodeling, transcriptional regulation, and post-transcriptional regulation. This repertoire of functions is achieved by the interaction of different protein complexes. For instance, lncRNAs can form complexes with chromatin modifying proteins or Dnmts and recruit the complex to a target sequence (Morris and Mattick 2014). The best studied example for lncRNA function is the X-chromosome inactivation by the lncRNA Xist. Xist RNA coats the inactive X-chromosome which recruits polycomb repressive complex 2 and results in extensive histone methylation and therefore silencing of the genes (Silva et al. 2003; Plath et al. 2003).

1.1.2 DNA methylation as hallmark of cellular memory

DNA methylation as a form of epigenetic memory in the central nervous system

DNA methylation, once established, is a stable mark and is propagated through multiple cell divisions. As mentioned in chapter 1.1.1 on page 12, DNA methylation is dynamic during development and cell differentiation. However, DNA methylation profiles are partially retained in the descendant cells. This *epigenetic memory* was comprehensively investigated in neural stem cells (NSCs) (Ziller et al. 2015; Mo et al. 2015; Sanosaka et al. 2017). NSCs have the capacity to self-renew and progressively give rise to the major cell types in the mammalian central nervous system (CNS) such as neurons, astrocytes and oligodendrocytes (Martynoga, Drechsel, and Guillemot 2012). At early gestation (before embryonic day 9 (E9)) NSCs divide symmetrically to expand the progenitor pool. During midgestation (E9-E10) the cells are programmed to differentiate into neurons but lack the potency to differentiate into glia cells. During this step NSCs develop into neural progenitor cells (NPCs). At late gestation (E18.5), NPCs begin to give rise to astrocytes and oligodendrocytes. The potency of NSCs/NPCs to give rise to distinct cell types in a temporarily coordinated fashion critically depends on epigenetic mechanisms. Promoters of genes involved in the astrogliogenesis onset, such as glial fibrillary acidic protein (*Gfap*) and S100 calcium-binding protein B (*S100b*), are methylated in the early and midgestation and become unmethylated at E18.5 (Namihira, Nakashima, and Taga 2004; Takizawa et al. 2001). Consistent with the differential methylation, the knock-out of Dnmt1 in NPCs leads to a global hypomethylation (including hypomethylation of gliogenic genes), which in turn leads to precocious onset of astrogliogenesis due to the expression of respective genes (Fan et al. 2005). Correspondingly, genes involved in the generation of neurons were unmethylated in NPCs and neurons but gained methylation upon differentiation into astrocytes (Sanosaka et al. 2017).

Interestingly, Sanosaka et al. 2017 showed that the hypomethylation in course of E11.5 to E18.5 NPC development was not significantly associated with gene expression change of associated genes. Moreover, the DNA methylation profiles of NSCs and NPCs were closely correlated to neurons, while the gene expression profile of NSCs and NPCs showed strong differences to neurons. This suggests that the

observed methylation profiles in the precursors reflects the differentiation potential rather than the functional properties. In line with this observation, lowly methylated regions associated with transcription factors crucial in early neurogenesis are established in NSCs/NPCs and retained in mature neurons. Respective genes are expressed in NSCs/NPCs but are downregulated in mature neurons. This demonstrated that DNA methylation at these loci mirrors the origin of the cell and reflects the transient expression of respective genes (Sanosaka et al. 2017).

Mo et al. 2015 identified a unique DNA methylation pattern as a signature of epigenetic memory from the comparison of fetal cortex with adult cell types of the CNS. They determine DNA methylation valleys (DMVs)[4] present in fetal and adult cortical cells at transient expressed developmental genes. Interestingly, comparison of the DMVs in the cell types revealed that DMVs remain unmethylated large regions but gain DNA methylation at the DMV boundaries during development. This hypermethylation was mostly pronounced at transcription factor expressed only in NPCs and immature neurons establishing neuronal subtype identity. This specific hypermethylation in neurons represents a signature of past gene expression (Mo et al. 2015).

Genome wide DNA methylation organization as a form of epigenetic memory

Besides local traces of epigenetic memory described for neuronal development, DNA methylation pattern at larger scales is a hallmark for lineage-specific epigenetic sate. The methylome can be classified into unmethylated regions (UMRs), lowly methylated regions (LMRs), partially methylated regions (PMDs) and fully methylated regions (FMRs) (Lister et al. 2009; Stadler et al. 2011). PMDs are large regions of 100 kb to 20 Mb, covering gene-poor and heterochromatic regions and represent up to 75% of the entire genome. In addition, PMDs are characterized by a reduced average DNA methylation. Comprehensive analysis of 195 methylomes by Salhab et al. 2018 identified PMD arrangements as a cell type classifier and showed further high similarities between PMD arrangements in cells of the same lineage. This demonstrated that genome wide methylome organization in the progenitors is

[4] Unmethylated genomic domains spanning at least 5kb of developmental genes (Xie et al. 2013).

maintained throughout development and can be used to assign the origin of a cell (Salhab et al. 2018).

1.1.3 Methods to profile epigenomes and transcriptomes

This chapter outlines general technical principles to generate epigenetic and transcriptomic data. With recent innovation in sequencing and computational technology various techniques evolved to map cell type-specific epigenomes (**Figure 1-5**). In the following paragraphs a general overview of the techniques will be given, and the methods used for this work will be described in more detail.

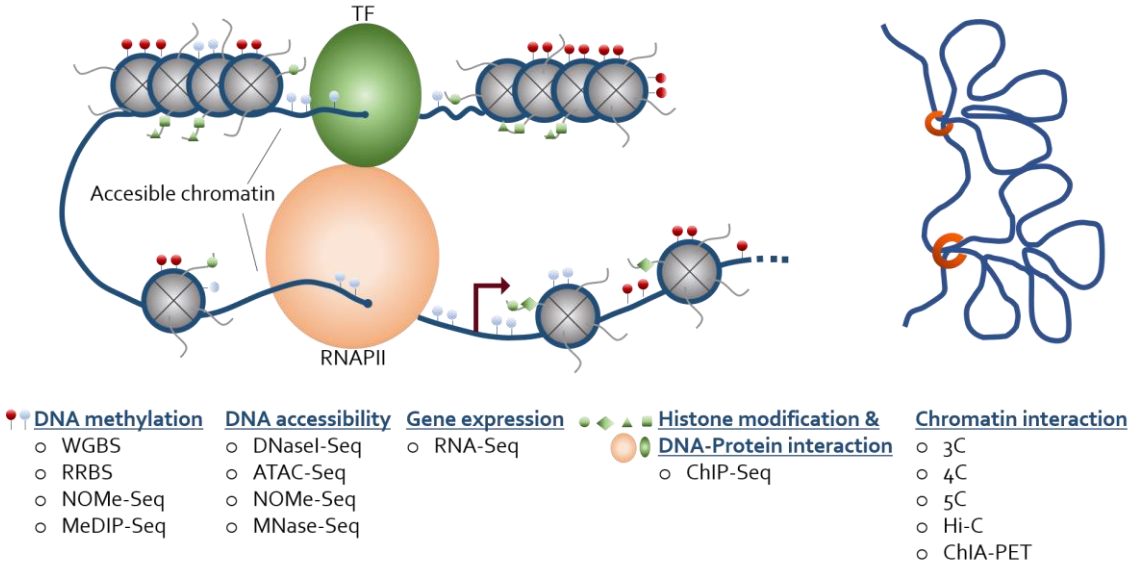


***Figure 1-5 Technology for profiling epigenomes and transcriptomes**. Different epigenetic and chromatin states can be quantified using color-corresponding technologies.*

Next Generation Sequencing (NGS)

In course of the Human Genome Project, sequencing of the human genome using Sanger sequencing was completed after 13 years of international collaboration. In contrast, using NGS technology it is now possible to sequence an entire human genome with high accuracy and dramatically decreased costs within a single day. The principle behind NGS is similar to Sanger sequencing. During DNA synthesis fluorescence

labeled deoxyribonucleotide triphosphates are incorporated into the newly synthesized strand based on the template. During each cycle the strand is extended by one base. The incorporation of the nucleotides leads to a signal emission of the respective fluorophore. The decisive advantage of NGS is that, instead of sequencing a single DNA fragment, it can detect the signals emitted from millions of DNA fragments in a massively parallel fashion. The workflow for all NGS application is in principle the same. The genome is fragmented randomly, or DNA fragments are enriched based on specific properties. Next, DNA fragments with known properties and sequences (sequencing adapters) are transferred onto the genomic fragments in order to generate libraries ready for sequencing. The libraries are subjected into a sequencing device and captured on a surface with immobilized oligonucleotides complementary to the sequencing adapters. After sequencing, the readout is translated into DNA sequences that can be aligned to the reference genome.

DNA accessibility

A widely used method to profile genome-wide DNA accessibility is the Assay for Transposase Accessible Chromatin Sequencing (ATAC-Seq, **Figure 1-6**) (Buenrostro et al. 2013). ATAC-Seq is based on the activity of the hyperactive, and sequencing-adapter preloaded *Tn5* transposase in the cell nucleus. In the nucleus it binds to open chromatin, fragments the DNA, and simultaneously inserts sequencing adapters. After purification of the DNA fragments, the tagged fragments are amplified by Polymerase Chain Reaction (PCR) using sequencing primers and sequenced using next generation sequencing. The enrichment of mapped reads at certain loci in the genome represents an open chromatin state at this region. The procedure in detail and the analysis of the data are described in chapter 2.2.4 on page 46 and chapter 2.2.9 on page 51. Besides ATAC-Seq, there are various other methods to analyze DNA accessibility, e.g. DNase Sequencing, MNase Sequencing, and FAIRE Sequencing (Boyle et al. 2008; Kaplan et al. 2009; Giresi et al. 2007). These methods require a high number of cells and adjustment of the reagents for each cell type. Hence, ATAC-Seq was chosen for the work in this thesis due to the low cell number requirement and the straightforward procedure.

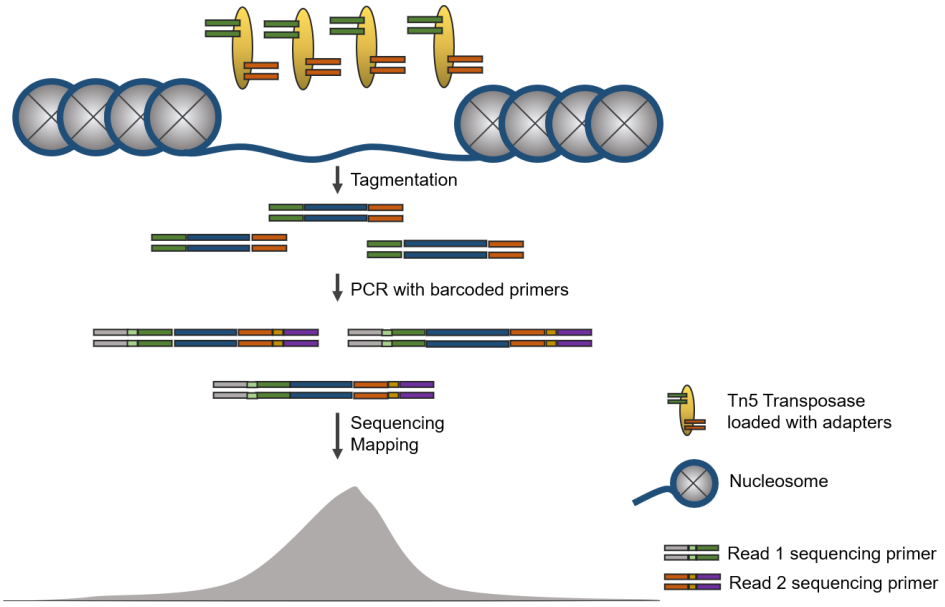


Figure 1-6 ATAC-Seq method to profile DNA accessibility*. The transposase introduces adapters into accessible regions, the DNA fragments are then amplified with sequencing primers and sequenced using next generation sequencing. The enrichment of sequencing reads at a certain genomic position indicates DNA accessibility at this locus.*

DNA methylation

A variety of sequencing-based methods emerged to profile DNA methylation. They include immunoprecipitation of DNA fragments containing 5-mC using anti-5mC antibodies (Methylated DNA Immunoprecipitation, MeDIP), methylation-sensitive cleavage, and sodium bisulfite treatment of the DNA. The respective methods differ in the resolution, number of analyzed CpGs and experimental costs. Currently it is not possible to directly distinguish between cytosine and 5-mC during sequencing. However, DNA methylation status at single base resolution can be explored using bisulfite sequencing. For this method, the DNA is treated with sodium bisulfite which converts cytosines into uracils that are read as thymines after PCR. In contrast to unmethylated cytosines, 5-mC is not converted by sodium bisulfite treatment and is read as a cytosine after PCR (**Figure 1-7**). This approach can be used to detect CpG methylation in targeted regions of interest or at a genome wide level. The genome wide

approach is widely used to characterize the methylomes of cells/tissues and to discover changes in the methylome. As mentioned in chapter 1.1.1 on page 10, the majority of all CpGs is highly methylated across all cell types and the methylation status of certain CREs like promoters and enhancers shows the highest degree of change across cell types or cell states. This observation was made using Whole Genome Bisulfite Sequencing (WGBS), the most comprehensive method to study DNA methylation by covering nearly all CpGs (Lister et al. 2009; Ziller et al. 2013). However, it is not necessary to perform expensive WGBS experiments, if the interest of a study lies in e.g. gene regulation difference. As a cost-efficient alternative Meissner et al. 2005 developed a method that enriches for CpG rich regions and therefore covers the majority of CPIs, promoters, and putative enhancers. For Reduced Representation Bisulfite Sequencing (RRBS) endonucleases are used to fragment the DNA. The enzymes are selected based on the recognition site that enriches for DNA fragments with present CpGs. The commonly used enzyme *Msp*I (C^CGG) generates methylomes that represent 20% of all CpGs with a strong enrichments of methylation information at promoters. For this work, the enzyme *Hae*III (GG^CC) was selected, as it yields a broader coverage across the genome (4-5 million CpGs in mouse, 6-8 million CpGs in human) and provides a good approximation to WGBS. The workflow to generate sequencing libraries is represented in **Figure 1-7** and described in detail in chapter 2.2.5 on page 47. Briefly, in the first step the DNA is digested with the endonuclease *Hae*III and subsequently ligated to sequencing adapters. In the next step the DNA fragments are treated with sodium bisulfite, which converts unmethylated cytosines into uracil while 5-mCs are not converted. After amplification of the bisulfite treated single stranded DNA and subsequent sequencing, it is possible to distinguish between methylated and unmethylated CpG positions based on the comparison to the reference sequence. In case the reference cytosine is read as thymine, the position was unmethylated, conversely, a cytosine readout indicates a methylated CpG position. The detailed processing of the data is described in chapter 2.2.10 on page 53.

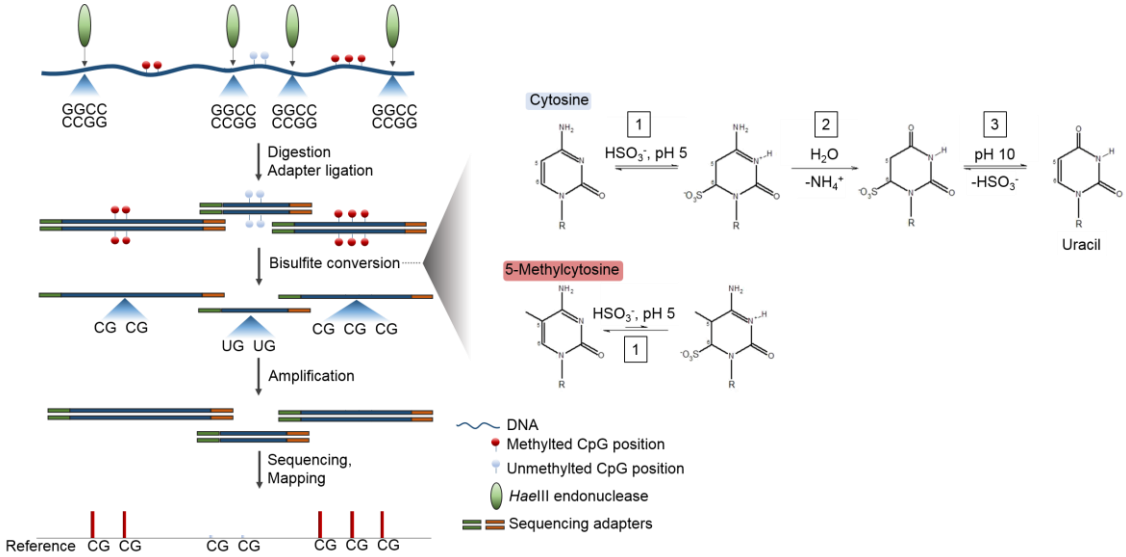


Figure 1-7 RRBS method overview. *DNA is digested by the* endonuclease *HaeIII, the fragments are adenylated to subsequently ligate sequencing adapters. The fragments are treated with sodium bisulfite which converts unmethylated cytosine into uracil. The conversion is a three-step process. First, the reactive agent hydrogen sulfite HSO3- binds to the C5 position of cytosine by nucleophile addition (1). Nucleophilicity of C5 position of 5-mC is reduced due to the positive inductive effect of the methyl group. Therefore, the addition of HSO3- to 5-mC is much slower compared to cytosine, 5-fC or 5-caC. By subsequent deamination (2) and desulfonation (3) cytosine is converted into uracil. The bisulfite treated DNA is amplified by PCR and sequenced. After PCR the converted uracils are read as Ts (thymines) and 5-mC positions are read as Cs (cytosines) in the sequencing procedure. The mapping to a reference genome reveals the 5-mC positions.*

RNA Sequencing

Various technologies have been developed to approach and quantify the transcriptome, including sequencing based approaches. The principle for the various available protocols is based on a reverse transcription of the RNA into copy DNA (cDNA). The different RNA species are separated based on size exclusion or hybridization of the poly adenylated (polyA) mRNA. The subsequent library preparation differs between the protocols. The workflow used for this thesis is represented in **Figure 1-8** and described in detail in chapter 2.2.6 on page 49. Briefly, the mRNA is captured by oligo(dT) oligonucleotides and reverse transcribed into cDNA. The double stranded cDNA is fragmented using the hyperactive *Tn5* transposase which directly introduces sequencing adapters to the DNA fragments. The library is amplified

by PCR and sequenced using NGS. The sequencing reads are aligned to reference genome and the reads mapped to the exons of a genes are used to quantify the expression of the respective gene. The data processing is described in more detail in chapter 2.2.11 on page 53.

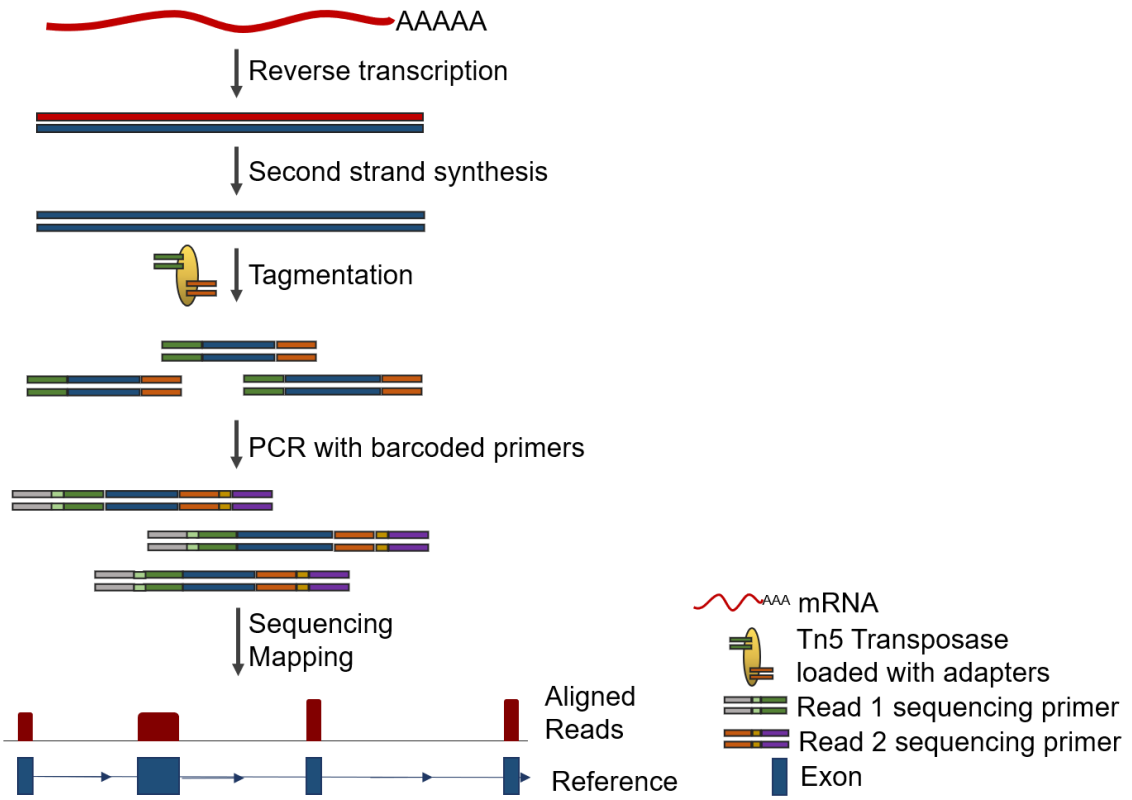


***Figure 1-8 mRNA-Sequencing workflow**. polyA selected mRNA is reverse transcribed into double stranded cDNA which is subsequently fragmented and linked to sequencing adapters by the hyperactive transposase Tn5. After sequencing the reads are aligned to the reference and the reads aligned to the exons of a gene/transcript are used for expression quantification.*

1.2 Biological background on astrocytes

The CNS consists of four major cell types, namely neurons, astrocytes, oligodendrocytes and microglia. Astrocytes are the most abundant glial in the adult human and mouse brain and are implicated in various functions such as metabolic, homeostatic, blood flow modulating, as well as structural functions. These ubiquitous cells extrude thousands of processes that ensheath synapses, nodes of Ranvier, blood vessels and form contact via gap junctions with astrocytes in the close neighborhood. They take up nutrients through the blood vessel and deliver them to neurons. For example, the up taken glucose is stored in the form of glycogen and metabolized to lactate to provide an essential energy source to adjacent neurons (Brown and Ransom 2007). Furthermore, astrocytes are the main source of essential structural components of the plasma membrane like cholesterol and lipoproteins for neurons since the liver produced cholesterol cannot pass the blood-brain-barrier. Through the contact of astrocytic processes with synapses, astrocytes are involved in synapse formation, removal, and function. For instance, astrocytes secrete Hevin that promotes the formation of excitatory synapses by connecting pre- and post-synaptic adhesion proteins. On the other hand, astrocytes can secrete Sparc that antagonizes Hevin and therefore can suppress synapse formation (Kucukdereli et al. 2011). Through membrane-bound receptors that recognize engulfment signals and phagocytosis activity, astrocytes contribute actively to synapse elimination (Chung et al. 2013, 2016). Astrocytes respond to excitatory and inhibitory neurotransmission by intracellular and intercellular Ca^{2+} elevations and in turn modulate neuronal activity. The term tripartite synapse (**Figure 1-9**) refers to this astrocytic modulatory activity in synaptic transmission and includes sensing of synaptic activity through membrane-bound receptors (e.g. AMPA, NMDA and $GABA_{a/b}$ receptors), taking up synaptic neurotransmitters via transporters (e.g. GLAST, GLT-1 and GAT-3), and releasing gliotransmitters (e.g. glutamate, D-serine and ATP) (Perea, Navarrete, and Araque 2009; Haydon and Parpura 2009). In addition, astrocytes modulate excitability of neurons by rapidly decrease the extracellular K^+ concentration enabling fast repetitive neurotransmission (Walz 2000). Consequently, the awareness of such a coordinated

network comprising both neurons and glia changed the historical view on brain function to be a result exclusively conducted by neurons.

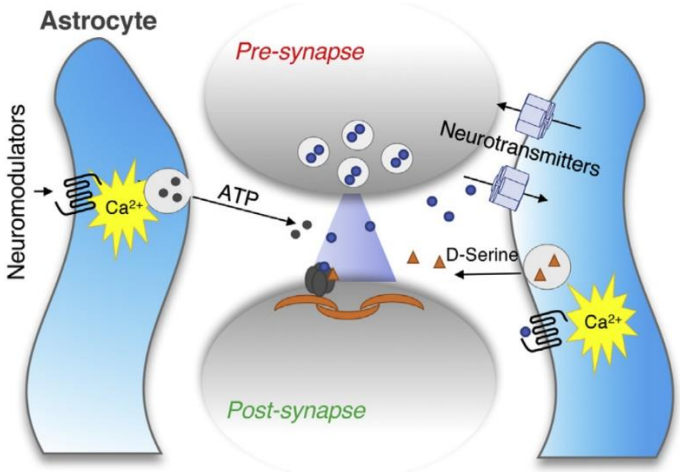


Figure 1-9 Tripartite synapse. *Astrocytes together with pre- and post-synapse form the tripartite synapse. Astrocytes are involved in the synaptic transmission by taking up and releasing neurotransmitters. The astrocytic response to neuronal signaling is transmitted through intracellular Ca^{2+} elevations. Allen and Eroglu 2017*

1.2.1 Morphological astrocyte diversity

The functional complexity described above is accompanied by morphological and regional diversity of astrocytes in the CNS. The earliest morphological heterogeneity was mentioned in 1893 by William Lloyd Andriezen who described differences between protoplasmic astrocytes of the grey matter and fibrous astrocytes of the white matter (Andriezen 1893). The morphological heterogeneity was further refined by Ramon y Cajal who stained mammalian cortices using a gold chloride-sublimate method that stains GFAP intermediate filaments (García-Marín, García-López, and Freire 2007). However, the lack of approaches and technologies to explore the functional relevance of morphologically distinct astrocytes led to the view that astrocytes are a functionally homogenous cell population. Only in the past decades, with the advances in cell culturing, cell labeling and molecular approaches, the functional and molecular diversity of astrocytes began to be appreciated.

Today, nine different astrocytic morphologies across the mammalian CNS are reported along with an array of functional features reflecting their neuronal

environment. *Protoplasmic astrocytes* are found in the grey matter of the CNS, possess highly branched "bushy" processes that form perivascular endfeet, enwrapping blood vessels, and form contacts with neurons. One single protoplasmic astrocyte is in contact with about 100,000 synapses in mice and exhibits the above described functions of a tripartite synapse (Bushong et al. 2002). *Fibrous astrocytes* are present in the white matter of the CNS, have an elongated shape that is in line with myelinated fibers of the white matter, and form less elaborate processes that are in contact with blood vessels and nodes of Ranvier. This astrocyte type is characterized by a high expression of the intermediate filament protein GFAP in the entire cell body, while protoplasmic astrocytes GFAP is often found in the perivascular endfeets (Oberheim, Goldman, and Nedergaard 2012). The function of fibrous astrocytes remains to be explored in detail. *Bergmann glia* is a specialized, unipolar astrocyte type of the cerebellum found in the Purkinje cell layer. Eight Bergmann glia cells surround one Purkinje neuron and send out long and branched processes through the molecular layer of the cerebellum to the pial surface. During development, the radial processes serve as scaffold for migrating granule neurons of the cerebellum. In the mature brain, the processes contact numerous synapses and exhibit synapse modulating functions of the above described tripartite synapse (Yamada and Watanabe 2002). *Radial glia* is a bipolar shaped neural precursor cell type that exhibits many features of astrocytes such as glycogen granules, expression of intermediate filaments GFAP and vimentin, and expression of glutamate transporter Glast. Radial glia has two processes with one at the ventricular wall and the other extruding to the pial surface. As precursor cell they give rise to neurons and glial cells during development and constitute a scaffold for neuronal migration. Besides the developing brain, Radial glia is found in the subventricular zone (Malatesta, Appolloni, and Calzolari 2008). Furthermore, specialized astrocyte types were described in the retina (*Müller glia*), in the granular layer of the cerebellum (*velate glia*), at the pial surface (*perivascular glia*), in the third ventricle (*tanycytes*), and as an epithelial layer that lines the brain ventricles (*ependymal glia*).

Functionally highly specialized astrocytes are found in the chemoreceptive regions of the ventral surface of the medulla oblongata. They sense changes in the blood pH and release ATP to adjust the respiratory rate (Gourine et al. 2010). The

hypothalamus contains hormone responsive astrocytes that are sensitive to hunger and thirst hormones and modulate neuronal activity to regulate glucose uptake (García-Cáceres et al. 2016).

1.2.2 Molecular astrocyte diversity

The advent of transgenic mouse lines to target astrocytes in combination with gene expression analysis extended the understanding of the molecular differences between regional astrocytes. In 2017, three independent studies using different approaches to isolate astrocytes investigated the transcriptomes of astrocytes from distinct brain regions (Morel et al. 2017; Chai et al. 2017; John Lin et al. 2017). Morel et al. 2017 demonstrated that astrocytic gene expression follows the dorsoventral axis with gradual changes in expression and with only a low number of genes to be exclusively expressed in only one region. Moreover, by co-culturing astrocytes and neurons from cortical and subcortical regions in matched or mismatched cultures, they demonstrated a selectively region-matched modulation of neurite growth and synaptic activity. A comprehensive, integrated study on striatal and hippocampal astrocytes comparing their physiology, morphologies, proteomes, and transcriptomes revealed properties such as morphology, electrophysiology, and Ca^{2+}-signaling to be specific within the respective neural circuit (Chai et al. 2017). Very recently, a transcriptome study of astrocytes on single-cell level revealed seven molecularly distinct sub types with a clear spatially specialized distribution that developmental patterning of the neural tube (Zeisel et al. 2018).

This interregional diversity is accompanied by intraregional heterogeneity demonstrated in mouse brain cortex and cerebellum. Farmer et al. 2016 compared the gene expression profiles of cerebellar astrocyte population present in the Purkinje cell layer (Bergmann glia) or granular cell layer (velate glia). In physiological conditions, these cells have distinct expression profiles driven by the presence or absence of the Purkinje cell-derived sonic hedgehog (SHH) protein. By artificial increase of SHH in the granular cell layer, the gene expression profile of velate astrocytes becomes more similar to Bergmann glia profile (W. T. Farmer et al. 2016). The cerebral cortex is

organized in six layers and contains layer-specific astrocytes. By the disruption of the layer-generation, astrocytes lose layer-specific properties as a result of missing environmental signals (Lanjakornsiripan et al. 2018). These observations indicate that the formation of intraregional subpopulations is highly dependent on the environmental cues.

1.2.3 Developmental patterning of astrocytes

The above described morphological and molecular differences of reginal astrocytes were long thought to be mainly determined by the signals derived from distinct neuronal environments. However, a growing body of evidence indicates that the molecular properties of astrocytes strongly relate to their derivation from regionally specialized (patterned) radial glia (see chapter 1.2.2 on page 26). During embryonic brain development, neuroepithelium cells are organized into domains that produce domain-restricted neuronal subtypes (see also chapter 1.1.2 on page 15). These domains are established by extrinsic morphogens such as bone morphogenetic protein (BMP) for dorsal orientation and SHH for ventral orientation. In turn, dorsoventral segmentation of the progenitor cells is reflected in the expression of homeodomain and basic helix–loop–helix (bHLH) transcription factors that refine domain boundaries (Campbell 2003; Sur and Rubenstein 2005). This interplay of transcriptions factors leads to the differentiation of radial glia into specified neuronal subtypes. After the neurogenic phase, radial glia cells start to differentiate into astrocytes. Fate mapping experiments of astrocytes in the spinal cord and the brain demonstrated that astrocytes derive from spatially pre-specified progenitors, show little to no lateral migration and are unable to populate adjacent domains (Tsai et al. 2012; Hochstim et al. 2008; Magavi et al. 2012). The usage of the same transcription factor code in neurons and astrocytes might ensure regionally matched interactions and neural circuit formation.

1.3 Biological background on B-cells

The mammalian immune system is a multilayered complex system comprising innate and adaptive immune responses to defend the organism from foreign substances (antigens). Innate immune response represents the immediate protection against antigens and is mediated on the cellular level by macrophages, neutrophils, eosinophils, basophils, natural killer cells, mast cells, and dendritic cells. Adaptive immunity is a highly sophisticated biological response mediated by B- and T-cells after exposure to a specific antigen that evaded innate immune system (Moser and Leo 2010). B-cells mediate immune response by producing antigen-specific antibodies and secreting pro- and anti-inflammatory cytokines and operate as antigen presenting cells. There are three major classes of mature B-cells that are distinguished by their ontogeny and anatomic localization: follicular (FO) B-cells, marginal zone (MZ) B-cells and B1 B-cells (**Figure 1-10**). FO and MZ B-cells develop from transitional 2 (T2) B-cells that originate from hematopoietic precursor cells from the bone marrow, while B1 B-cells arise from progenitors in the fetal liver and are maintained by self-renewing during adulthood. FO B-cells reside in the lymphoid follicles of the spleen and lymph nodes, whereas MZ B-cells are located around the splenic marginal sinus where they are in close association with blood-borne pathogens. B1 B-cells reside mostly in the peritoneal and pleural cavities and produce polyreactive or polyspecific antibodies (IgM) to bind both self-antigens and microbial antigens (Hoffman, Lakkis, and Chalasani 2016).

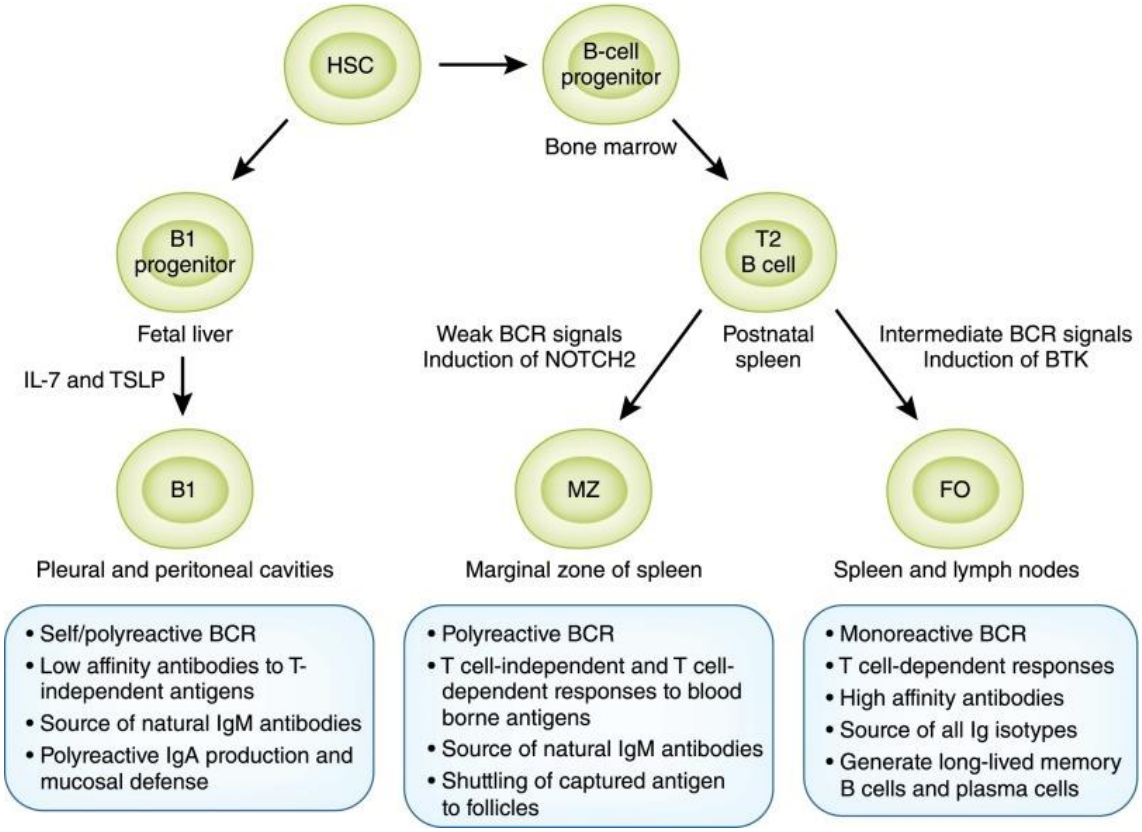


Figure 1-10 B-cell subtypes. *Hematopoetic stem cells (HSC) are the common precursor for all B-cells. They give rise to progenitor cells of B1 B-cells in the fetal liver and progenitor cells of transitional 2 (T2) B-cells in the bone marrow. T2 B-cells further differentiate into marginal zone (MZ) and follicular (FO) B-cells in the spleen. This lineage commitment is determined by B-cell receptor (BCR) signals. The three B-cell classes exhibit distinct properties in antigen-recognition, as well as distinct antibody classes production (Hoffman, Lakkis, and Chalasani 2016).*

In order to produce antibodies, B-cells have to differentiate into plasmablasts (PBs) and further into plasma cells (PCs). The differentiation is initiated by the activation of the B-cell through antigen-receptor interaction. The differentiation takes place either rapidly outside of the follicle to generate a fast but less specific response, or after the entry of the activated cell into the follicle where the cell undergoes extensive proliferation, affinity maturation of immunoglobulin antigen–binding sites and immunoglobulin class-switching to generate highly specific antibodies (Nutt et al. 2015).

Besides antibody production, B-cells mediate immune responses by the secretion of pro-inflammatory cytokines such as interleukin (Il) 1, Il2, Il4, Il6, Il12, IFNγ, TNFα, as well as anti-inflammatory cytokines Il10 and Il35 (Fillatreau 2013). Consequently, B-cells play multifaceted roles in infection, autoimmune reactions, and cancer.

1.4 Aim of this thesis

This work focuses on the analysis of epigenomic profiles of various cell types such as astroglia subpopulations from the adult murine brain, as well as a multitude of B-cell subpopulations and integrates the respective gene expression profiles. The epigenomic and transcriptomic patterns shed light on the characteristics of the respective cell populations and thereby contribute to our understanding of the factors responsible for the establishment of cell identity. Moreover, DNA methylation signatures can be used to infer relations between the epigenomes of different cell populations to explore the origin of respective cell populations.

Chapter 3.1. presents the comprehensive epigenomic and transcriptomic analyses on astrocyte diversity. Astrocytes were, for a long time, considered a homogenous, glue-like cell population and despite important progress, they remain understudied. This thesis seeks to determine whether astroglia populations from two distinct, mature brain regions display diversity in their molecular machinery and if so, how the differences were established. To this end, two exemplary brain regions were chosen: the cerebral cortex (CTX) and the cerebellum (CB). Both regions differ in cellular composition, connectivity, and are involved in distinct central nervous functions. Putative region-specific programs were expected to be found by isolating astrocytes from CTX and CB, and exploring their DNA methylation, open chromatin, and transcriptome in an integrative manner.

Chapter 3.2. dissects the molecular characteristics of *in vivo* regulatory B-cells (Bregs), employing gene expression and DNA methylation analyses. Bregs support immunological tolerance by immunosuppressive activity. The suppression of immunopathology is exerted by the production of Interleukin-10, which in turn restricts expansion of other lymphocytes, and the production of pro-inflammatory cytokines. The identification of Lag3-expressing plasma cells as Bregs led to two questions: (I) What are the molecular differences between $Lag3^{+}$ plasma cells and $Lag3^{-}$ plasma cells? (II) And which B-cell subtype gives rise to these regulatory plasma cells?

Besides the investigation of the above described biological questions, this thesis

extends the background of regulatory concepts and epigenetic mechanisms involved in the establishment of cell identity. It provides a general overview of the complexity of epigenetic gene regulation.

Chapter 2.

Materials and Methods

2.1 Materials

Table 1 Antibodies used in this work

Antibodies	*Source*	*Catalogue number*
Anti-Goat Alexa Fluor®555 conjugated	Invitrogen	A21432
Anti-Lhx2	Merck	ABE1402
Anti-NeuN Antibody, clone A60, Alexa Fluor®488 conjugated	MerckMillipore	MAB377X
Anti-Rabbit Alexa Fluor®546 conjugated	Invitrogen	A10040
Anti-Zic1	RDsystems	AF4978

Table 2 Chemicals and enzymes used in this work

Chemicals, Enzymes	*Source*	*Catalogue number*
Adenosine 5′-Triphosphate (ATP)	New England Biolabs	P0756L
Bovine serum albumin	Sigma Aldrich	A9647-100G
Calcium chloride ($CaCl_2$)	Roth	5239.1
Chloroform	VWR	22920.K2
cOmplete, EDTA-free Protease Inhibitor Cocktail	Sigma Aldrich	11873580001
Deoxyribonucleotide triphosphate (dNTPs)	Solis Biodyne	02-21-00400
4′,6-Diamidino-2-phenylindole (DAPI)	Roche	10236276001
Dithiothreitol (DTT)	Invitrogen	Y00147
DNA Polymerase I	New England Biolabs	M0209L
DNase I	Sigma Aldrich	D4263
Ethanol	Fisher chemicals	E/0650DF/17
Ethylene glycol-bis(2-aminoethylether)-N,N,N′,N′-tetraacetic acid (EGTA)	Fluka	3778
Ethylenediaminetetraacetic acid (EDTA)	Sigma Aldrich	E9884
Glycogen	Thermo Fischer	10814010
HaeIII	New England Biolabs	R0108M
Hanks' Balanced Salt solution (HBSS)	Sigma Aldrich	RNBG4539
HBSS 10x	Thermo Fischer	14185052
Hoechst 33342	Sigma Aldrich	B2261-25MG
HotStar Taq	Qiagen	203207
ImmunMount	Thermo Fischer	9990402
Klenow Fragment exo-	New England Biolabs	M0212L

Chemicals, Enzymes	*Source*	*Catalogue number*
Magnesium acetate ($Mg(C_2H_3O_2)_2$)	Merck	5819
M-MLV Reverse Transcriptase RNase H-	Promega	M3683
Normal goat serum	Jackson Immuno Research	005-000-121
10% NP-40	Abcam	ab142227
Papain	Sigma Aldrich	P3125
Pefabloc SC	Sigma-Aldrich	76307
Percoll	Sigma Aldrich	P1644-25ML
Phenol-chloroform-isoamyl (25:24:1)	Sigma Aldrich	P2069-100ML
Phenylmethanesulfonylfluoride (PMSF)	Sigma Aldrich	10837091001
Potassium chloride (KCl)	Merck	1.04936.0500
Potassium dihydrogenphosphate (KH_2PO_4)	Merck	P8416
Propidium iodide	Miltenyi Biotec	130-093-233
Proteinase K	Sigma Aldrich	
RNase H	New England Biolabs	M0297L
Saccharose	Roth	4621.1
Sodium chloride (NaCl)	ZChL	701313
Sodium dihydrogenphosphate (NaH_2PO_4)	Merck	S9638
T4-ligase	New England Biolabs	M0202M
T4-ligase	New England Biolabs	M0202L
Tris(hydroxymethyl)aminomethane hydrochloride (TrisHCl)	Roth	90.90.3

Table 3 Devices used in this work

Devices	*Source*
Agilent 2100 Bioanalyzer	Agilent
AxioScanZ.1	Zeiss
BD FACSaria III	Becton Dickinson
HiSeq 2500	Illumina
LSM 710	Zeiss
Nanodrop2000c	Thermo Fischer
PCR thermocycler	Eppendorf/ Applied, Biosystems
Qbit® Fluorometer	Invitrogen

Table 4 Commercially available Kits used in this work

Kits	*Source*	*Catalogue number*
Agencourt AMPure XP	Beckman Coulter	A63881
Agilent High Sensitivity DNA Kit	Agilent	5067-4626
Agilent RNA 6000 Pico	Agilent	5067-1514
CellTrics strainer (20 or 150 µm)	Symex Partec	04-0042-2315, 04-0042-2319
Direct-zol RNA Miniprep Plus Kits	Zymo research	R2072
EASYstrainer™ for 50 ml tubes, 70 µm mesh size	Greiner	542 070
EZ DNA Methylation-Gold Kit	Zymo research	D5005
MinElute PCR Purification Kit	Qiagen	28006
mRNA Capture Kit	Sigma-Aldrich	11787896001
NEBNext High-Fidelity 2X PCR Master Mix	New England Biolabs	M0541S
Nextera DNA Library Preparation Kit	Illumina	FC-121-1031
Nextera Index Kit	Illumina	FC-121-1011

Kits	*Source*	*Catalogue number*
Qubit® dsDNA HS Assay Kit	Thermo Fisher Scientific	Q32851
TruSeq DNA Single Indexes Set A	Illumina	20015960

Table 5 Mouse models used in this work

Mouse models/strains	*Source*	*Catalogue number*
Mouse: TgN(hGFAP-EGFP)GFEA (GFAP-EGFP)	AG Kirchhoff	N/A
Mouse: TgN(PLP-DsRed1)PRDB (PLP-DsRed)	AG Kirchhoff	N/A
Mouse: (C57BL6/J x DBA2/N)F1	AG Walter	N/A

Table 6 Oligonucleotides used in this work

Oligonucleotides
TruSeq PCR primer 1 5'AAT GAT ACG GCG ACC GAG ATC TAC AC
TruSeq PCR primer 2 5'CAA GCA GAA GAC GGC ATA CGA GAT

Table 7 Software and R packages used in this work

Software	*Source*	*Web page*
2-step STAR alignment	(Dobin and Gingeras 2015)	
Alibaba2		http://gene-regulation.com/pub/programs/alibaba2/
BWA	(Heng Li and Durbin 2010)	http://bio-bwa.sourceforge.net/
ChipSeek	(Chen et al. 2014)	http://chipseek.cgu.edu.tw/index_show.py
ChromH3M	(Salhab et al. 2018)	https://github.com/asalhab/ChromH3M/releases
circlize	(Gu et al. 2014)	https://cran.r-project.org/web/packages/circlize/index.html
csaw	(Lun and Smyth 2016)	https://bioconductor.org/packages/release/bioc/html/csaw.html
Cutadapt	(Martin 2011)	http://dx.doi.org/10.14806/ej.17.1.200
Cytoscape	(Shannon et al. 2003)	https://cytoscape.org/
DAVID	(Huang, Sherman, and Lempicki 2009a, 2009b)	https://david.ncifcrf.gov/
deepTools	(Ramírez et al. 2014)	https://deeptools.readthedocs.io/en/develop/
edgeR	(M. D. Robinson, McCarthy, and Smyth 2010)	https://bioconductor.org/packages/release/bioc/html/edgeR.html
featureCounts	(Liao, Smyth, and Shi 2014)	http://subread.sourceforge.net/
Galaxy	(Afgan et al. 2018)	https://usegalaxy.org/
Galaxy Operations		https://github.com/galaxyproject/gops

Software	*Source*	*Web page*
GEM mapper	(Marco-Sola et al. 2012)	https://github.com/smarco/gem3-mapper
genevenn		http://genevenn.sourceforge.net/
Ggplot2	(Wickham 2016)	http://ggplot2.org
graphics		https://stat.ethz.ch/R-manual/R-devel/library/graphics/html/00Index.html
GREAT	(McLean et al. 2010)	http://bejerano.stanford.edu/great/public/html/
IGV	(J. T. Robinson et al. 2011; Thorvaldsdottir, Robinson, and Mesirov 2013)	http://software.broadinstitute.org/software/igv/
Jaspar	(Khan et al. 2018)	http://jaspar.genereg.net/
macs2	(Yong Zhang et al. 2008, 2)	https://pypi.org/project/MACS2/
MethylCtools	(Hovestadt et al. 2014)	https://github.com/hovestadt/methylCtools
methylKit	(Akalin et al. 2012)	https://bioconductor.org/packages/release/bioc/html/methylKit.html
MethylSeekR	(Burger et al. 2013)	https://bioconductor.org/packages/release/bioc/html/MethylSeekR.html
Panther	(Mi et al. 2017, 2013)	http://www.pantherdb.org/
pheatmap		https://CRAN.R-project.org/package=pheatmap
Picard tools		http://broadinstitute.github.io/picard
R	(R Development Core Team 2008)	https://www.r-project.org/

Software	*Source*	*Web page*
RColorBrewer		https://cran.r-project.org/web/packages/RColorBrewer/index.html
REVIGO	(Supek et al. 2011)	http://revigo.irb.hr/
RNA-seQC	(DeLuca et al. 2012)	
RSAT	(Nguyen et al. 2018)	http://rsat.sb-roscoff.fr/
samtools (version 1.3)	(H. Li et al. 2009)	https://sourceforge.net/projects/samtools/files/samtools/1.3/
stats		https://www.rdocumentation.org/packages/stats
STRING db	(Szklarczyk et al. 2017)	https://string-db.org/
Trap	(Thomas-Chollier et al. 2011)	http://trap.molgen.mpg.de/
Trim Galore		http://www.bioinformatics.babraham.ac.uk/projects/trim_galore
Zen	Zeiss	

Table 8 Solutions and buffers used in this work

Solutions and buffers	*Components*
0.5 M PIPES stock buffer[5]	0.5 M PIPES in ddH_2O
1X TE Buffer	10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0
10x GKN Buffer[5]	80 g/l NaCl, 4 g/l KCl, 35.6 g/l Na_2HPO_4 12 H_2O, 7.8 g/l NaH_2PO_4 2 H_2O, 20 g/l D-(+)-glucose, pH 7.4
10X Phosphate buffered saline (PBS)	80 mM Na2HPO4, 20 mM KH2PO4, 1.5 M NaCl, 30 mM KCl, pH 7.4
25 % Percoll solution	2.5 ml SIP, 7.5 ml 1X PBS pH 7.4
ATAC Reaction Buffer	25 µl TD Buffer 2X, 2.5 µl TD Enzyme, 22.5 µl MilliQ
Blocking Solution	3% bovine serum albumin, 2 % normal goat serum in 1X PBS; pH 7.4
DNA extraction Solution A	75 mM NaCl, 25 mM EDTA
DNA extraction Solution B	10 mM EDTA, 10 mM Tris-HCl (pH 8.0), 1 % SDS
Homogenization buffer	10 mM Tris-HCl (pH 8), 0.1 mM EDTA, 3 mM $Mg(Ac)_2$, 5 mM $CaCl_2$, 0.1 % NP-40, 0.32 M sucrose, 1X cOmplete EDTA-free Protease Inhibitor Cocktail (Sigma Aldrich), 0.16 mM DTT, 10 mM PMSF
Nuclei extraction buffer	60 mM KCl, 15 mM Tris-HCl, 15 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 0.5 mM Spermidine (free base), 1X cOmplete Protease Inhibitor Cocktail, 0.1 % NP-40
Papain activation buffer[5]	55 mM L-Cystein HCl, 11 mM EDTA, pH 7.4

[5] Provided by AG Kirchhoff

Solutions and buffers	*Components*
PIPES working solution[5]	0.5 M PIPES stock buffer (PIPES 0.5 M (Sigma Aldrich), 1 M NaOH, 1.2 M NaCl and 50 mM KCl solution, 45 % glucose in 50 ml ddH_2O
RRBS lysis buffer	10 mM TrisHCl, 5 mM EDTA
Stock Isotonic Percoll (SIP)	1 ml 10X PBS, 9 ml Percoll, adjust pH to 7.4

[5] Provided by AG Kirchhoff

2.2 Methods

2.2.1 Preparation of single cell suspension from fresh mouse brain

Mice were anesthetized with ketamine/xylazine (140 mg/kg and 10 mg/kg body weight) and perfused with ice cold HBSS by Dr.Carmen Kasakow. After perfusion the brain was directly removed, and cerebral cortex and cerebellum were dissected and minced into small tissue pieces and transferred into 10 ml ice cold HBSS. The tissue was spun down at 300 xg, 4 °C for 1 minutes and the supernatant was removed down to 5 ml. Next, 5 ml PIPES working solution were added to the cells gently mixed by inverting. The tissue was centrifuged at 300 xg, 4 °C for 5 minutes. In the meantime, the papain solution was prepared by adding 156 µl papain (16 U/ml) into 4.5 ml PIPES working solution supplemented with 500 µl of activation buffer. The papain was activated for 20 minutes at 37 °C. After activation, 100 µl DNaseI (80 Kunitz units/mL) were added to the papain solution and the mixture was drawn into a 10 ml syringe. After pelleting the tissue, the supernatant was removed down to 5 ml and the papain/DNaseI was added to the suspension by pushing the solution through an attached 40 µm syringe filter. The suspension was mixed by gentle inversion, placed into the MACSmix™ Tube Rotator and incubated at 37 °C for 50 minutes with low speed rotation. During the incubation DNaseI solution II was prepared by adding DNaseI (final concentration 25 Kunitz units/mL) to 6 ml of wash buffer (0.5 % BSA in DMEM). After the 50 minutes digestion the tissue was gently triturated by using p1000 with cut and flamed tips. After the first trituration step the DNaseI solution II was added to the suspension followed by the second digestion step at 37 °C, 15 minutes with low speed rotation. The papain/DNaseI solution was removed by centrifugation at 300 xg for 10 minutes at 4 °C and discarding the supernatant. 2 ml wash buffer were added to the cell pellet and the cells were triturated gently using a p1000 to a homogeneous single cell suspension. The cells were filtered through a pre-wetted 70 µm strainer and centrifuged at 300 xg for 10 minutes at 4 °C. The supernatant was discarded, and the cells were washed with 40 ml wash buffer (0.5 % BSA in DMEM) to quench the digestion reaction. Subsequently, the cells were centrifuged at 300 xg for 20 minutes at 4 °C, the supernatant was discarded, and the cells were resuspended in 1 ml 1X GKN buffer.

2.2.2 Nuclei preparation for NeuN Sorting

Neuronal and non-neuronal nuclei of F1(C57BL6/J x DBA2/N) mice were extracted from frozen forebrain tissue. The tissue was transferred into a douncer filled with 5 ml pre-cooled homogenization buffer and dissociated on ice with 50 strokes in 1 minute. To remove undissociated tissue and cell debris, the suspension was filtered sequentially through 150 µm and 20 µm CellTrics strainer. The nuclei were centrifuged at 500xg, 4 °C for 7 minutes and the supernatant was discarded. The nuclei pellet was resuspended in 25 % Percoll solution and centrifuged for 15 minutes at 4° C with 750 xg and low brake to remove the remaining cell debris and myelin. The debris formed a layer on top of the suspension which was discarded. Nuclei pellet was kept on ice for 10 minutes to facilitate a gentle resuspension of nuclei in 10 ml PBS supplemented with Protease Inhibitor Cocktail. After the PBS wash, nuclei were resuspended and incubated in 500 µl blocking solution for 30 minutes at 4 °C in the dark. The anti-NeuN-Alexa488 antibody was diluted 1:1000 in 500 µl blocking solution and blocked for 30 minutes at 4 °C in the dark. After blocking, the antibody solution was added to the nuclei and the immune reaction was carried out for 1 h at 4 °C in the dark. The antibody solution was removed, and the nuclei were resuspended in 500 µl blocking solution containing 1 µg/ml propidium iodide (PI).

2.2.3 Fluorescence activated cell sorting (FACS)

Before setting up the sorting strategy, following conditions had to be adjusted on the BD FACSaria III. The collection device was cooled to 5 °C. The sample reservoir was cooled to 4 °C and agitation of the sample reservoir had to be switched on. A 70 µm nozzle and a 2.0 grey filter had to be installed. The stream flow speed was set to 3 and the drop delay was determined with AccuDrops for 95-100% events in the side stream to achieve high accuracy during the sort. After adjustment of the machine settings, cell-/ nuclei-suspensions were filtered through a prewetted 50 µm strainer followed by DNA staining with Hoechst 33342 1:1000 (for cells) or PI 1:100 (for nuclei) before installing the tubes into the BD FACSaria III. Cells from the TgN(hGFAP-EGFP)$_{GFEA}$ (GFAP-EGFP) mice were sorted for a high EGFP signal measured with the FITC 530/30 filter, a high

Hoechst 33342 signal measured with the Hoechst blue 450/20 filter, and in addition for the size according to the forward scatter to exclude debris and cell nuclei. Cells from the TgN(PLP-DsRed1)$_{PRDB}$ (PLP-DsRed1) mice were sorted for a high DsRed signal measured with the mCherry 610/20 filter, a high Hoechst 33342 signal measured with the Hoechst blue 450/20 filter, and in addition for the size according to the forward scatter to exclude debris and cell nuclei. The cells were sorted into 5 ml BD tubes containing 250 µl 1X GKN. Cells were transferred into 1.5 ml tubes after sorting and centrifuged at 300 xg, 4 °C for 10 minutes. The supernatant was discarded, and the cell pellets were frozen on dry ice and stored at -80 °C for RRBS and mRNA-Seq. For ATAC-Seq the cells were directly used for library preparation. Cell nuclei from the NeuN labeling were sorted for a high or low Alexa Fluor® 488 signal measured with the FITC 530/30 filter, and a high PI signal measured with the PE 585/15 filter. Duplets and clumps displayed gradually higher PI signals and were excluded accordingly. The nuclei were sorted into 5 ml BD tubes containing 250 µl 1X HBSS. Nuclei were transferred into 1.5 ml tubes after sorting and centrifuged at 500 xg, 4 °C for 10 minutes. The supernatant was discarded, and the nuclei pellets were frozen on dry ice and stored at -80 °C for DNA preparation. For ATAC-Seq the nuclei were directly used for library preparation. Sorting accuracy was evaluated by flow cytometry of a small aliquot of the sorted populations.

2.2.4 ATAC-Seq library preparation

After FACS, 50,000 cells were centrifuged at 300 xg, 4 °C for 10 minutes and the supernatant was discarded. The pellet was carefully resuspended in Nuclei Extraction Buffer and kept on ice for 10 minutes. To remove the extraction buffer, nuclei were pelleted at 500 xg, 4 °C, for 10 minutes, and the supernatant was discarded. For sorted NeuN$^+$ and NeuN$^-$ nuclei, the nuclei extraction steps were skipped. Nuclei were carefully resuspended in ATAC Reaction Buffer using a p200 with cut and flamed tips. The transposase reaction was incubated at 37 °C for 30 minutes. The reaction was stopped with the addition 250 µl of the PB Buffer from the Qiagen MinElute PCR Purification Kit. This buffer contains guanidine hydrochloride which inactivates the transposase and releases it from the DNA, resulting in tagmented and adapter-linked DNA fragments.

The DNA was purified with the Qiagen MinElute PCR Purification Kit following the manufacturer's manual and eluted in 10 µl nuclease free water. The library was amplified by PCR for 9-10 cycles using 12.5 µl of the NEBNext High-Fidelity 2X PCR Master mix and 0.5 µl of 10 µM Index-adapters (5'AATGATACGGCGACCACCGAGATCTACAC[i5] TCGTCGGCAGCGTC3' and 5'CAAGCAGAAGACGGCATACGAGAT[i7]GTCTCGTGGGCTCGG3') in 25 µl total volume. The library was purified with 1x Agencourt AMPure XP beads according to manufacturers' protocol and eluted in 20 µl Qiagen elution buffer. The concentration was measured using the Qbit HS Kit and the quality was assessed by analyzing the fragment distribution using the Agilent 2100 Bioanalyzer according to manufacturers' protocol. The libraries were sequenced on the Illumina HiSeq 2500 with 2x100 bp paired-end reads.

Thermocycler program:

Temp.	Min	
72°C	5:00	
98°C	0:30	
98°C	0:10	9-10 cycles
63°C	0:30	
72°C	1:00	
72°C	5:00	
4°C	hold	

2.2.5 RRBS library preparation

The frozen cell pellets were thawed in 16 µl RRBS lysis buffer and 2 µl proteinase K (1 mg/ml) were added to digest DNA-bound proteins. The reaction was carried out at 65

°C for 4 h. After cooling down the reaction to room temperature (RT), the protease activity was inhibited by the addition of 2.8 µl Pefabloc SC (8.14 mM) and incubation for 1 h at RT. Alternatively, the procedure was done with 100-200 ng DNA directly. After inhibition of the proteinase K, the reaction was supplemented with 3 µl NEB 10X Cutsmart buffer, 3.2 µl H_2O, and 1 µl *HaeIII* endonuclease (50 U/µl) to digest the DNA over night at 37 °C. *HaeIII* was inactivated at 80 °C for 20 minutes. Adenine overhangs were added to the 3' end of the DNA fragments by the addition of 1 µl dATP (10mM) and the 1 µl DNA Polymerase Klenow fragment (5 U/µl), lacking exonuclease activity. The A-tailing was carried out at 37 °C for 30 minutes and the enzyme was inactivated at 75 °C for 20 minutes. Indexed Illumina TruSeq adaptors were ligated at 16°C overnight using 2µl 1:10 adapter dilutions, 4 µl ATP (10mM), 1 µl NEB 10X Cutsmart buffer, and 1 µl T_4-Ligase (2000U/µl). The enzyme was inactivated at 65 °C for 20 minutes and the DNA was purified using 1.5x Agencourt AMPure XP beads and eluted in 20 µl H_2O. Bisulfite conversion of the DNA was performed using the EZ DNA Methylation-Gold Kit following the manufacturers' protocol. The bisulfite-DNA was eluted in 24 µl pre-warmed H_2O (65°C). 19 µl of this elution was supplemented with 2.5 µl 10X HotStar PCR Puffer, 2.5 dNTPs (2.5 mM each), 0.25 µl forward primer (5'AATGATACGGCGACCACCGAGATCTACAC3'; 10 µM), 0.25 µl reverse primer (5'CAAGCAGAAGACGGCATACGAGAT3', 10µM), and 0.5 µl HotStarTaq. The library was amplified by 12-15 cycles of PCR and purified with 0.9x Agencourt AMPure XP beads. The library quality was assessed by analyzing the DNA fragment distribution using High Sensitivity DNA Assay on the Agilent 2100 Bioanalyzer following manufacturers' protocol. The library concentration was measured using the Qubit® dsDNA HS Assay Kit according to manufacturers' protocol. Sequencing was done with the Illumina HiSeq 2500 system with 100 bp single-end reads.

Thermocycler program:

Temp.	Min	
95°C	15:00	
95°C	0:30	
60°C	0:30	12-15 cycles
72°C	1:00	
72°C	7:00	
4°C	hold	

2.2.6 mRNA-Seq library preparation

The mRNA of sorted cells was extracted using the mRNA Capture Kit. First, the cells were lysed in 49 µl lysis buffer. The cell lysate was then heated at 70 °C for 5 minutes to disrupt secondary mRNA structures and directly put on ice to avoid formation of secondary structures. Next, the lysate was incubated with 1 µl of biotinylated oligo(dT)$_{20}$ (1:20) for 5 minutes at 37 °C to hybridize the poly-A tails and biotinylated oligo(dT)$_{20}$ and was then transferred into streptavidin coated tubes to immobilize the mRNA. The lysate was discarded, and the tube-bound mRNA was washed 3 times to remove potential DNA contamination. For reverse transcription the following first-strand solution was added to the tube-bound mRNA and incubated for 2 h at 37 °C: 28 µl nuclease free water, 10 µl MLV-Buffer 5X, 10 µl dNTPs 10 mM each, 1 µl RNasin 40 U/µl, 1 µl MLV-RT 200 U/µl. The first-strand solution was discarded, and the tube was washed once with 250 µl wash buffer. To generate double stranded cDNA, the RNA strand was removed by RNase H while the DNA Polymerase I synthesized the complementary DNA strand. The following second-strand solution was prepared, added to the tube and incubated at 16 °C for 2.5 h: 33.8 µl H_2O, 5µl T4-DNA Ligase Buffer 10X, 3.75 µl KCl 1 M, 3 µl dNTPs 10 mM, 3 µl DNA Polymerase I 10 U/µl, 0.5 µl T4-Ligase 400 U/µl (NEB), 1 µl RNase H 5 U/µl (NEB). After washing with 250 µl wash buffer, the cDNA was fragmented using 0.2 µl hyperactive Tn5 transposase from the Illumina Nextera DNA Library Preparation Kit in 50 µl reaction (24.8 µl H_2O, 25 µl Tn5 2X Buffer). The reaction was carried out at 55 °C for 5 minutes and stopped with the addition of 250 µl guanidine

hydrochloride containing Qiagen PB buffer. The adapter-linked cDNA was purified with the Qiagen MinElute PCR Purification Kit following the manufacturers' manual and eluted in 11.5 µl nuclease free water. The library was amplified by PCR for 9-12 cycles by the addition of 12.5 µl of the NEBNext High-Fidelity 2X PCR Master mix, 0.5 µl Illumina indexed Primer 1 (1:10; 5'AATGATACGGCGACCACCGAGATCTACAC[i5]TCGTCGGCAGCGTC3'), and 0.5 µl Illumina indexed Primer 2 (1:10, 5'CAAGCAGAAGACGGCATACGAGAT[i7]GTCTCGTGGGCTCGG3'). The libraries were purified with 0.8x Agencourt AMPure XP beads following manufacturers' manual, eluted in 12 µl 0.1X TE and sequenced on the Illumina HiSeq 2500 with 1x100 bp single-end reads.

Thermocycler program:

Temp.	Min	
72°C	3:00	
98°C	0:30	
98°C	0:10	9-12 cycles
63°C	0:30	
72°C	3:00	
72 °C	5:00	
4°C	hold	

2.2.7 DNA extraction

DNA was extracted from sorted nuclei and purified from proteins with the phenol/chloroform extraction method. The Volume of the nuclei suspension was adjusted to 250 µl with 0.1X TE and mixed with 125 µl Solution A and 125 µl Solution B. For protein digestion 20 µl Proteinase K (20 mg/ml) was added to the sample and the reaction was incubated at 60 °C over night with agitation. An equivalent volume of phenol/chloroform/isoamyl alcohol (25:24:1, v/v) solution was added, the mixture was then mixed by overhead rotation for 15 minutes. After 15 minutes of centrifugation at 13,000 rpm, the aqueous phase was carefully transferred into a fresh tube. Again, an

equivalent volume of phenol/chloroform/isoamyl alcohol (25:24:1, v/v) was added, mixed by overhead rotation and centrifuged for 15 minutes at 13,000 rpm. The aqueous phase, free from proteins, was again transferred into a new tube. An equivalent volume of chloroform/isoamyl alcohol (24:1) was added, mixed by overhead rotation and centrifuged for 15 minutes at 13,000 rpm. The aqueous phase was again transferred into a new tube. The DNA was precipitated by the addition of 1 µl glycogen (20 µg/µl), 20 µl NaCl (5 M), and 500 µl ice-cold absolute ethanol. The solution was mixed by gentle inversion and in case of a visible DNA precipitant, the DNA was carefully transferred into a fresh tube and washed twice with ice-cold ethanol (70 %). In case of no visible DNA precipitant, the solution was placed at -20 °C over night and centrifuged for 40 minutes at 4 °C at 13,000 rpm. The supernatant was discarded, and the pellet was washed with 70 % ethanol, spun down again and the supernatant was discarded. The pellet has been air-dried and resolved in 20-50 µl 0.1X TE.

2.2.8 RNA extraction

RNA extraction was performed using the Direct-zol RNA Miniprep Plus Kits following the manufacturers' manual. The RNA quality was assessed by analysis of the RNA Integrity Number (RIN) using the Agilent RNA 6000 Pico Kit on the Agilent 2100 Bioanalyzer.

2.2.9 ATAC-Seq data analysis

After sequencing of the samples on the Illumina HiSeq 2500, the sequencing reads were filtered for quality (read tails Q > 20) and trimmed for adapters sequences using Trim Galore! (version 0.4.2). Trimmed reads were aligned to the mouse reference genome (mm10) using GEM mapper (version 1.376 beta). The SAM file format was then converted into BAM file format using samtools (version 1.3). MarkDuplicate (version 1.115) from Picard tools was used to mark the PCR duplications and peaks were called using macs2 (version 2.1.0.20140616). The parameters to call peaks were set as follows: --nomodel, --shift -125, --extsize 250. Normalization of the coverage files was

performed using the library size using the *bamCoverage* command from deep Tools (v1.5.9.1). The processing of the data up to this point was done by Abdulrahman Salhab.

The visualization of the coverage files was done by using the normalized coverage BigWig files as input into deeptools. Heatmaps were generated using *compute matrix* command with center of region as reference point extended by 1 kb upstream and downstream. Coverage profiles were generated using the deeptools *plotProfile* command.

The differential analysis of ATAC-Seq data was done using the R package csaw. The workflow requires the conversion of reads into read counts in two user-specified fixed-width genomic intervals (window sizes). The first window size defines genomic intervals that will be treated as regions to test enrichment over the background. The background is defined as the sum of read counts in a large user-specified window. Dr. Karl Nordström generated 100 bp and 10 kb read count files by using featureCounts. These files were then imported into csaw and were used to remove composition biases as the first step. This was done using the *normOffsets* wrapper function which uses the trimmed mean of M-values (TMM) method (Mark D Robinson and Oshlack 2010) to correct for any systematic enrichment differences between samples. This allows to compare samples that have a higher background with low-background samples. To reduce the running time of the differential analysis, the windows are then filtered for a count fold change of >2 compared to the background. Based on the assumption that these represent background, the majority of the windows is then removed from the analysis. The biological variance in the data is modeled using *estimateDisp* and *glmQLFit* commands which use quasi-likelihood (QL) and negative binomial methods to estimate the dispersion. Differential enrichment was tested with the *glmQLFTest* command which uses QL F-test to compute p-values (S. P. Lund et al. 2012). Significant adjacent or overlapping windows were merged into clusters up to 3 kb width with gaps no greater than 150 bp using *mergeWindows* and combined p-values were computed using *combineTests* function. The resulting counts per million (cpm) values and fold changes were used as quantitative values to represent and analyze the overall ATAC-Seq data or differentially accessible regions (DARs).

2.2.10 RRBS data analysis

In the first step of the RRBS data processing, the sequencing reads were trimmed for adapters using the Cutadapt wrapper Trim Galore! (version 0.4.2) in RRBS mode. To align the bisulfite converted DNA sequences, the mm10 reference genome was in silico converted (C to T for top strand and G to A for bottom strand) using MethylCtools. In addition, the Cs of the first sequencing read were converted to Ts and the Gs of the second read were converted to As. The alignment was then done using BWA (bwa-0.6.2-tpx), and the reads were re-converted to the original state. Reads with alignment scores < 1 were removed. The processing of the data up to this point was done by Dr. Karl Nordström.

Differential analysis of DNA methylation was carried out using the R package methylKit. In the first step, the data is filtered with the commands *filterByCoverage* and *unite* for CpG sites that are covered at least 10x and are present in at least three replicates of a group. If a group consists of less than three replicates, the sites of all available replicates are considered. After filtering, the regions to test for differential methylation were defined as 1 kb regions with at least three CpGs using the function *tileMethylCounts*. The function *calculateDiffMeth* uses logistic regression to calculate p-values and the SLIM method to calculate adjusted p-values (q-values) (H.-Q. Wang, Tuominen, and Tsai 2011). Differentially methylated regions (DMRs) were then defined as regions with at least 25 % methylation difference with a q-value ≤ 0.01.

The segmentation of the RRBS data was done by Abdulrahman Salhab using the R package MethylSeekR with default parameters (Salhab et al. 2018). Using the ChromH3M workflow combinatorial states were identified for partially methylated regions.

2.2.11 mRNA data analysis

After sequencing, the reads (FastQ format) were filtered for low quality (phred score=20). After filtering adapter sequences were trimmed using Trim Galore! (version 0.3.3). The reads were then aligned to the mm10 reference genome using 2-step STAR

alignment. After alignment, PCR duplications were determined using MarkDuplicate from Picard tools (version 1.115). The quality of the data was controlled with the output of RNA-seQC. In order to work with quantitative values in subsequent analysis, featureCounts was used to count and assign the reads to genes from Gencode annotation (vM2). The processing of the data up to this point was done by Dr. Karl Nordström.

Read counts were normalized into counts per million (cpm) mapped reads using the R package edgeR. The data was first filtered for genes with cpm values ≥2.5 in at least two replicates of a group. Composition biases were removed using the *calcNormFactors* function which uses the TMM method for correction of systematic enrichment differences. Biological variance was determined by estimating the dispersion using *estimateDisp* and a generalized linear model was fit using *glmQLFit. glmQLFTest* function tested for the differential expression and computed p-values with the QL F-test. The read counts per gene were normalized for sequencing depth and gene length with the *rpkm* command resulting in the quantitative values Reads Per Kilobase per Million (RPKM). Differentially expressed genes (DEGs) were defined as genes with significant (p-value ≤ 0.01 and FDR ≤ 0.05) difference in gene expression without restricting the fold change. The resulting RPKM values and fold changes were used to represent and analyze the overall gene expression or expression difference, respectively. In addition, the expression difference was visualized using row-z-scores which are defined as numbers of standard deviations from the mean expression of the respective genes. Z-scores were calculated with the formula:

$$z = (x - \mu) / \sigma$$

$x = RPKM_{geneX}$(Sample X)

μ= Mean($RPKM_{geneX}$(all Samples))

σ= Gene expression standard deviation

Heatmaps of row-z-scores were generated using the R package *pheatmap*.

2.2.12 Correlation of gene expression

Correlation between the expression values of the astrocyte mRNA-Seq data with published external mRNA-Seq data was calculated using the R function *cor* which returns Pearson coefficients. The external data was downloaded as processed FPKM values and joined with the data in this thesis based on the gene name.

2.2.13 Annotation of genomic intervals to genes and genomic features

Annotation of ATAC peaks, DARs, CpGs, and DMRs to genes and genomic location was done using ChipSeek which utilizes the HOMER function *annotatePeaks.pl*. The annotation to genes is determined by the distance to the nearest TSS (-1kb to +100bp) and the region is further assigned to the genomic location which the center of the peak/region occupies (e.g. TSS, exon, intron, intergenic).

2.2.14 Integrative analysis

The first step in integrative analyses was to assign DMRs to DARs or DMRs and DARs to DEGs. The subsequent analyses were performed using the intersected genes/ regions. For pairwise comparison, DMRs and DARs were intersected based on the genomic position with an overlap of at least 1bp using the command *intersect* from Galaxy Operation Tools. The assignment of expression values to DNA accessibility or DNA methylation values was done based on gene symbols. Two different methods to link DARs and DMRs to differential gene expression were chosen. First, to retrieve the genes that have distinct chromatin accessibility and distinct DNA methylation, common gene symbols between the DEGs/ DARs/ DMRs lists were found using the online tool *genevenn.* The obtained gene list comprises DEGs with DARs and DMRs that do not necessarily overlap. In the second method, the DMRs that overlap with DARs by at least 1bp are selected and assigned to the respective DEG based on the gene symbol. The epigenetic changes linked to gene expression changes were represented in a heatmap with row scaled values. Row-z-scores were calculated as described in chapter 2.2.11 for each dataset separately using the DNA methylation, RPKM, and cpm values,

respectively. The z-scores were then combined based on identifiers created from the combination of gene symbol and chromosomal position. The rows were clustered using Manhattan distance and average linkage.

2.2.15 Gene Ontology Term enrichment analysis

Gene Ontology (GO) Term enrichment of was performed using DAVID functional annotation in case of DEGs. The GO terms for DMR and DAR associated genes was either analyzed using the online tool GREAT, which annotated the given genomic intervals to genes with the closest TSS, or using the gene annotation from ChipSeek as input for DAVID (see chapter 2.2.13).

The reduced visualization of GO terms was done with REVIGO using GO terms and associated p-values as input. Default parameters were selected (whole UniProt database, *simRel* score, medium sized list). The output was extracted and imported into R to modify the aesthetics (see chapter 2.2.17 on page 58).

2.2.16 Transcription factor binding motif analysis

Overrepresentation of transcription factor binding motifs in DARs, DMRs or specified genomic intervals was assessed using RSAT with following parameters: Cut peak sequences: +/-100; Compare discovered motifs with known motifs from databases: JASPAR core nonredundant vertebrates, cisBP mouse, RSAT non-redundant vertebrates, Hocomoco (Mouse TFs). For the other parameters the default settings were not changed. Trap (multiple sequences) and alibaba2 were used to find motifs in a small number of given genomic intervals as it was the case for the *Il10* locus. For Trap, mouse promoters were selected as background models. Both, Jasper and transfac databases were used for the same region analysis. Alibaba2 was used with default parameters.

2.2.17 Data visualization

Principal component analysis

Principal component analysis (PCA) is a technique used to emphasize biological or technical variation in large data sets. For this thesis, the *prcomp* function was used in R on epigenetic or transcriptomic quantitative values stored in matrices with samples as rows and genes/ CpGs/ genomic intervals as columns. The analysis of the features, contributing to the separation of respective samples, was done by extracting the squared variable loadings of the PCA object (*pca$rotation**2*) and ordering the initial data matrix based on the largest loading values. The PCA object generated by prcomp was visualized using the *autoplot* function from the ggfortify R package.

Heatmaps/ Patternmaps

A heatmap/patternmap converts values into colors to visualize and simplify large datasets. For this thesis, the *pheatmap* function from the pheatmap R package was used with varying parameters for the clustering of the rows/columns.

Venn diagrams

The overlap of genomic intervals was assessed using the *Venn diagram from bed files* function from the internal Galaxy EpiGenetics ToolBox. The Intersection of gene names from different gene lists was done using GeneVenn.

Hierarchical clustering

Unsupervised hierarchical clustering of epigenetic and/or transcriptomic data were generated using the *hclust* function in R. The dendrograms were visualized using *fviz_dend* function from the factoextra R package.

Violin plots, scatter plots

Violin plots were generated using *ggplot* with the *geom_violin* function. The distribution difference of respective data points was tested with Wilcoxon rank sum test using the *wilcox.test* function

Scatter plots

Scatter plots were generated using *ggplot* with the *geom_point* function. Revigo output was imported into R and also visualized using the *geom_point* function, with point sizes representing -log10(pValue).

Bar diagrams

Bar diagrams were generated using the *ggplot geom_bar* function or using the column chart function in excel.

Boxplot

Bar diagrams were generated using the *ggplot geom_boxplot* function or the R boxplot function.

Circular genome maps

Circular visualization of the genomic distribution of DMRs was done using the circlize R package. Initially, *circos.initializeWithIdeogram* is used to generate circular chromosome ideograms. Next, genomic positions are indicated as points in circular tracks using *circos.genomicTrackPlotRegion* with the *circos.genomicPoints* command.

Protein-Protein Interaction Networks

Protein-Protein interactions were investigated using the STRING database. The resulting networks were exported as tabular formats and imported as network into Cytoscape. Additionally, a table containing information whether the respective gene is associated with a DMR or DAR is imported and joined with the network based on gene symbol.

Genome Browser visualization

The data is represented in in the genome browser IGV. Biological replicates were merged into one track using the function *Overlay Tracks*, or visualized as single replicates in separated tracks. Methylation values range was set from 0-100, methylation coverage range was set from 0-30. The chromatin and cDNA enrichment

data were visualized either in autoscale mode or the scale range was set for a minimum of background.

2.2.18 Immunohistochemistry

The immunohistochemistry was done by Dr. Carmen Kasakow and Dr. Laura Stopper from the Kirchhoff group. hGFAP-EGFP mice were anesthetized with ketamine/xylazine 140 mg/kg and 10 mg/kg body weight to perform the subsequent intracardial perfusion wit PBS and 4 % formaldehyde in 0.1 M phosphate buffer (pH 7.4). The brain was isolated and fixed overnight. Sagittal slices were prepared with a vibratome (Leica VT1000S) and permeabilized and blocked. The slices were incubated with primary antibody over night at 4 °C, washed, and incubated with the secondary antibody. Subsequently the slices were stained with DAPI or ToPro3 and mounted with ImmuMount. Images were recorded with confocal laser scanning microscopy (LSM 710). z-Stacks of images were taken at 1 µm intervals, processed with Fiji and displayed as maximum intensity projections. For quantitative analysis of the signals, regions-of-interest (ROIs) were selected based on nuclear signal. The means of three consecutive z-layers were determined and compared with GraphPad Prism 7. The Intensity values were tested for normality using the Shapiro-Wilk normality test and the ROUT-test for outliers.

2.2.19 Statistical Analysis

Statistical analysis of the sequencing data was done in R using the statistical tests implemented within the respective functions/packages. The variance of ATAC-Seq and mRNA-Seq data were tested *glmQLFTest* command which uses QL F-test to compute p-values. The DNA methylation data was tested for significant difference using the function *calculateDiffMeth* which uses logistic regression to calculate p-values and the SLIM method to calculate adjusted p-values (q-values). The distribution difference in of data represented as violin plots were tested with Wilcoxon rank sum test using the *wilcox.test* function. The significant enrichment of GO Terms was determined by

DAVID with a P-value cut off of ≤0.01. Immunofluorescence data were analyzed with the Mann-Whitney-U-test with the following p-values * p= 0.01-0.05, ** p= 0.001-0.01, *** p= 0.0001-0.001, ****p< 0.00001. Data are represented as means ± STDEV of single cell values.

Chapter 3.

Results and Discussion

3.1 Epigenetic control of region-specific transcriptional programs in mouse cerebellar and cortical astrocytes

3.1.1 Introduction

For a long time, astrocytes were viewed as a functional homogenous cell population to support neurons throughout the brain. However, in the past decade functional diversity of astrocytes in distinct brain regions became evident (W. Todd Farmer and Murai 2017). It was demonstrated that gene expression of astrocytes follows the dorsoventral axis. In addition, astrocytes preferentially stimulate neurite growth and are involved in the synaptic activity of region matched neurons (Morel et al. 2017). Moreover, striatal and hippocampal astrocytes were shown to have distinct and region-specialized properties like altered morphology, electrophysiology, and Ca^{2+}-signaling within neural circuits (Chai et al. 2017). The molecular heterogeneity of intraregional subpopulations was shown to be highly dependent on the interaction with the neuronal environment. For instance, disruption of neuronal layers in the neocortex subsequently leads to the loss of layer-specific properties of cortical astrocytes (Lanjakornsiripan et al. 2018). In the cerebellum, neuron derived sonic hedgehog (Shh) controls the gene expression profile of Bergmann glia and may drive the gene expression in velate astrocytes towards Bergmann glia profiles (W. T. Farmer et al. 2016).

However, the underlying mechanisms ofthe regional specification of astrocytes remain to be uncovered. Epigenetic modifications are key mechanism to establish and maintain cell identity by modulating gene expression. Functional and cell-specific program changes in the genome can be approached by exploring genome wide

distribution of DNA methylation and its local variation. In addition, DNA methylation profiles allow to indirectly infer the developmental history of cells (see chapter 1.1.2 on page 15). Next to DNA methylation, chromatin modifications are excellent indicators of the functional state of genes and genomes. The integration of epigenetic modifications paves the way to understand the regulatory programs of regional astroglial populations to identify differences explaining their functional, epigenomic, and regional diversity.

This thesis addressed the molecular diversity of astrocytes from two distinct adult brain regions using genome wide approaches. For this purpose, using the well documented hGFAP-EGFP mouse model (EGFP expression under the control of the human glial fibrillary acidic protein promoter; Nolte et al. 2001; **Figure 3-1**), protoplasmic astrocytes from the cerebral cortex and astrocytes from the cerebellum (mainly Bergmann glia) were isolated using fluorescence activated cell sorting (FACS). Subsequent comprehensive epigenomic and gene expression profiling followed by integrative analyses resolved commonalities and differences of both astroglial populations (**Figure 3-2**).

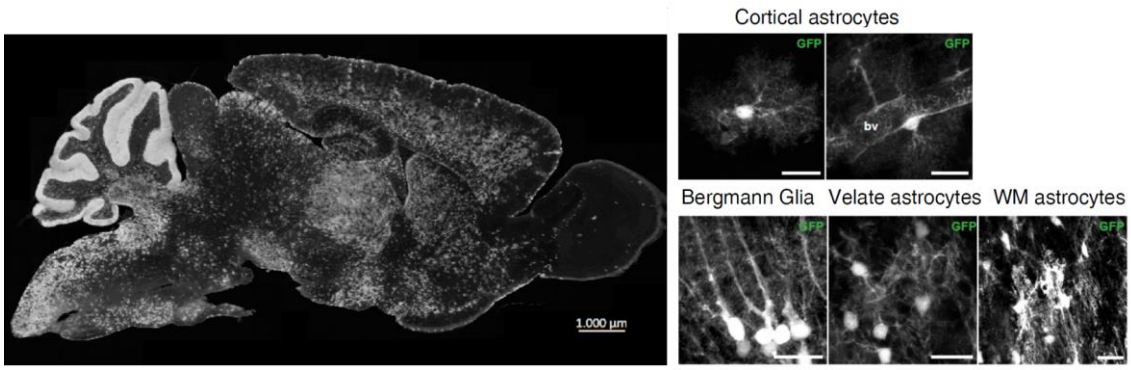


Figure 3-1 Sagittal illustration of hGFAP-EGFP mouse brain *(left) with EGFP⁺ astrocytes in CTX (upper panel right) and CB (lower panel right). bv=Blood vessel, WM=White matter; scale bars=20 µm. Immunofluorescence images were produced by Dr. Carmen Kasakow in terms of a cooperation.*

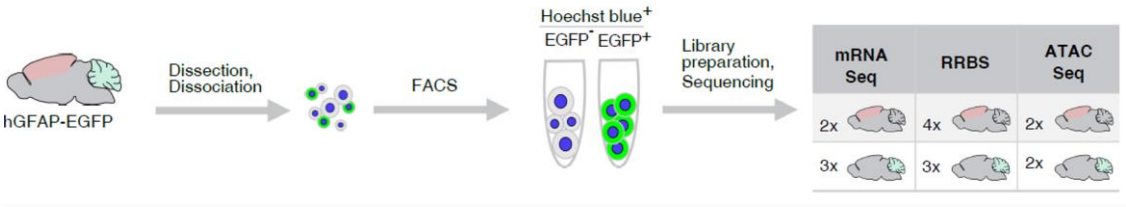


Figure 3-2 Schematic of the experimental setup *to characterize molecular astrocyte diversity.*

3.1.2 Isolation of astrocytes from the cerebral cortex and the cerebellum

The isolation of astrocytes required dissociation of the tightly connected brain tissue into single cell suspensions. The cells were prepared as described in chapter 2.2.1 on page 44 and sorted based on the fluorescence signal of EGFP and Hoechst blue to obtain astrocytes. The sorted population was evaluated under a fluorescence microscope which revealed intact cells with EGFP expression in the cytosol (**Figure 3-3**)

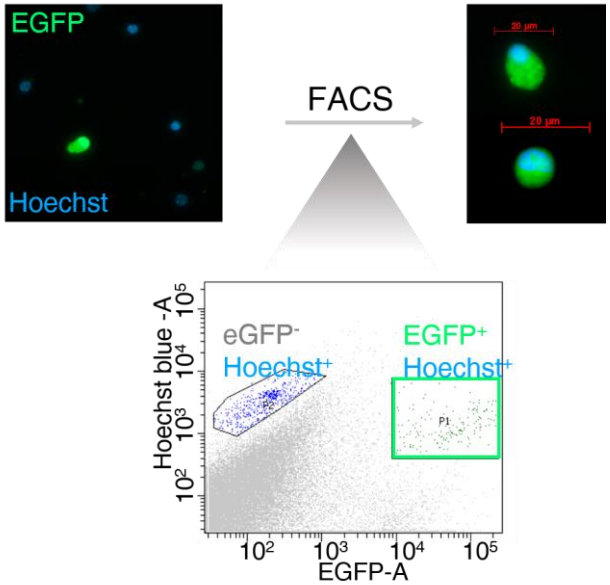


Figure 3-3 Fluorescence activated cell sorting (FACS) of hGFAP-EGFP mouse cerebellum. *The single cell suspension (left) consists of EGFP⁺ and EGFP⁻ cells. Using FACS the EGFP⁺ cells were gated for high EGFP and Hoechst blue signals and were collected for subsequent procedures.*

To characterize the molecular programs of astroglial populations from CTX and CB, the sorted cells were used for the preparation of replicate libraries for sequencing on a Illumina HiSeq2500 to profile gene expression by mRNA-Seq, genome-wide DNA methylation by RRBS, and DNA accessibility by ATAC-Seq (see chapter 2.2.4 above, chapter 2.2.5 on page 47, and chapter 2.2.6 on page 49 for more details). In addition, RRBS and ATAC-Seq libraries were constructed with $NeuN^+$ and $NeuN^-$ nuclei[6] to compare astroglial profiles with neuronal and glial profiles. Moreover, oligodendrocytes were sorted based on high DsRed fluorescence signal using a single cell suspension prepared from the cortex of a PLP-DsRed1 mouse and used for genome-wide DNA methylation comparisons. Before analyzing the molecular diversity between astrocytes from distinct brain regions, the data quality and purity of the samples had to be assessed. Various features of the mapped sequencing reads were analyzed to assess the data quality. For mRNA-Seq the exonic rate was similarly high and the intronic rate similarly low among the samples. The number of detected genes was comparably high and almost no ribosomal RNA was detected. Taken together, the sequencing data represented transcriptomes with high quality (**Table 9**).

RRBS on the sorted cells/nuclei covered 3-4 million CpGs per sample and yielded a mean coverage above 10x. This assured a broad and representative DNA methylation profile (**Table 10**). The reads from ATAC-Seq had high mapping rates (>90%) in the astrocyte samples and slightly lower mapping rates for the NeuN-sorted nuclei (>80%). Moreover, the fraction of reads in peaks was higher in astrocytes from the cortex than in the cerebellum, indicating distinct distribution of accessible fragments in astrocytes (**Table 11**).

[6] NeuN is a neuron-specific nuclear protein that can be targeted by antibodies to label and isolate neuronal cell nuclei.

Table 9 Sequencing Information on mRNA-Seq

Sample	Number of Reads	Mapping rate	Ribosomal RNA rate	Detected Genes	Exonic rate	Intronic rate
CB1	37289395	0.939	0.003	21,091	0.564	0.155
CB2	95298868	0.958	0.002	23,536	0.677	0.118
CB3	56239574	0.958	0.002	22,062	0.617	0.149
CTX1	69592154	0.97	0.003	21,525	0.599	0.133
CTX2	79904620	0.964	0.004	24,653	0.653	0.141

Table 10 Sequencing Information on RRBS

Sample	Number of Reads	Number of Sites	Number of Sites with coverage≥10x	Mean coverage
CB1	43525776	3554016	1237528	10.91
CB2	52660387	4165161	1720049	12.09
CB3	94150304	4462068	1787610	18.78
CTX1	43491013	4068652	1400616	10.06
CTX2	63717524	2972471	1129551	17.92
CTX3	51185762	3721804	1409414	10.85
CTX4	66036845	4216310	1787610	12.38
NeuN_positive	67338205	4187436	1927446	14.91
NeuN_negative	61340311	3954107	1822102	15.72
Oligodendrocyte	60313767	4308408	1699619	10.94

Table 11 Sequencing Information on ATAC-Seq

Sample	Number of Reads	Mapping rate	Fraction in peak
CB2	115455035	0.97008	0.096
CB4	108172651	0.918528	0.109
CTX3	73433999	0.974331	0.157
CTX4	81505152	0.972028	0.162
NeuN_positive	72577417	0.804753	0.050
NeuN_negative	86352354	0.837109	0.068

After evaluation of the sequencing quality, the purity of the EGFP$^+$ cells had to be confirmed. Examination of the mRNA-Seq data for molecular signatures specific for astrocytes, endothelial cells, microglia, neurons, oligodendrocytes or oligodendrocyte precursor cells (OPCs) revealed a high enrichment of astrocyte-specific genes such as *Aqp4*, *Gja1* (connexin 43), *Slc1a3* (Glast), and *Gjb6* (connexin 30) (**Figure 3-4A**). In contrast, endothelium-, microglia-, and neuron-specific genes were depleted. Traces of oligodendroglial and OPC marker expression were detected in both astrocyte populations with slightly higher presence in the cortical astrocytes. This can be assigned to a minor contamination with highly transcribed oligodendroglial mRNAs that were floating in the single cell suspension after brain dissociation and a weak GFAP-EGFP transgene activity in few OPCs. In addition to the validation of marker expression, the global transcriptome data was correlated to published transcriptomes of astrocytes, and other major brain cell types (**Figure 3-4B**). The astroglial transcriptome profiles of this thesis correlated highly with transcriptomes of astrocytes derived from other isolation methods. Furthermore, the correlation with the transcriptomes of other cell types, such as neurons, microglia or endothelium was low.

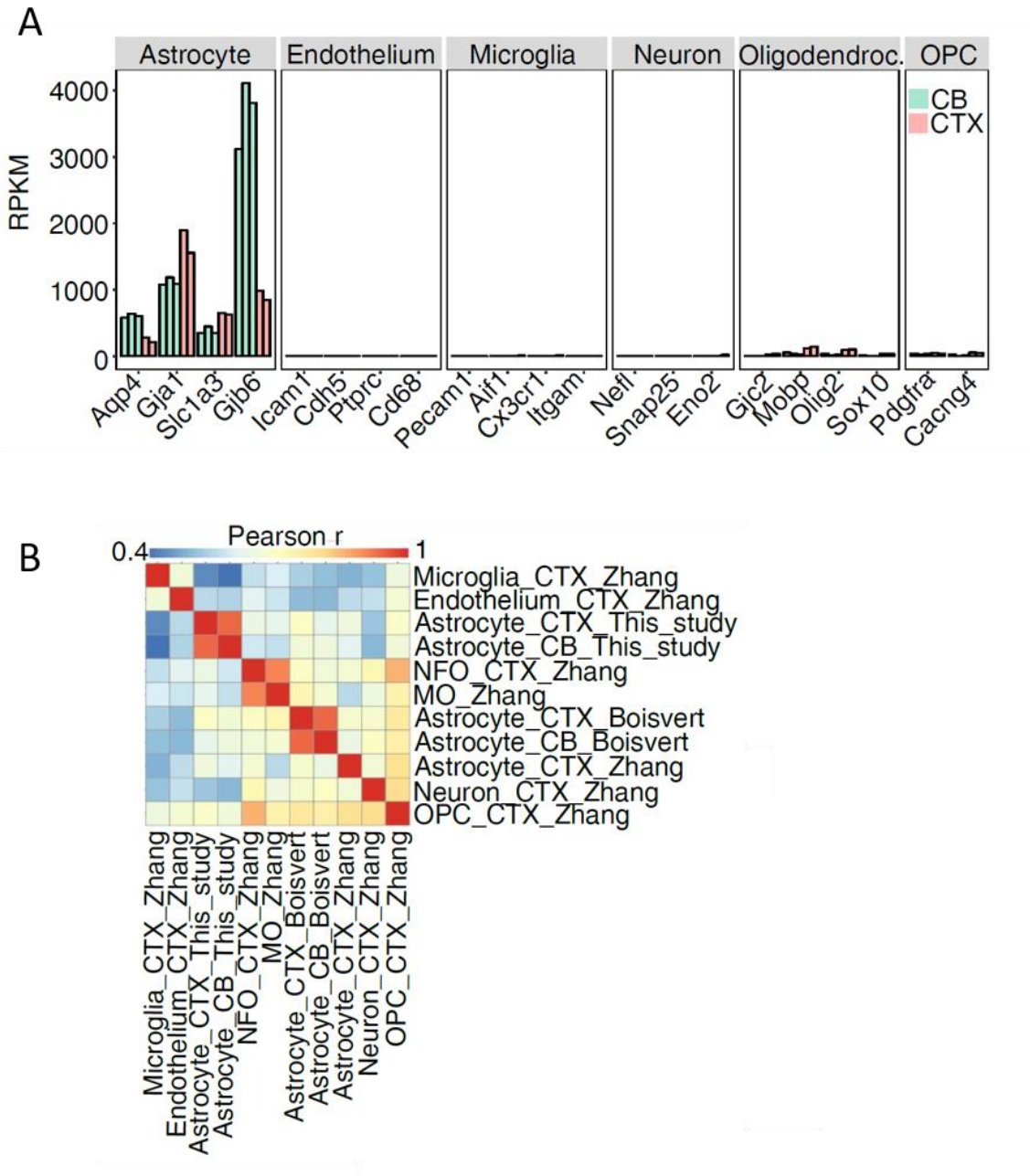


***Figure 3-4 Validation of astroglial transcriptomes. A**. Gene expression of cell type-specific genes measured in RPKM. **B**. Correlation of mean gene expression data (this thesis) with published data sets (Boisvert et al. 2018; Y. Zhang et al. 2014). CTX=cortex, CB=cerebellum, NFO=newly formed oligodendrocytes, MO=myelinating oligodendrocytes, OPC=oligodendrocyte precursor cell.*

Besides the transcriptomes, the genome-wide DNA methylation and accessibility was compared between astrocytes and other major brain cell types. Additionally, the epigenetic signatures around cell type-specific genes were evaluated to confirm astroglial identity of the sorted cells (**Figure 3-5**). Unsupervised principal component analysis (PCA) of the DNA methylation of all covered CpGs revealed a clear separation of astrocyte methylomes from $NeuN^+$, $NeuN^-$, and oligodendrocyte methylomes in the first major component explaining 33% of the DNA methylation variance in the data. Examination of the methylation level of the 500 most variable CpGs contributing to the separation showed astrocyte-specific hypomethylation of these CpGs (**Figure 3-5A**). Of note, a part of these CpGs are annotated to genes encoding members of the Notch-, Tgfß-, and Wnt-signaling pathways known to promote astrogliogenesis (e.g. Fgfr3, Hes1, Notch2/3, Dll4, Ncor2, Prdm16, Sfrp1, Sox9, Tcf7l1, Tcf7l2). Similarly, PCA of genome-wide DNA accessibility binned in 10 kb windows clearly separated astrocytes from NeuN+ and NeuN- cell populations in the first principal component explaining 84% of the variation in the data (**Figure 3-5B**). In order to investigate the commonalities of open regions in both astrocyte population, shared open regions (n=20,648) were subjected to a transcription factor motif enrichment analysis. A significant enrichment for binding sites of the nuclear factor 1 (Nfi) family was observed in 24% of the common open regions (p-Value = 10^{-413}) (**Figure 3-5C**). Nfi members are essential factors for astroglial differentiation by regulating the expression of astrocyte-specific genes. Furthermore, Nfia is involved in DNA-demethylation of gliogenic gene promoters such as *Gfap* and *Olig1* and promotes astrogliogenesis (Namihira et al. 2009; Sanosaka et al. 2017). Both, genome-wide DNA methylation and genome wide DNA accessibility, indicated that astrocytes from CTX and CB share a broad spectrum of epigenomic signatures involved in establishment and maintenance of astrocytic cell identity.

Examination of local epigenetic signatures at astrocyte-specific genes such as *Aldh1l1*, *Slc1a3*, and *Gfap* showed low DNA methylation and high DNA accessibility at regulatory elements in both astrocyte populations, while neurons and oligodendrocytes showed high DNA methylation and little to no DNA accessibility at these loci (**Figure 3-5D**). In contrast, the promoters of neuron-specify *Tubb3* gene and

oligodendrocyte-specific *Mag* gene were highly methylated and closed in both astrocytes, while neurons and oligodendrocytes showed unmethylated and open promoters at respective gene (**Figure 3-5E**).

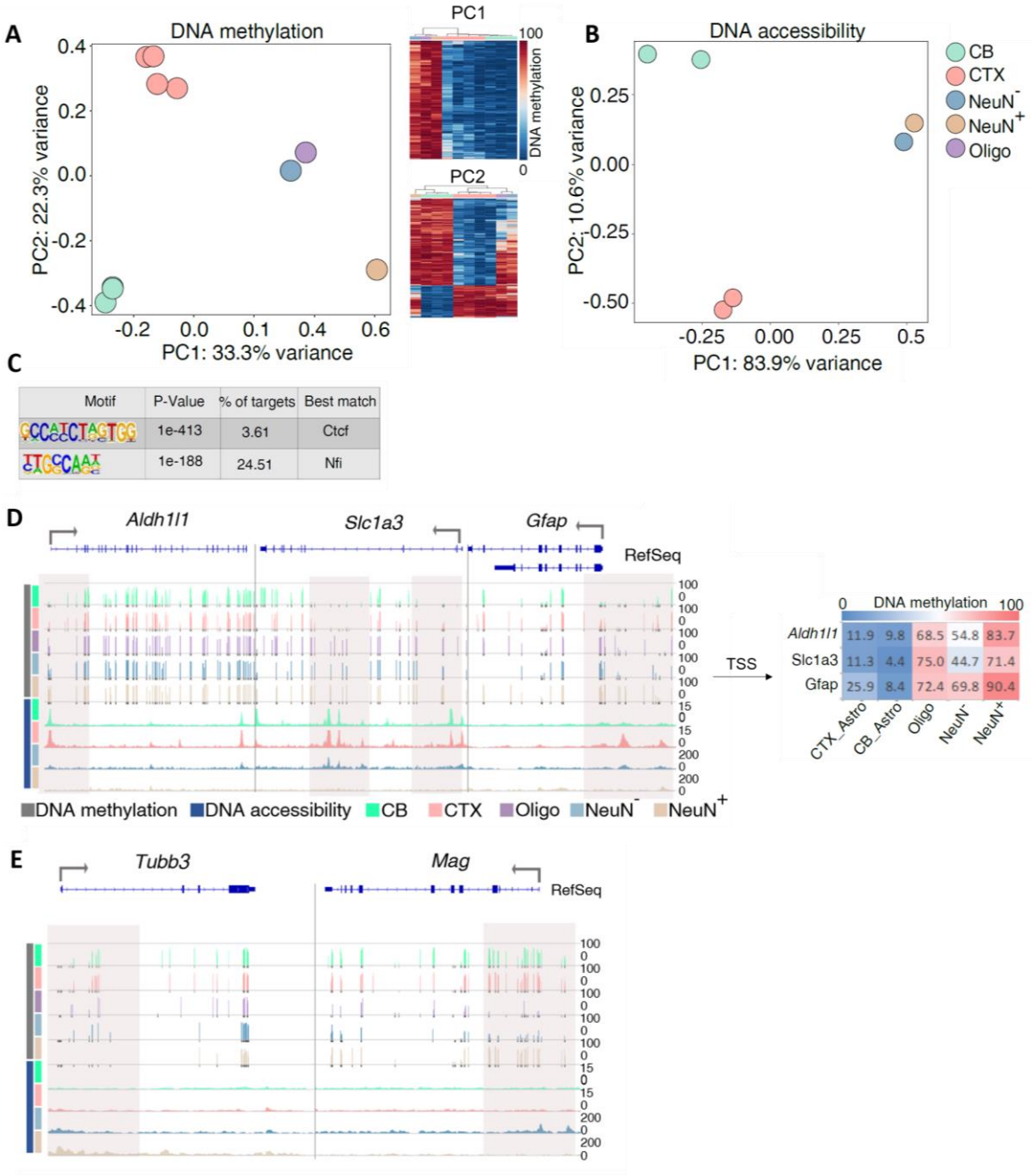


Figure 3-5 Validation of astroglial epigenomes. ***A***. *PCA of genome wide DNA methylation data with heatmaps representing the methylation values of the 500 most variable CpGs in PC1 and PC2.* ***B***. *PCA of genome wide DNA accessibility data.* ***C***. *Enrichment of binding motifs in common open regions of CB and CTX astrocytes.* ***D***. *DNA methylation and DNA accessibility of Aldh1l1, Slc1a3, Gfap (left) with mean DNA methylation values of the transcription start site ±1kb of Aldh1l1, Slc1a3, Gfap in respective cell type (right).* ***E***. *DNA methylation and DNA accessibility of neuronal Tubb3 and oligodendroglial Mag.*

Taken together, the reporter- and FACS-based isolation enabled generation of highly pure astrocyte epigenomes and representative transcriptomes, allowing further integrated analyses of cell type-specific epigenetic program differences. Before performing integrated analyses of the datasets, each dataset was analyzed separately to characterize the differences between CTX and CB astrocytes on gene expression level, DNA methylation, and DNA accessibility.

3.1.3 Astrocytes from cerebellum and cortex show common transcriptional programs

The first pairwise comparative analysis of CTX and CB astrocyte transcriptome data showed an overall high correlation of gene expression levels (r=0.89; **Figure 3-6A**). The scatterplot already showed some genes that are considerably higher expressed in CTX or CB. In CTX astrocytes genes encoding Apolipoprotein E (*Apoe*), Carboxypeptidase E (*Cpe*), Insulin-like growth factor binding protein-2 (*Igfbp2*), and Regulator of cell cycle (*Rgcc*) were higher expressed than in CB astrocytes. These genes are involved in metabolism and cell proliferation and have been described to be highly expressed in cortical astrocytes (Y. Zhang et al. 2014; X. Liu et al. 1994, 2; Sakers et al. 2017; Boyles et al. 1985). However, the decreased expression of these genes in the CB and, in turn, the respective potential functional decrease has not been discussed before. In addition, oligodendroglial genes such as *Mag*, *Mog*, *Cldn11* were found to be higher expressed in CTX astrocytes. Conversely, the Bergmann glia-specific Tgfß ligand *Gdf10* and the neuropeptide Y (*Npy*) were considerably higher expressed in CB astrocytes than in CTX astrocytes. Npy was recently shown to be expressed in Bergmann glia in a topographic zone-restricted manner following the parasagittal domains of Purkinje cells (Reeber, Arancillo, and Sillitoe 2018). The function of Gdf10 in Bergmann glia is not yet fully understood. However, it was shown that *Gdf10* is directly regulated by neuron-derived Shh protein, a crucial signal for the establishment and maintenance of Bergmann glia gene expression profiles, which indicated the involvement of Gdf10 in Bergmann glia specification (W. T. Farmer et al. 2016; Mecklenburg et al. 2014, 10). Other genes that deviated strongly from the expression level in CTX astrocytes were the radial glia

marker *Hopx* and Vimentin (*Vim*), *S100a10*, the transcription factors *Zic1* and *Msx2*, the protease inhibitor *A2m* and a thyroid hormones transporter, *Slc04a1*, that is usually expressed in cortical microglia. These genes have multiple functions, but the functions in CB astrocytes are not characterized.

Next, the highest expressed genes in CTX and CB astrocytes were compared (**Figure 3-6B**, **Figure 3-6C**). The highest expressed gene in both cell types was *Gm10800*, a predicted gene with unknown function. Inspecting the genes, 12 genes were present in both top 20 list (*Apoe*, *Cst3*, *Gm10800*, *Gm10801*, *Gm21738*, *Gpm6a*, *Gpr37l1*, *Itm2b*, *Mt1*, *Mt3*, *Plpp3*, *Sparcl1*). In general, these 12 highest transcribed genes were similarly high transcribed in both regional astrocytes. Exceptions constitute the gene encoding the inhibitor of cysteine proteinases Cst3 (higher expressed in CTX astrocytes), and genes with higher expression in CB astrocytes such as *Sparcl1* (synapse formation), *Itm2b* (deposition inhibition of beta-amyloid), and *Gpm6a* (neurogenesis, synapse formation). These top common genes represent some of astroglial core functions such as neuroprotection (*Gpr37l1*, *Mt1*, *Mt3, Itm2b, Cst3*), and synapse formation and elimination (*Apoe*, *Sparcl1, Gpm6a*). Although highly expressed, *Gm10800*, *Gm10801*, *Gm21738*, and *Plpp3* are still of poorly defined function in astrocytes.

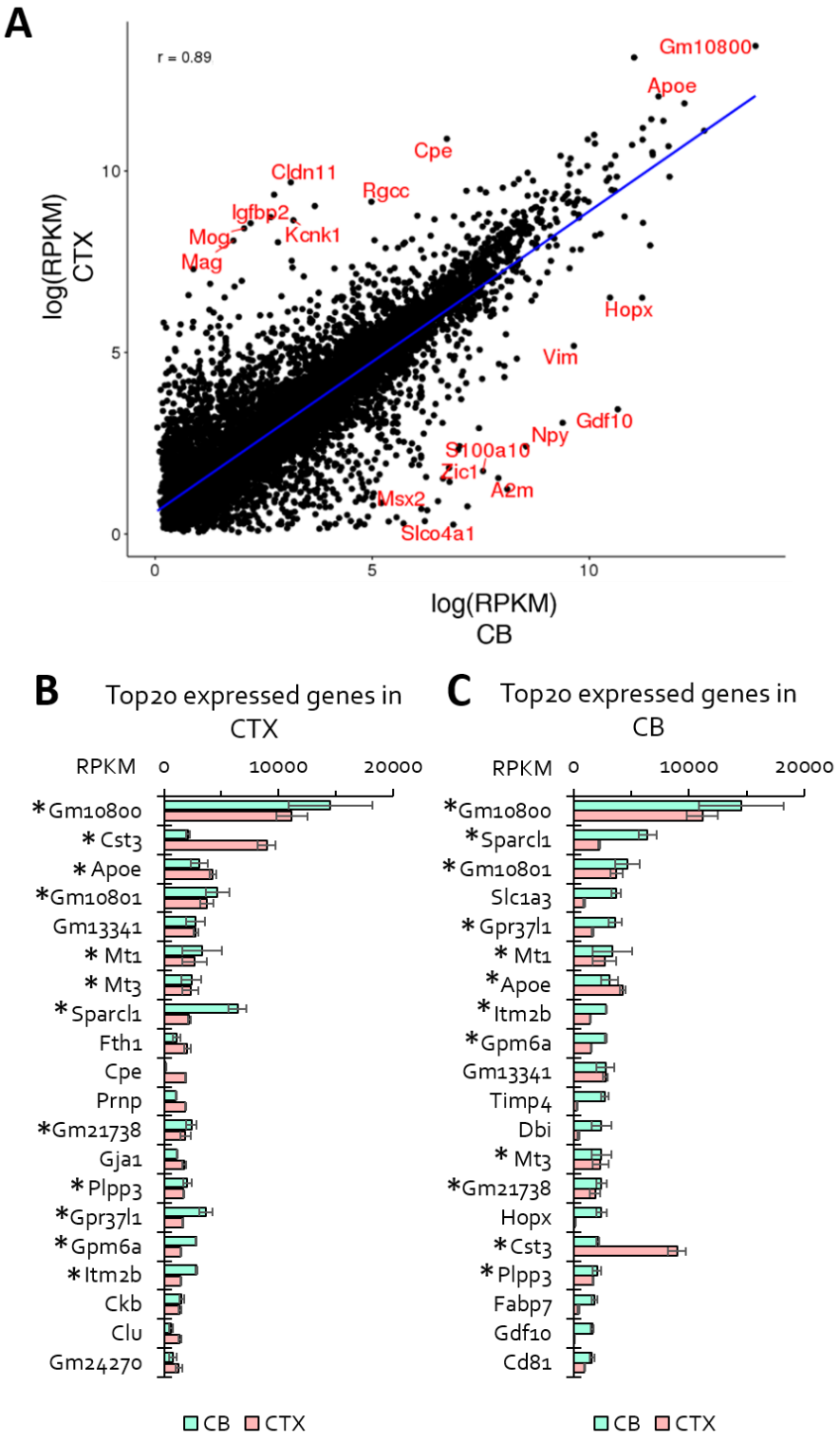


Figure 3-6 Pairwise comparison of CTX and CB astrocyte transcriptomes*. A. Correlation of gene expression levels of all detected genes. B. Expression values of the 20 highest expressed genes in CTX astrocytes. C. Expression values of the 20 highest expressed genes in CB astrocytes. r= Pearson correlation coefficient; mean ± STDEV, * indicates genes present in both top20 lists.*

Analysis of the global gene expression showed that the overall transcriptional activity was highly similar (**Figure 3-7A**) between both astrocyte populations. When classifying the detected genes (n=12,801) based on the mean expression into lowly expressed genes (RPKM > 0 ≤ 10), medium expressed genes (RPKM > 10 ≤ 100), and highly expressed genes (RPKM >100), no apparent quantitative differences was observed (**Figure 3-7B**). In addition, the assessment of global gene expression showed that the majority of the genes were lowly expressed in both astrocyte types. Of note, more than 90% of lowly expressed genes in CTX or CB were also lowly expressed in the other cell type and are involved in metabolic processes (**Figure 3-8**). Functional analysis of the lowly expressed genes only in one of the two cell types revealed processes related to development. Comparison of the medium and highly expressed genes indicated that, in percentage terms, the intersection in these classes was smaller than in lowly expressed genes (70-83%; **Figure 3-9**, **Figure 3-10**). These intersected genes with similar gene expression levels were strongly enriched for metabolic processes, suggesting that the majority of homeostatic and core functional processes are shared between CTX and CB astrocytes. However, this comparison also pointed out differences in expression of genes involved in various functions. The enrichment analysis of non-overlapping genes of medium expressed genes resulted partially in the same GO Terms (e.g. sensory perception, cellular processes, metabolic processes), which means that distinct gene sets were contributing to the same function (**Figure 3-9**). The enrichment analysis of non-overlapping genes of highly expressed genes was more diverse between CTX and CB astrocytes. In CB astrocytes processes related to transport were found to be enriched, while in CTX astrocytes the enriched processes were related to axon ensheathment, and development (**Figure 3-10**).

In summary, the examination and first comparisons of CTX and CB astrocyte transcriptomes revealed a broad spectrum of commonalities, while also pointing to gene expression differences related to various functions.

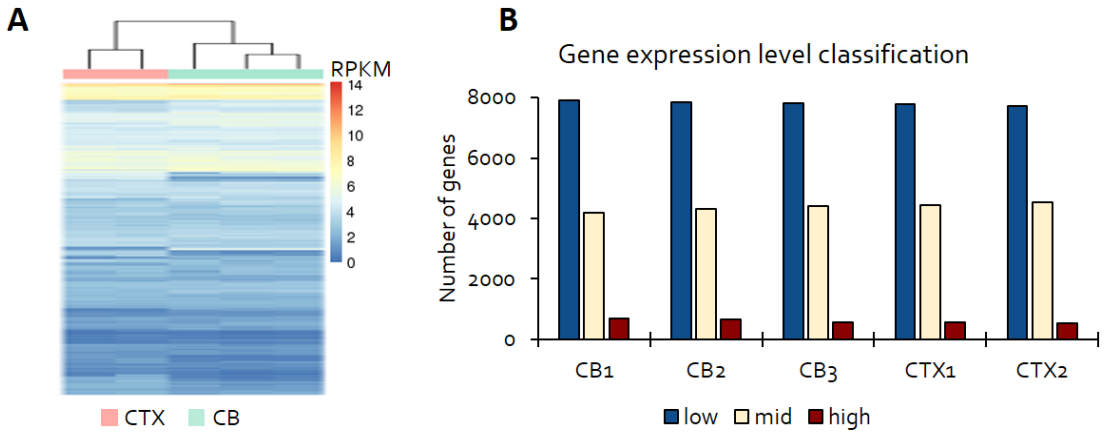
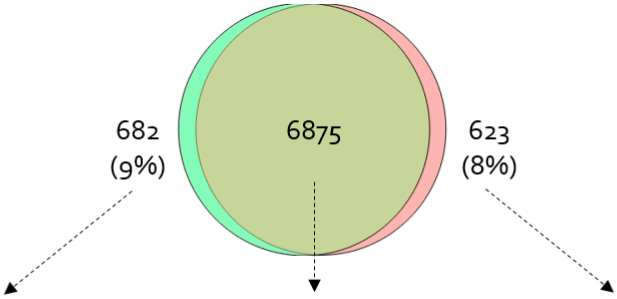


Figure 3-7 Global gene expression in CTX and CB astrocytes. ***A***. *Heatmap representing all 12,801 detected genes measured in logRPKM.* ***B***. *Display of numbers of lowly, medium, and highly expressed genes in CB and CTX astrocytes. Low: RPKM > 0 ≤ 10, mid: RPKM > 10 ≤ 100, high: RPKM >100.*

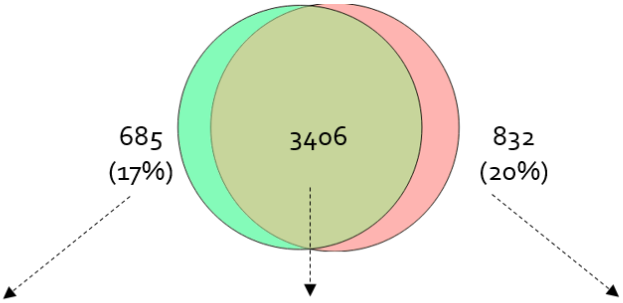


GO Biological process	-log (pValue)
nervous system development	10.1
neurogenesis	9.7
system development	9.4
regulation of developmental process	8.5
generation of neurons	8.3
multicellular organism development	8.1
anatomical structure development	7.8
regulation of cellular component organization	7.6
sensory perception of chemical stimulus	7.4
sensory perception of smell	7.4

GO Biological process	-log (pValue)
sensory perception of chemical stimulus	106.0
sensory perception of smell	106.0
G protein-coupled receptor signaling pathway	66.1
sensory perception	62.3
macromolecule metabolic process	47.6
cellular macromolecule metabolic process	42.2
cellular metabolic process	39.4
nitrogen compound metabolic process	39.2
cellular component organization or biogenesis	39.1
regulation of metabolic process	37.7

GO Biological process	-log (pValue)
sensory perception of chemical stimulus	6.3
sensory perception of smell	5.3
positive regulation of biological process	3.2
positive regulation of cellular process	3.1
nervous system development	2.7
cellular process	2.6
positive regulation of developmental process	2.4
regulation of cell communication	2.3
regulation of signaling	2.3
sensory perception	2.1

Figure 3-8 Intersection and GO Term enrichment of lowly expressed genes in CB and CTX astrocytes.

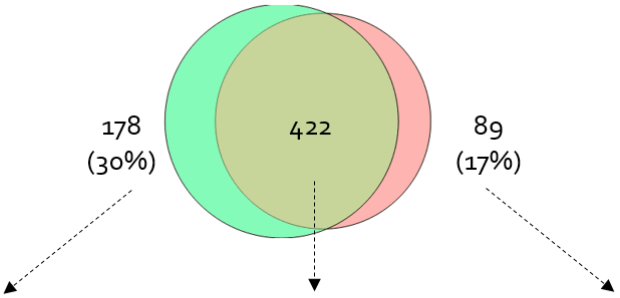


GO Biological process	-log (pValue)
sensory perception of chemical stimulus	7.6
sensory perception of smell	6.4
nervous system development	6.1
cellular process	5.6
metabolic process	4.5
positive regulation of biological process	4.3
positive regulation of multicellular organismal process	4.3
regulation of multicellular organismal process	4.0
cellular metabolic process	3.9
positive regulation of developmental process	3.8

GO Biological process	-log (pValue)
cellular metabolic process	103.3
metabolic process	98.7
organic substance metabolic process	87.6
primary metabolic process	83.3
nitrogen compound metabolic process	78.8
sensory perception of chemical stimulus	65.8
macromolecule metabolic process	62.5
sensory perception of smell	61.7
cellular nitrogen compound metabolic process	53.4
organonitrogen compound metabolic process	51.5

GO Biological process	-log (pValue)
organic substance metabolic process	11.3
sensory perception of chemical stimulus	10.6
metabolic process	10.4
sensory perception of smell	10.3
small molecule metabolic process	10.2
organonitrogen compound metabolic process	9.5
primary metabolic process	9.0
cellular metabolic process	8.8
negative regulation of cellular process	8.5
cellular process	8.4

Figure 3-9 Intersection and GO Term enrichment of medium expressed genes in CB and CTX astrocytes.



GO Biological process	-log (pValue)
nitrogen compound transport	4.3
organic substance metabolic process	4.0
transport	3.7
metabolic process	3.6
intracellular protein transport	3.5
cellular metabolic process	3.4
amide transport	3.3
organic substance transport	3.3
establishment of localization	3.2
protein transport	3.1

GO Biological process	-log (pValue)
ATP metabolic process	20.8
purine nucleoside triphosphate metabolic process	20.2
nucleoside triphosphate metabolic process	20.1
peptide metabolic process	20.0
purine ribonucleoside triphosphate metabolic process	19.9
ribonucleoside triphosphate metabolic process	19.6
translation	19.2
peptide biosynthetic process	18.3
organonitrogen compound biosynthetic process	18.2
cellular amide metabolic process	18.2

GO Biological process	-log (pValue)
nervous system development	4.6
ensheathment of neurons	3.8
axon ensheathment	3.8
myelination	2.5
system development	1.9
small molecule metabolic process	1.9
cellular process	1.5
multicellular organism development	1.4
organic substance biosynthetic process	1.4

Figure 3-10 Intersection and GO Term enrichment of highly expressed genes in CB and CTX astrocytes.

3.1.4 Transcriptional differences between astrocytes from cerebellum and cortex

The previous chapter described shared gene expression profiles between CTX and CB astrocytes, and core astrocyte functions, while also pointing to differences between them (e.g. the dendrogram in **Figure 3-7A**). The following chapter describes significant gene expression differences between CTX and CB astrocytes that were identified performing a differential analysis using the R package edgeR (see chapter 2.2.11 on page 53). The analysis resulted in 3,104 differentially expressed genes (DEGs) with 1,727 genes higher expressed in CTX and 1,377 genes higher expressed in CB astrocytes (**Figure 3-11**). The main biological processes associated with these DEGs were related to ion transport, metabolism, cytoskeleton organization, glial morphogenesis, signaling and development. Of note, many terms (e.g. ion transport, cell migration, nervous system development) were enriched in both populations when performing GO term enrichment analysis with genes that were regionally higher expressed. This suggested a similar functionality of the cells but with different gene sets involved in these processes. In addition to shared functionalities, genes higher expressed in CB were involved in nucleosome and cilium assembly, and actin filament organization. On the other hand, genes higher expressed in CTX were involved in axon guidance, lipid metabolism, and filopodium assembly.

Figure 3-11 Differentially expressed genes between CTX and CB astrocytes. *On the left, the heatmap displays the 3,104 identified differentially expressed genes. Row z-scores were calculated using log2 of RPKM. On the right, GO terms for higher expressed genes in CB (top) and CTX astrocytes (bottom). Red boxes indicate common terms, yellow boxes indicate unique terms for respective brain region.*

When classifying the identified DEGs based on the mean expression into lowly expressed genes (RPKM > 0 ≤ 10), medium expressed genes (RPKM > 10 ≤ 100), and highly expressed genes (RPKM >100), 58%-62% of DEGs were classified as lowly expressed, 31%-37% of DEGs were classified as medium expressed, and 5%-7% of DEGs were classified as highly expressed. Comparing the cell types, one can observe a trend that in cortical astrocytes more DEGs were expressed at medium level and less DEGs were expressed at low or high level (**Figure 3-12A**). The comparison of expression levels of the respective differentially expressed genes and clustering them into five clusters based on k-means clustering resulted in the observation that the majority of the DEGs were found in the same class (**Figure 3-12B**). 1,991 DEGs were lowly expressed in both CTX and CB astrocytes (cluster 2), 730 DEGs were expressed at medium level in both CTX and CB astrocytes (cluster 3), and 108 DEGs were highly expressed in both CTX and

CB astrocytes (cluster 4). Cluster 1 represented 413 DEGs that were lowly expressed in CTX but had medium expression levels in CB. Cluster 5 summarized 128 genes that were expressed at a medium level in CTX astrocytes but at high levels in CB. This clustering of gene expression levels suggested a gradual expression difference between CTX and CB astrocytes for the majority of the DEGs rather than a difference based on presence and absence of these genes. In fact, only 13 genes were found to be regionally exclusive expressed (**Table 12**; criteria for exclusive expression: mean RPKM > 1 in one cell type while RPKM < 0.1 in the other). These "uniquely expressed" genes are involved in various functions and mostly are not well characterized in the brain. Noteworthy, as shown in **Table 12**, the expression of these genes was fairly low in the respective astrocyte type. One exception forms *Emx2*, a homeobox-containing transcription factor gene involved in early spatially specification of the forebrain.

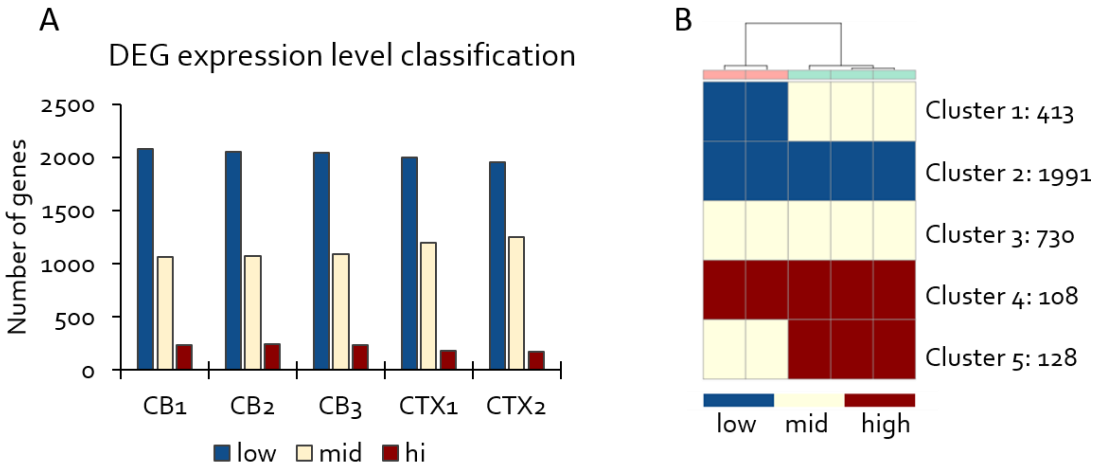


Figure 3-12 Gene expression level of differentially expressed genes between CTX and CB astrocytes. A*. Number of differential genes classified into lowly, medium, and highly expressed.* ***B****. Differential genes clustered based on gene expression level into 5 clusters. Low: RPKM > 0 ≤ 10, mid: RPKM > 10 ≤ 100, high: RPKM >100.*

Table 12 Exclusevly expressed genes in CTX or CB astrocytes. *LogFC=log2(foldchange). Colorcode represents low to high expression (blue to red). Gene symbols are colorcoded based on higher expression in the respective cell type (reddisch= higher expressed in CTX, green= higher expressed in CB)..*

Symbol	logFC	PValue	FDR	CB	CTX	Description
Emx2	-7.2	2.13E-06	0.00014368	0.09	14.66	Transcription factor
Klri1	-6.9	8.58E-05	0.00140723	0.04	6.63	Lectin-like receptor
D030046No 8Rik	-5.8	5.86E-05	0.00111289	0.09	4.61	Unknown
Ddn	-5.7	0.00011267	0.00168121	0.07	3.83	Involved in synaptic plasticity
Gm43517	-3.9	0.00011035	0.00166135	0.09	1.25	Unknown
Pcp2	6.3	0.00032526	0.00347784	2.77	0.03	Modulator for G protein-mediated cell signaling
Irx3	6.0	0.00062992	0.0056315	4.15	0.06	Transcription factor
Sfrp2	5.7	2.20E-05	0.00057692	5.10	0.10	Modulators of wnt signaling
Mstn	4.8	0.00313913	0.01768676	1.00	0.03	Secreted ligand of the tgf-beta receptor
Tmem30b	4.8	0.0004155	0.00412086	2.39	0.09	Membrane organization
Prokr2	4.4	5.57E-05	0.00107196	2.03	0.10	Receptor for prokineticin 2
Clic6	4.3	8.13E-05	0.00136016	1.16	0.06	Chloride intracellular channel
Gm26794	3.6	0.00044645	0.00436261	1.09	0.09	Unknown

Table 13 displays the top 20 differentially expressed genes in CTX and CB astrocytes and represents the molecular basis by which cortical and cerebellar astrocytes differ. The transcription factor genes *Foxg1*, *Lhx2* and *Emx2* were among the genes with highest fold change in expression in CTX. In CB astrocytes we found *Dao*, encoding D-amino acid oxidase, *Slco4a1,* a thyroid hormone transporter, and *Cnpy1*, the FGF signaling regulator 1, and the transcription factor genes *Irx5* and *Zic4* as genes with the highest fold change in expression.

Table 13 Top 20 differentially expressed genes ordered based on the absolute expression fold change between CTX and CB astrocytes*. LogFC=log2(foldchange). Gene symbols are colorcoded based on higher expression in the respective cell type (reddisch= higher expressed in CTX, green= higher expressed in CB).*

Symbol	logFC	abs. logFC	PValue	FDR	CB	CTX	Description
Dao	9.18	9.18	4.08E-06	0.00020828	115.63	0.20	D-amino acid oxidase
Foxg1	-8.54	8.54	2.54E-07	5.82E-05	0.14	55.60	Transcription factor
Lhx2	-8.17	8.17	2.41E-07	5.82E-05	0.19	50.52	Transcription factor
Slco4a1	8.04	8.04	1.13E-08	3.30E-05	73.07	0.28	Thyroid hormones transporter
Cnpy1	7.85	7.85	4.82E-07	7.53E-05	51.71	0.22	FGF signaling
Slc38a1	7.71	7.71	4.58E-07	7.43E-05	145.15	0.69	Transporter of glutamine
Gria1	7.66	7.66	2.89E-07	6.39E-05	274.65	1.36	Glutamate ionotropic receptor AMPA subunit
Opalin	-7.54	7.54	5.82E-08	3.30E-05	0.84	156.12	Oligodendrocytic myelin protein
Gdf10	7.36	7.36	1.29E-08	3.30E-05	1609.17	9.80	TGF beta receptor ligand
Syt10	7.26	7.26	4.48E-06	0.00021871	23.76	0.15	Ca^{2+} sensor
Wif1	7.22	7.22	4.79E-05	0.00097328	40.56	0.27	Inhibits WNT proteins
Emx2	-7.17	7.17	2.13E-06	0.00014368	0.09	14.66	Transcription factor
Hapln1	-7.09	7.09	2.06E-08	3.30E-05	0.70	94.57	Hyaluronan and proteoglycan link protein
Irx5	7.07	7.07	1.45E-07	5.08E-05	27.23	0.20	Transcription factor
Mybpc1	7.05	7.05	1.71E-08	3.30E-05	75.95	0.57	Filament-associated protein
Gm27239	-7.00	7.00	0.00038471	0.0039059	0.12	11.16	Unknown
A2m	6.96	6.96	2.14E-08	3.30E-05	238.32	1.91	Protease inhibitor and cytokine transporter
Abi3bp	6.94	6.94	2.33E-07	5.82E-05	46.16	0.37	Unknown
Klri1	-6.92	6.92	8.58E-05	0.00140723	0.04	6.63	Lectin-like receptor
Zic4	6.89	6.89	9.06E-07	0.00010029	24.92	0.21	Transcription factor

In summary, mRNA-Seq data demonstrated that cortical and cerebellar astrocytes are molecularly distinct cell populations. These populations share a broad spectrum of common functions. However, differential analysis demonstrated that genes responsible for these functionalities are distinct in both astrocytes.

3.1.5 Astrocytes from cerebellum and cortex differ in local DNA methylation

The first comparative analyses indicated commonalities between CTX and CB astrocytes with respect to local DNA methylation of cell type-specific genes (**Figure 3-5D**, **Figure 3-5E**) and genome-wide DNA methylation (PC1 in **Figure 3-5A**). Still, the PCA on DNA methylation revealed a clear separation of CTX and CB astrocytes in the second principal component (PC2 in **Figure 3-5A**). The methylation state of the CpGs contributing to this separation was closer to neurons than to CTX astrocytes, while $NeuN^-$ cells and oligodendrocytes were equidistant to CB and CTX astrocytes. These

CpGs were associated with genes linked to neuronal tube development (pValue $<1^{-10}$). A closer inspection of global DNA methylation revealed a comparable high methylation (median methylation > 90%) in both astrocyte populations (**Figure 3-13A**, **Figure 3-13B**). However, similar to the PCA in **Figure 3-5A**, hierarchical clustering of the methylomes pointed to detectable differences between these cells. Moreover, a genome wide DNA methylation segmentation analysis (see chapter 2.2.10 on page 53), classifying the genome into highly and partially methylated domains (see chapter 1.1.2 on page 16), separated CTX and CB showing that both astrocyte populations slightly differ in the genome wide methylome organization (**Figure 3-13C**).

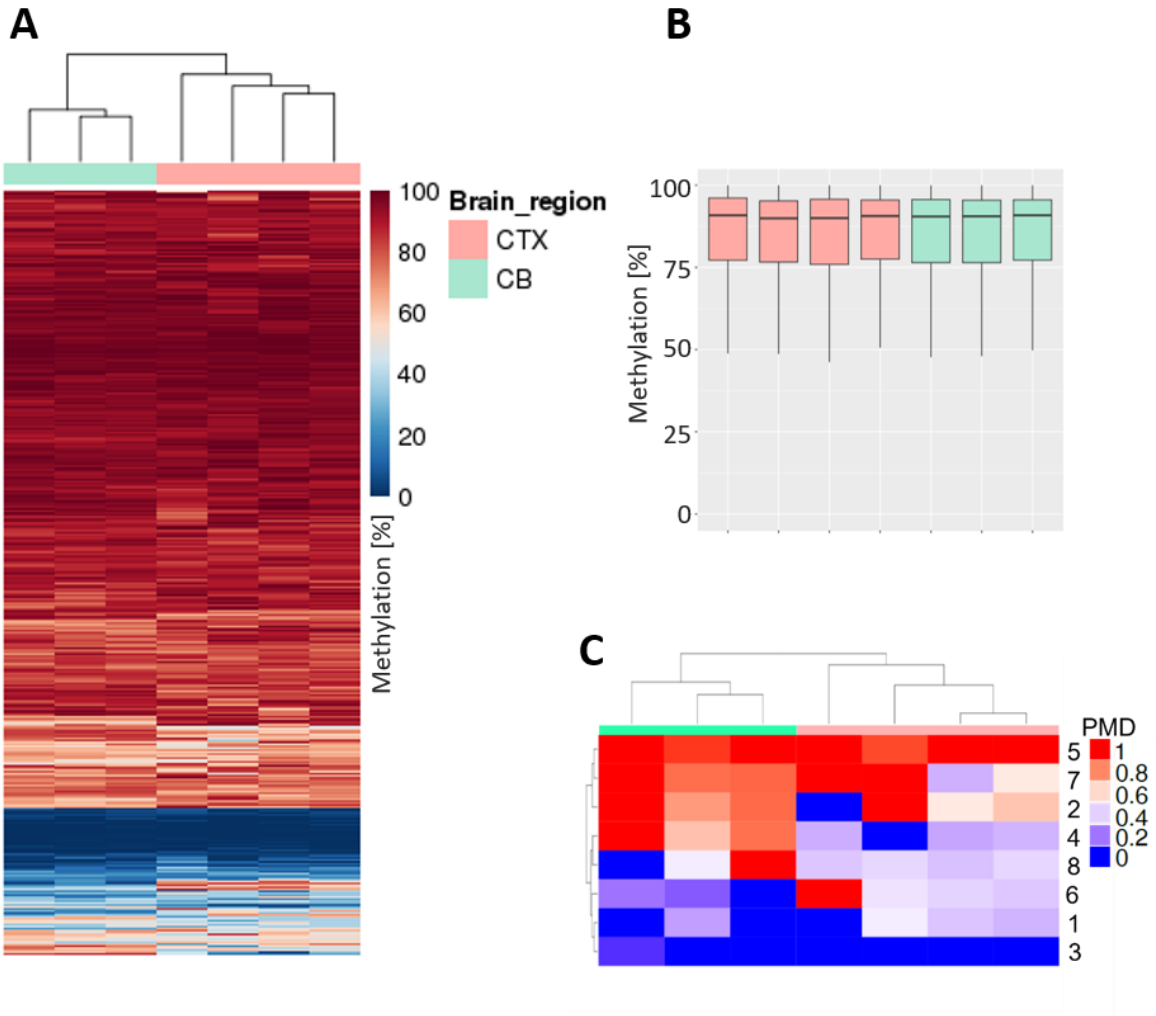


Figure 3-13 Global DNA methylation. *A. Heatmap representing DNA methylation of the genome binned in 1kb tiles. B. Distribution of DNA methylation values of all detected CpGs (coverage ≥10x) in each sample. C. Hierarchical clustering and emission probabilities for partially methylated domains calculated by ChromH3M.*

In order to investigate the DNA methylation differences in more detail, differential analysis using the R package methylKit was performed (**Figure 3-14A**; see also chapter 2.2.10 on page 53 for more details). This analysis resulted in a total of 5,363 differentially methylated regions (DMRs) with 3,052 hypomethylated regions in CB and 2,311 hypomethylated regions in CTX (**Figure 3-14B**). More than 90% of the identified DMRs were located in intronic and intergenic regions, most likely demarcating proximal and distal gene regulatory elements (**Figure 3-14C**). Of note, a high proportion of genes associated to DMRs coded for transcription factors and nucleic acid binding proteins followed by genes encoding hydrolases, receptors, transporters and cytoskeletal proteins (**Figure 3-14D**). The main biological processes associated with DMRs were related to developmental processes (**Table 14**).

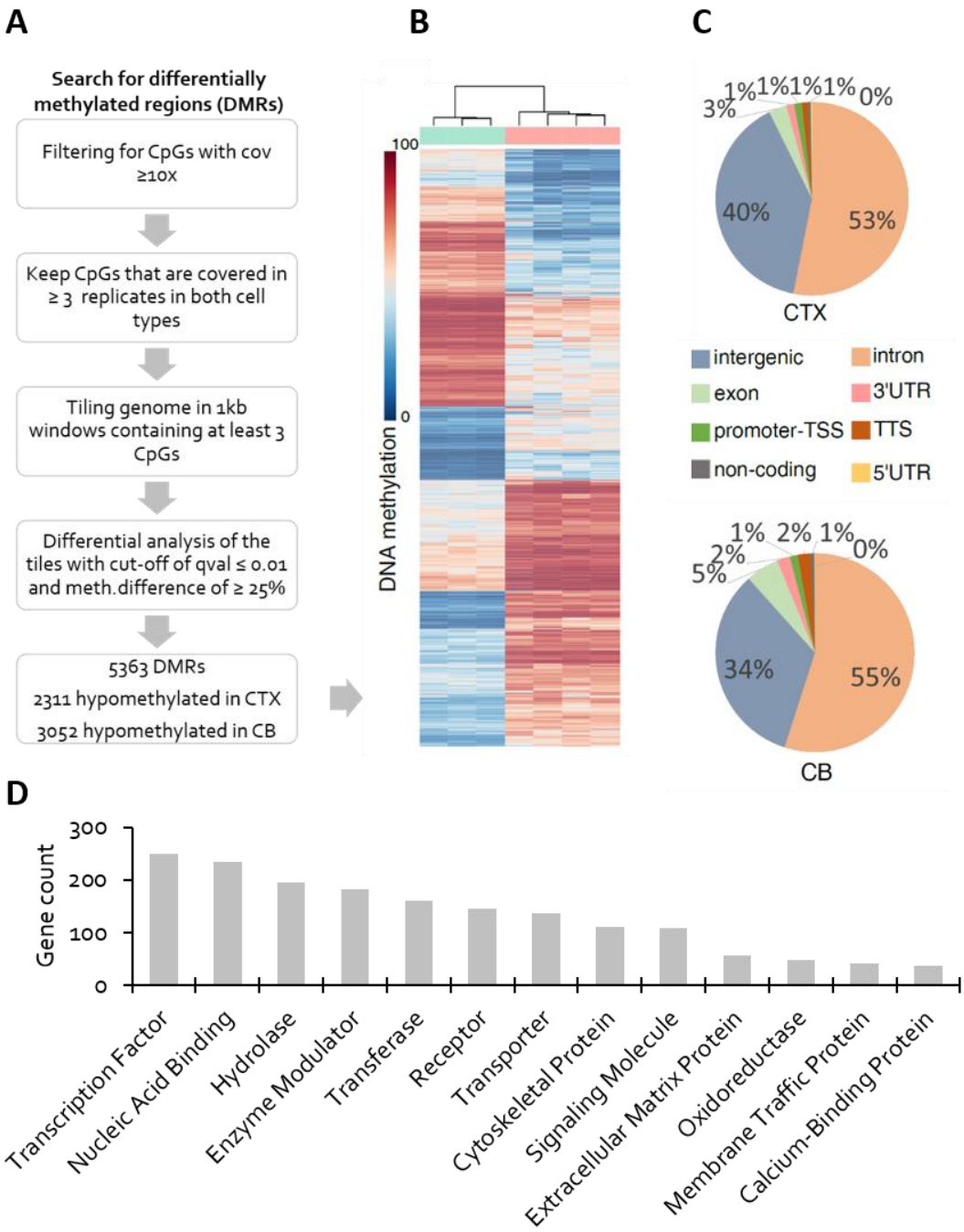


Figure 3-14 Differentially methylated regions (DMRs). A*. Strategy to identify and define DMRs.* ***B****. Heatmap representing the 5,363 DMRs discriminating CB astrocytes from CTX astrocytes.* ***C****. Genomic distribution of the hypomethylated regions in CTX (top) and CB astrocytes (bottom).* ***D****. Classification of DMR-associated genes into protein classes. Cov= Coverage; TSS= Transcription start site; TTS= Transcription termination site; UTR= Untranslated region.*

Table 14 GO Term enrichment for DMRs between CTX and CB astrocytes.

GO Biological process Term Name	P-Value	Observed Gene Hits
Respiratory system development	2.8E-31	95
Lung development	1.6E-26	84
Respiratory tube development	2.9E-26	86
Negative regulation of cellular response to growth factor stimulus	4.2E-25	33
Regulation of cellular response to growth factor stimulus	6.4E-24	59
Somatic stem cell maintenance	1.4E-23	26
Stem cell maintenance	8.7E-23	48
Mammary gland development	4.6E-20	58
Embryonic cranial skeleton morphogenesis	1.5E-19	20
Mesoderm development	9.2E-19	47

3.1.6 Astrocytes from cerebellum and cortex differ in global and local DNA accessibility.

Similar to the global DNA methylation, the first comparative analyses indicated commonalities between CTX and CB astrocytes with respect to local DNA accessibility of cell type-specific genes (**Figure 3-5D**, **Figure 3-5E**) and genome-wide DNA accessibility (PC1 in **Figure 3-5B**), while the second principal component revealed a clear separation of accessibility of CTX and CB astrocytes (PC2 in **Figure 3-5B**). In line with this, unsupervised hierarchical clustering on the read counts in 100bp windows across the genome discriminated between CB and CTX astrocytes (**Figure 3-15A**). In order to investigate the DNA accessibility differences in more detail, differential analysis using the R package csaw was performed (see chapter 2.2.9 on page 51 for more details). This analysis resulted in a total of 17,097 differentially accessible regions (DARs) between CTX and CB astrocytes with 15,866 regions more open in CTX and 2,311 more open regions in CB (**Figure 3-15B**). In addition to the strikingly high number of DARs, the peak height and width of open chromatin sites were more pronounced in CTX (**Figure 3-15B,** lower panel). More than 90% of the identified DARs were located in intronic and intergenic regions, again indicating epigenetic differences in proximal and distal gene regulatory elements (**Figure 3-15C**). Interestingly, the classification of the DARs into protein classes showed a very similar picture as for DMRs. A high proportion of genes associated to DARs coded for nucleic acid binding proteins, followed by genes encoding hydrolases, transcription factors, receptors, transporters and cytoskeletal proteins (**Figure 3-15D**). The biological processes associated with more open regions in CTX were related to various processes such as somatic stem cell maintenance, cell adhesion and migration, and Notch signaling (**Table 15, Table 14**). More open regions in CB were associated with genes related to early brain development (**Table 16**).

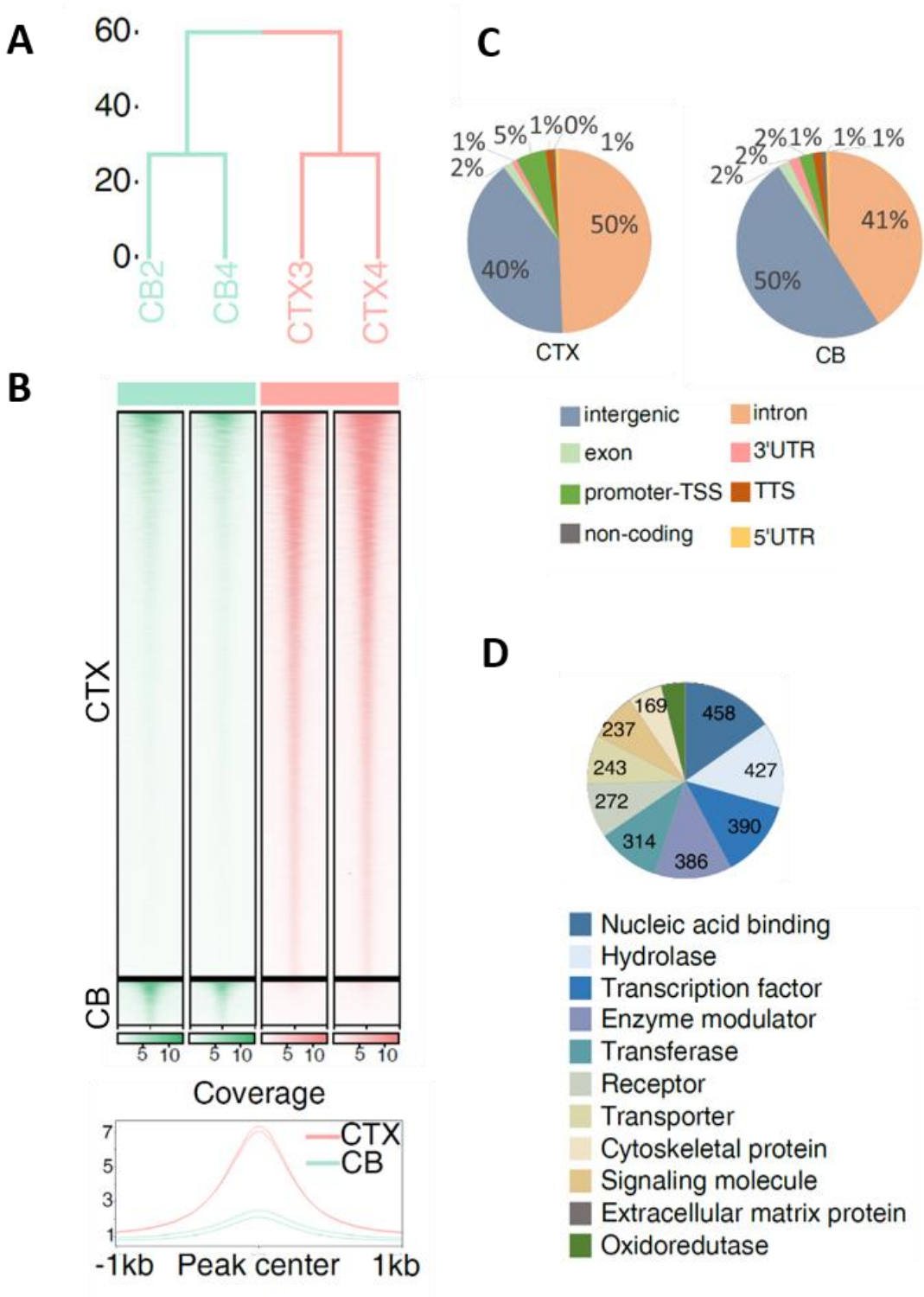


Figure 3-15 Differential DNA accessibility in CB and CTX astrocytes. ***A****. Unsupervised hierarchical clustering of global DNA accessibility.* ***B****. Heatmap representing the accessibility around the center of differentially accessible regions (DARs) ±1kb (15,866 regions more open in CTX, and 1,231 regions more open in CB). Profile of the coverage across DARs (±1 kb) is represented beneath the DARs heatmap.* ***C****. Genomic annotation of more open regions in the respective astrocyte type.* ***D****. Classification of DMR-associated genes into protein classes. TSS= Transcription start site; TTS= Transcription termination site; UTR= Untranslated region.*

Table 15 GO Term enrichment for more accessible regions in CTX

GO Biological Processes Term Name	P-Value	Observed Gene Hits
Somatic stem cell maintenance	7.02E-41	32
Regulation of Notch signaling pathway	2.12E-26	30
Regulation of generation of precursor metabolites and energy	1.38E-21	32
Filopodium assembly	8.68E-19	13
Cerebral cortex radially oriented cell migration	4.03E-17	20
Homotypic cell-cell adhesion	4.72E-17	20
Negative regulation of osteoblast differentiation	6.56E-16	28
Cerebral cortex radial glia guided migration	3.22E-15	13
Regulation of oligodendrocyte differentiation	7.84E-15	18
Cerebellum morphogenesis	1.04E-13	22
Negative regulation of gliogenesis	1.34E-13	21
Ventral spinal cord interneuron differentiation	1.13E-12	12
Dentate gyrus development	1.55E-12	12
Positive regulation of mrna catabolic process	2.82E-12	13
Regulation of nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay	3.43E-12	11
Negative regulation of cell-cell adhesion	1.34E-11	20
Positive regulation of Notch signaling pathway	1.41E-11	15
Positive regulation of nuclear-transcribed mRNA catabolic process, deadenylation-dependent decay	1.44E-11	10
Lamellipodium assembly	2.45E-11	17
Fatty acid elongation	8.16E-11	8

Table 16 GO Term enrichment for more accessible regions in CB

GO Biological Processes Term Name	P-Value	Observed Gene Hits
Morphogenesis of embryonic epithelium	6.10E-10	29
Mammary gland specification	1.65E-08	3
Epithelial tube formation	2.47E-08	25
Embryonic epithelial tube formation	4.01E-08	24
Tube formation	1.51E-07	26
Regulation of binding	1.82E-07	27
Chondrocyte development	2.52E-07	7
Craniofacial suture morphogenesis	2.69E-07	5
Primary neural tube formation	3.69E-07	15
Neural tube development	3.87E-07	26
Spinal cord development	5.69E-07	15
Bone development	5.81E-07	25
Negative regulation of DNA binding	2.20E-06	9
Neural crest cell migration	4.12E-06	10
Regulation of smoothened signaling pathway	6.47E-06	12
Neural tube formation	1.23E-05	17
Cell differentiation in spinal cord	1.37E-05	12
Regulation of protein processing	1.79E-05	9
Specification of nephron tubule identity	1.85E-05	4
Pattern specification involved in metanephros development	2.28E-05	3

The strong difference in the openness indicated a major difference in chromatin organization. In line with this, by inspecting gene expression data, an enhanced expression of genes involved in chromatin modification (*Hat1*, *Prmt2*, *Ino80c*), telomere control and chromatin compaction (*Hgmb1*, *Hmgn1*), and histone chaperones (*Anp32e, Npm1, Nap1l1, Asf1a)* and all major members of the cohesion complex (*Smc3, Stag1, Rad21* and *Ctcf*) was observed in CB astrocytes. On the other hand, histone modifiers such as *Hdac7* and *Hdac11*, *Kmt5a* and *Kdm5b* were higher expressed in CTX astrocytes (**Table 17**). Moreover, an enhanced expression of core histone variants H2a, H2b, H3.3 and H4, the linker histone H1, together with replication-independent histone variants was observed in the cerebellum. Evaluation of expression values in recently published single cell and bulk transcriptome data confirmed these differences (Appendix **Figure 4-2**, **Figure 4-4**, **Figure 4-5**) (Boisvert et al. 2018; Zeisel et al. 2018). Immunohistochemistry using antibodies against H3K27me3 and H3K27ac indeed revealed a significant difference between CB and CTX (**Figure 3-16**). These findings indicated that the chromatin of CB astrocytes may comprise an overall higher nucleosome density which affects the general DNA accessibility.

Intriguingly, the 12-fold increase in open chromatin sites was not associated with general increase of gene expression (**Figure 3-7**, **Figure 3-11**). Furthermore, the number of DMRs was rather balanced between CB and CTX astrocytes and genes associated with DNA methylation dynamics such as *Dnmts*, *Tets* or *Mecp2* were not differentially expressed (**Table 17**).

In conclusion, cortical and cerebellar astrocytes displayed an overall distinct chromatin organization that did not entirely correlate to gene expression changes or DNA methylation changes.

Table 17 Gene expression of genes involved in chromatin organization and DNA methylation dynamics.

Gene	logFC	PValue	FDR	CB	CTX
Hopx	4.72	3E-06	2E-04	2376.38	90.04
Hist1h4c	2.67	4E-03	2E-02	7.18	1.13
Hist1h4i	2.40	1E-05	5E-04	148.87	28.25
H1fx	2.27	3E-05	8E-04	17.56	3.63
Hist1h4k	2.21	4E-04	4E-03	33.91	7.32
Hist1h2ac	1.86	4E-04	4E-03	6.36	1.76
Hist1h4h	1.59	4E-03	2E-02	36.65	12.19
H2afv	1.49	6E-05	1E-03	197.00	70.16
Kdm5d	1.48	5E-04	5E-03	6.32	2.27
Kansl1l	1.45	1E-04	2E-03	10.81	3.96
Prmt2	1.38	7E-04	6E-03	24.58	9.48
Ino80c	1.36	3E-04	3E-03	23.91	9.31
Phf19	1.30	3E-03	2E-02	1.92	0.78
Ppp1r1b	1.21	4E-03	2E-02	21.19	9.15
Hist1h2bm	1.18	2E-03	1E-02	30.90	13.67
Lrrk2	1.12	3E-04	3E-03	4.93	2.26
Tspyl5	1.08	6E-03	3E-02	4.05	1.91
Pcgf5	1.07	2E-03	1E-02	16.18	7.69
Elof1	1.05	7E-04	6E-03	56.98	27.53
Hmgn1	1.01	2E-03	1E-02	97.19	48.39
Lpin1	0.86	6E-04	6E-03	15.99	8.81
Hmgb1	0.84	3E-03	2E-02	28.33	15.78
Hist1h2al	0.83	1E-02	4E-02	141.52	79.37
Rad21	0.82	4E-04	4E-03	65.92	37.32
L3mbtl3	0.82	1E-03	9E-03	7.78	4.42
Cbx3	0.82	2E-03	1E-02	17.12	9.73
Hist1h1c	0.80	4E-03	2E-02	69.06	39.59
Hat1	0.79	2E-03	1E-02	18.65	10.81
Dek	0.79	1E-03	8E-03	50.16	29.08
Pih1d1	0.78	1E-02	5E-02	19.93	11.61
Ube2e1	0.77	5E-03	3E-02	13.29	7.78
Pole3	0.72	4E-03	2E-02	19.03	11.59
Cdan1	0.69	4E-03	2E-02	3.92	2.43
Ube2n	0.68	2E-03	1E-02	33.62	21.03
Smc3	0.68	2E-03	1E-02	18.27	11.43
Nap1l1	0.64	3E-03	2E-02	38.30	24.50
Phf20	0.63	1E-02	5E-02	8.22	5.31
Spin1	0.61	3E-03	2E-02	19.23	12.58
Wdr61	0.60	4E-03	2E-02	23.92	15.75
Babam1	0.57	7E-03	3E-02	35.19	23.72
Nucks1	0.55	8E-03	3E-02	21.57	14.77
Mgea5	0.54	6E-03	3E-02	32.38	22.28
Prkca	0.54	1E-02	4E-02	5.88	4.05
Iws1	0.52	1E-02	4E-02	3.90	2.72
Smad4	0.50	1E-02	5E-02	31.01	21.86
Skp1a	0.46	9E-03	4E-02	184.33	134.44
Asf1a	0.45	2E-02	7E-02	38.58	28.18
Stag1	0.43	3E-02	9E-02	5.86	4.35
Npm1	0.43	1E-02	5E-02	106.51	79.23
Anp32e	0.42	3E-02	9E-02	22.79	17.08
Ctcf	0.41	3E-02	8E-02	16.91	12.68

Gene	logFC	PValue	FDR	CB	CTX
Sirt7	-0.48	1E-02	5E-02	8.31	11.62
Usp22	-0.54	1E-02	4E-02	16.27	23.67
Kdm5b	-0.56	1E-02	5E-02	3.18	4.69
Hltf	-0.58	1E-02	5E-02	3.14	4.70
Smyd2	-0.66	1E-02	4E-02	6.42	10.09
Usp16	-0.66	6E-03	3E-02	5.04	7.98
Zmynd8	-0.75	6E-03	3E-02	3.23	5.42
Paxbp1	-0.77	9E-03	4E-02	8.98	15.29
Hmgn3	-0.77	1E-02	4E-02	158.10	270.39
Rrp8	-0.85	3E-03	2E-02	5.35	9.62
Hdac7	-0.93	9E-03	4E-02	2.14	4.08
Cdyl	-0.95	4E-03	2E-02	3.30	6.34
Apbb1	-1.01	1E-03	1E-02	14.94	30.08
Prkd1	-1.09	6E-04	5E-03	8.43	17.89
Rcbtb1	-1.11	2E-04	2E-03	5.67	12.23
H2afj	-1.13	2E-04	3E-03	26.27	57.35
Per1	-1.20	3E-03	2E-02	1.33	3.07
Slk	-1.28	1E-04	2E-03	4.18	10.15
Zbtb7b	-1.29	7E-04	6E-03	1.38	3.38
Padi2	-1.30	8E-04	7E-03	2.16	5.31
Hdac11	-1.36	1E-03	9E-03	2.98	7.67
Sirt2	-1.41	1E-05	4E-04	59.50	157.80
Kmt5a	-1.42	5E-05	1E-03	5.15	13.77
Auts2	-1.63	1E-04	2E-03	1.29	4.01
Hr	-1.66	2E-04	2E-03	1.09	3.45
Chd3	-1.69	7E-04	6E-03	1.65	5.29
Dpf3	-1.72	1E-03	1E-02	1.34	4.36
Vegfa	-1.72	2E-04	3E-03	68.65	225.75
Brca1	-1.85	7E-04	6E-03	0.68	2.42
Snca	-1.96	7E-03	3E-02	2.84	11.01
Hist3h2ba	-2.18	2E-04	2E-03	7.61	34.44
Zfpm1	-2.26	2E-03	1E-02	0.37	1.73
Eya1	-2.30	2E-06	1E-04	3.56	17.51
Pim3	-2.37	5E-04	5E-03	3.50	17.90
Eya2	-2.38	3E-03	2E-02	0.39	1.99
Lef1	-2.48	6E-04	5E-03	0.25	1.39
Prkcd	-3.01	5E-07	8E-05	1.18	9.57
Smyd1	-3.09	3E-04	3E-03	0.18	1.50
Sall1	-5.58	6E-06	3E-04	0.37	16.87
Tet1	-0.33	0.11	0.24	1.61	2.02
Tet2	-0.06	0.79	0.87	4.06	4.24
Tet3	0.13	0.45	0.62	2.81	2.57
Mecp2	-0.07	0.67	0.79	4.94	5.18
Dnmt1	0.07	0.73	0.83	2.68	2.55
Dnmt3a	-0.47	0.03	0.08	2.76	3.83
Dnmt3b	-0.07	0.88	0.93	0.66	0.69

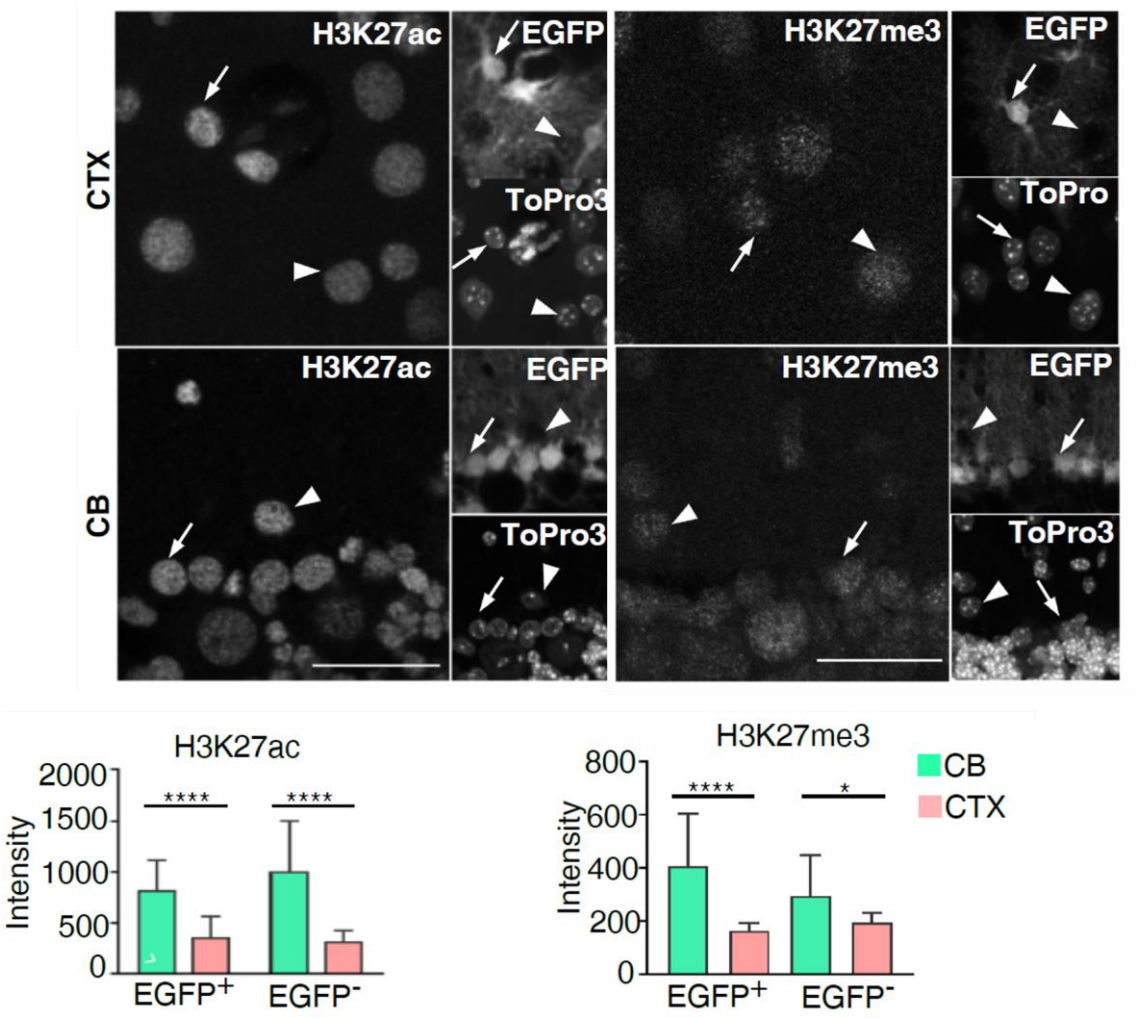


Figure 3-16 Immunofluorescence and signal quantification of H3K27ac (left) and H3K27me3 (right) in CTX and CB of hGFAP-EGFP mice. *Arrows indicate EGFP⁺ cells, arrow heads indicate EGFP⁻ cells. Number of analyzed cells (CB EGFP⁺/CTX EGFP⁺/CB EGFP⁻/CTX EGFP⁻) n= 33/16/15/17 (H3K27ac), n=26/13/20/28 (H3K27me3). Data represented as mean + STEDV. Mann-Whitney-U-test, * p=0.01-0.05, ** p=0.001-0.01, *** p=0.0001-0.001, ****p< 0.00001. Scale bar=20µm. Immunofluorescence staining and quantification were done by Dr. Laura Stopper in terms of a cooperation*

3.1.7 Correlation between epigenetics and gene expression reveals positive and negative regulation of cell type-specific expression

Next, the relationship between local epigenomic signatures in cross correlations to open chromatin was tested to identify loci that potentially regulate gene expression (**Figure 3-17**). Loci with a higher chromatin accessibility tend to exhibit an overall lower DNA methylation in the respective astrocyte groups (**Figure 3-17A** top). Moreover, the expression of genes associated to the respective more open region was higher (**Figure 3-17A** bottom). Conversely, inspection of hypomethylated regions in CB or CTX astrocytes showed a higher median DNA accessibility in the respective cell population in addition to a slightly higher expression of the associated genes (**Figure 3-17B**). This relation indicated that a cell type-specific openness and hypomethylation at regulatory regions, such as e.g. potential enhancers, is accompanied by a cell type-specific expression.

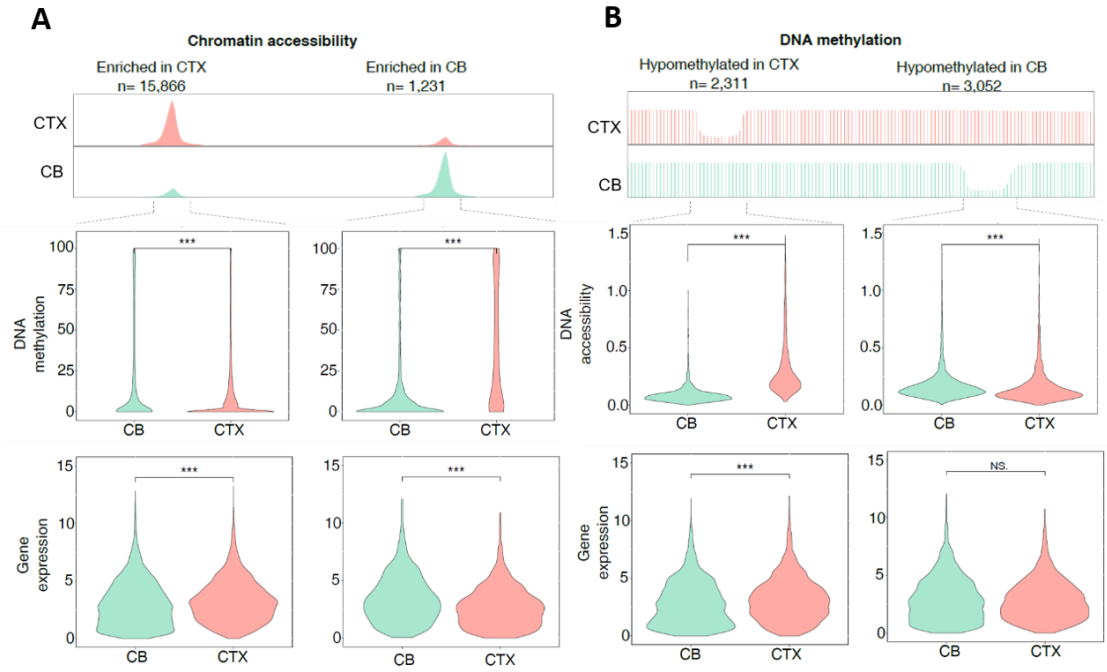


Figure 3-17 Integrating epigenetics and gene expression indicates regulation of cell type-specific expression. A*. Distribution of DNA methylation levels at DARs (top), and gene expression of DAR-associated genes (bottom). On the left for more open regions in CTX astrocytes and on the right for more open regions in CB astrocytes.* ***B****. Distribution of DNA accessibility at DMRs (top), and the gene expression of DMR-associated genes (bottom). On the left for hypomethylated regions in CTX astrocytes and on the right for hypomethylated regions in CB astrocytes.*

The previous analysis showed a general effect of epigenetic differences on global gene expression without accounting for gene-based resolution or statistical significance. To better understand the relation between open chromatin (DARs), DNA methylation status (DMRs) and expression changes (DEGs) further comparative tests were performed. Relating significant DNA methylation changes to significant gene expression changes results in 467 lower methylated regions linked to higher expression of the associated genes in CTX and 529 of lower methylated regions linked to higher expression of the associated genes in CB (**Figure 3-18A**, red boxes). The link between higher chromatin accessibility and higher gene expression in CTX (3,647 genes) is 10-fold higher than in CB (348 genes) (**Figure 3-18B**, red boxes). Of note, this comparison revealed a number of cases in which gene expression was associated to DMRs and DARs in a non-canonical correlation (**Figure 3-18A**, **Figure 3-18B**, blue boxes). 328 DMRs, showing higher methylation in CTX astrocytes, were associated with higher gene

expression in CTX, while 145 DMRs followed this correlation in CB astrocytes (**Figure 3-18A**, blue boxes). In CTX, decreased DNA accessibility was associated with genes showing higher gene expression in CTX in the case of 1,071 DARs (**Figure 3-18B**, blue boxes) and only 59 genes followed this association in CB. A third cross comparison between methylation and open chromatin showed that 383 hypomethylated regions overlapped with increased accessibility in CTX, and only 76 hypomethylated regions in CB followed this trend (**Figure 3-18C**).

Next, we performed GO term enrichment analysis with the genes shown in each quadrant of **Figure 3-18** (**Figure 3-18A-C**). Gene sets associated with lower methylation and increased gene expression in the cortex were enriched for processes involved in somatic stem cell maintenance (Q3 in **Figure 3-18A**), while higher gene expression linked with higher DNA accessibility was predominantly associated with gene sets related to development, and cell migration (Q3 in **Figure 3-18B**).

In CB astrocytes lower DNA methylation coupled to higher gene expression was associated with genes regulating peptidyl phosphorylation and cytoskeleton organization (Q2 in **Figure 3-18A**) and more open chromatin coupled to higher expression was found to be most strongly associated with morphogenesis (Q2 in **Figure 3-18B**).

In summary, these cross comparisons revealed positive and negative correlations between DMRs, DARs and gene expression. This observation suggested that epigenetically marked regions are acting as potential positive or negative gene regulators for cell type-specific expression in CB and CTX astrocytes.

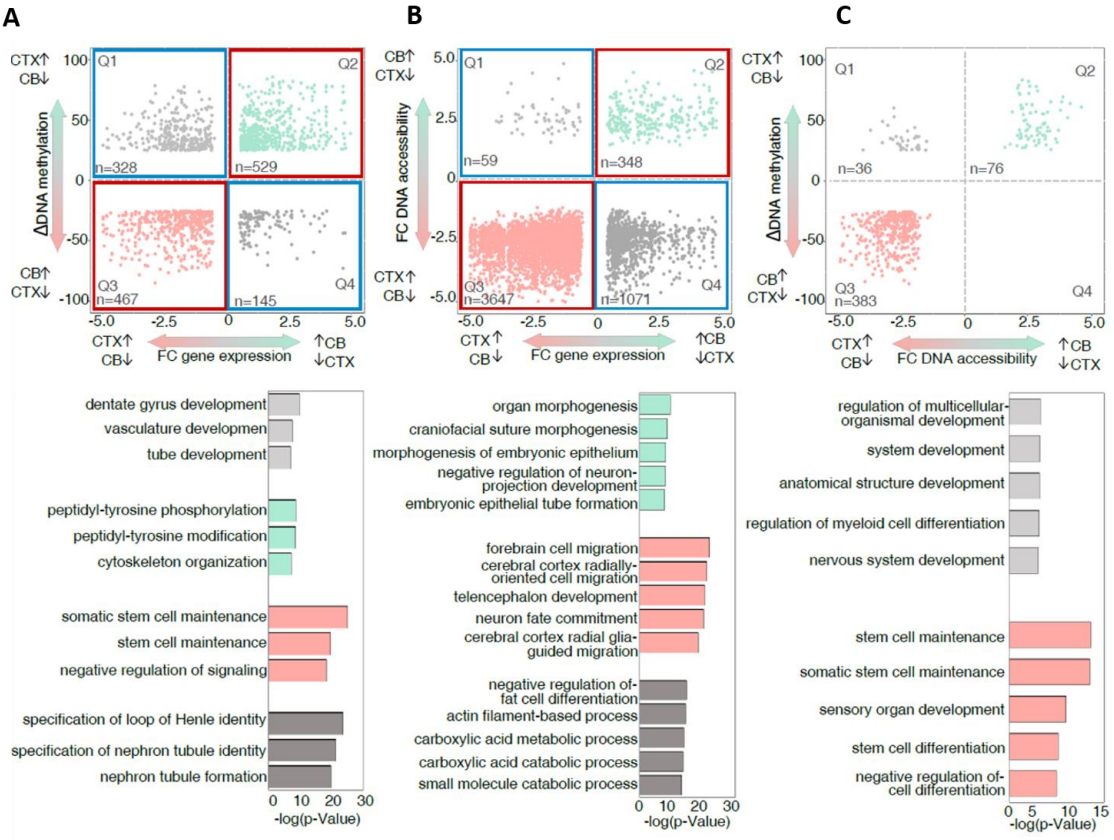


Figure 3-18 Integrating epigenetics and gene expression reveals positive and negative regulation of cell type-specific expression. ***A***. *Pairwise comparison of DNA methylation difference and gene expression difference.* ***B***. *Pairwise comparison of DNA accessibility difference and gene expression difference.* ***C***. *Pairwise comparison of DNA methylation difference and DNA accessibility difference.* ***A-C***. *Bar diagrams below respective scatterplots show enrichment of biological processes for the respective genes in Q1-Q4. No GO term enrichment for Q1 in B, and Q2 in C was detected. Positive expression FCs represent significantly higher expression or higher DNA accessibility in CB astrocytes. Negative expression FCs represent significantly higher expression or higher DNA accessibility in CTX astrocytes. Positive methylation difference represents significantly lower methylation in CB. Negative methylation difference represents significantly lower methylation in CTX astrocytes. ***pValue<0.001, NS=not significant (Wilcoxon Test). FC=fold change. Data represented as mean.*

3.1.8 Epigenetic differences indicate regional transcription factor control

As shown in **Figure 3-14** and **Figure 3-15**, more than 90% of the DMRs and DARs identified were located in intronic and intergenic regions, most likely demarcating proximal and distal gene regulatory elements. Moreover, the majority of DMR- and DAR-associated genes were classified into transcription factors (TFs) and DNA binding proteins. Together with the indicated positive and negative regulation of gene expression in **Figure 3-18** it was important to resolve potential regulators. To this end, TF binding motif analyses exploring DARs and DMRs sequences was performed. More open regions in CTX were particularly enriched for binding sites of the TFs Emx2 and Lhx2. Both TFs are known from tissue studies to be essential for the cerebral cortex regionalization and development (Bulchand et al. 2001; Mallamaci et al. 2000, 2) (**Figure 3-19A**). Of note, as reported in chapter 3.1.4, both were among the top 20 differentially expressed genes (**Table 13**; **Figure 3-19E**). Published single cell expression data confirmed that *Lhx2* is expressed in a large number of cortical astrocytes (Appendix **Figure 4-6**). In addition, Lhx2/Emx2 motif -containing open regions were much less methylated in CTX astrocytes (**Figure 3-19B**). *Lhx2* expression was verified by immunohistochemistry showing a higher expression in CTX compared to CB astrocytes (**Figure 3-19F**). The consistency on all four levels (enrichment of binding sites at open regions, hypomethylation, higher mRNA expression and higher protein expression) suggested an important role of Lhx2 in cortical astrocytes. Motif enrichment of more open region in CB astrocytes revealed an overrepresentation of Zic family binding sites (**Figure 3-19A**). Interestingly, Zic binding at enhancers was shown to be associated with the development of granule cells[7] to establish gene expression patterns of mature neurons (Frank et al. 2015). Along with Zic binding prediction, the expression of Zic family members was strongly enhanced in CB astrocytes (*Zic1* (63-fold), *Zic2* (4-fold) and *Zic4* (120-fold) ; **Figure 3-19E**) and, as reported in chapter 3.1.4, *Zic4* was in the top 20 most differentially expressed genes (**Table 13**). Published single cell expression data revealed that *Zic1* is expressed in nearly every Bergmann glia cell (Appendix **Figure 4-5**). Furthermore, the inspection of DNA methylation level of Zic-enriched sites showed hypomethylation of respective sites in CB astrocytes (**Figure**

[7] Neuron population in the cerebellum.

3-19B). Immunohistochemistry of Zic1 revealed a weak protein expression in Bergmann Glia along with strong signals in Purkinje cells and granule cells, but no signals in the cortical astrocytes (**Figure 3-19G**). The apparently prominent role of Zic family TFs in CB astrocytes was not reported before and remains to be clarified. The motif enrichment analysis also highlighted a strong enrichment of Nfi binding motifs in regions-specific enriched open regions (**Figure 3-19A**), confirming Nfi as an important astrocyte signature and suggesting the involvement of Nfi in region-specific astroglial programs.

For motif analyses of DMRs, regions with a methylation difference of at least 50% between CTX and CB astrocytes were subjected to the analysis, reasoning that this would restrict the results to major regulators of region-specific programs. 379 of these DMRs were hypomethylated in CTX and 573 hypomethylated in CB. Again, this analysis highlighted the strong enrichment of Lhx2 consensus motifs CTX astrocytes (**Figure 3-19C**) together with stronger enrichment of open chromatin at respective DMRs (**Figure 3-19D**). Hypomethylated regions of CB astrocytes were not significantly enriched for Zics but for binding motifs for Sox2/ Sox3/ FoxJ3 (**Figure 3-19C**). With the exception of Etv5, none of the found TFs in hypomethylated region in CB were differentially expressed (**Figure 3-19E**).

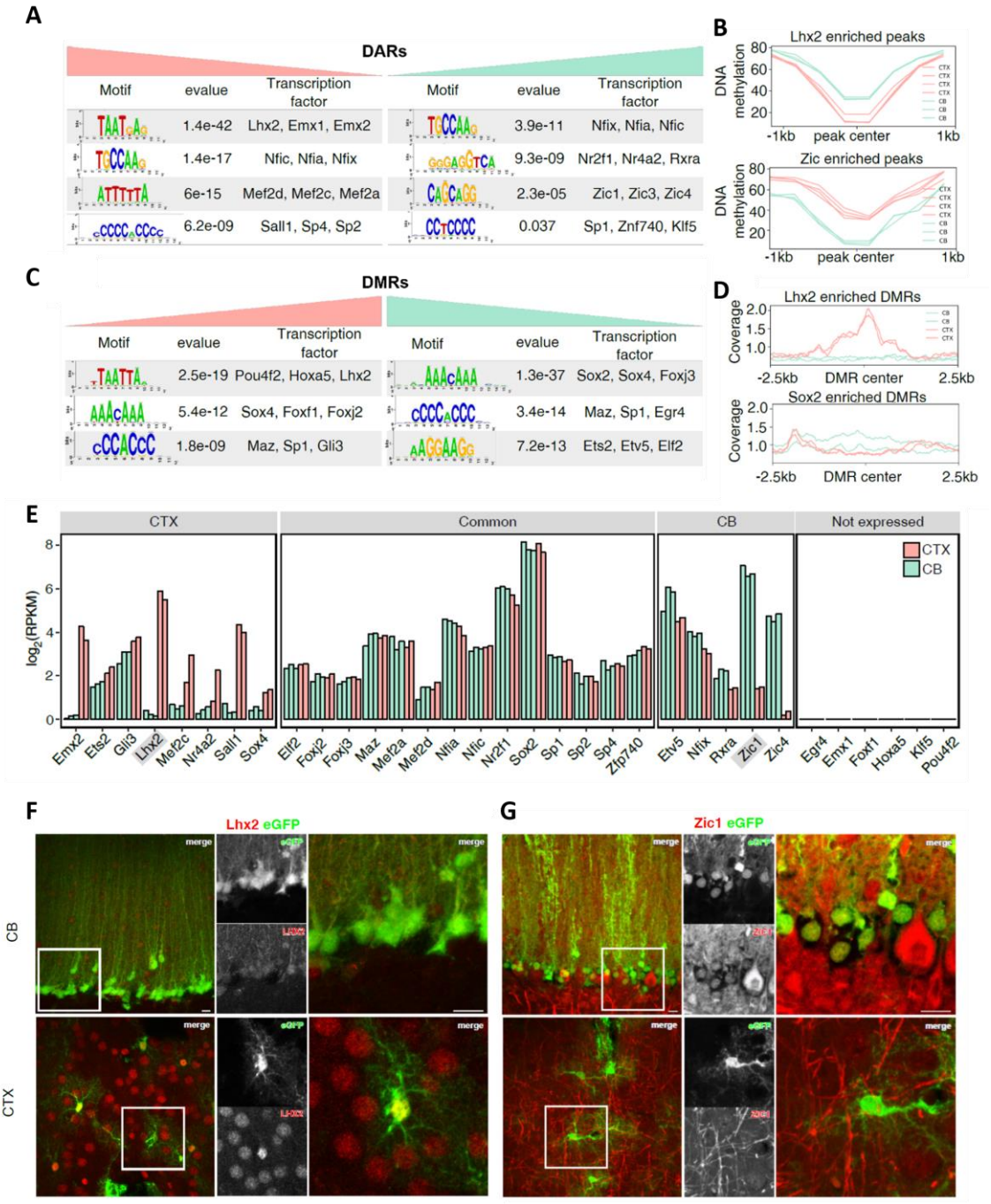
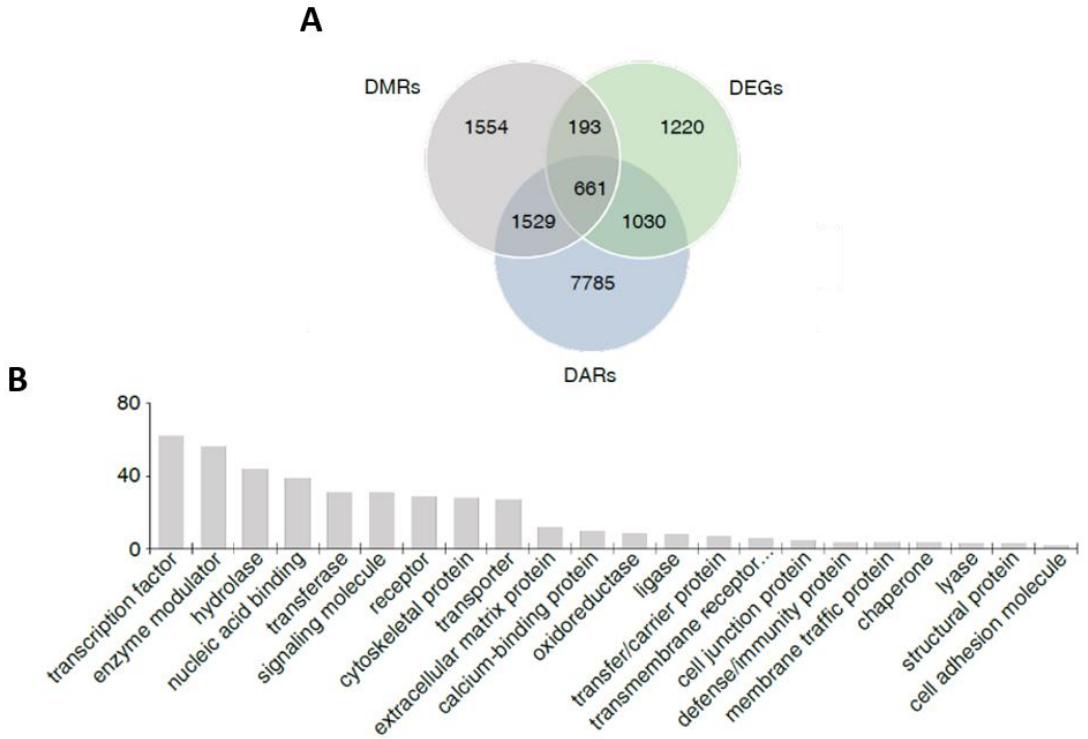


Figure 3-19 Epigenetic changes and transcription factor control. ***A****. Enriched transcription factor binding motifs in more accessible regions of CTX astrocytes (left) or CB astrocytes (right).* ***B****. DNA methylation around the center of open regions (±1 kb) enriched for Lhx2/Emx2 or Zic binding motifs.* ***C****. Enriched transcription factor binding motifs in the hypomethylated regions of CTX astrocytes (left) or CB astrocytes (right).* ***D****. DNA accessibility around the center of DMRs (±2.5 kb) enriched for Lhx2 or Sox2 binding motifs.* ***E****. Gene expression of transcription factors with enriched binding sites in DMRs or DARs grouped by higher expression in CTX astrocytes, comparable expression in both, higher expression in CB astrocytes, and not expressed.* ***F-G****. Immunofluorescence of Lhx2 and Zic1 in CTX and CB of hGFAP-EGFP mice. Scale bars= 20µm (left), 10µm (right). Immunofluorescence staining was done by Dr. Laura Stopper in terms of a cooperation*

3.1.9 Cortical and cerebellar astrocytes execute specific programs through complex epigenetic networks

To further investigate the interplay between DARs, DMRs, and astrocyte-specific expression programs, genes were analyzed that were linked to both DARs and DMRs (**Figure 3-20A**). 661 of such genes were identified and classified into protein classes. Consistently with the classification of DMRs and DARs alone, a high proportion of the intersected genes coded for transcription factors, nuclei acid binding proteins, hydrolases and receptors (**Figure 3-20B**). Functionally, nervous system development was identified as the most prominent common feature (**Figure 3-20C**). The fact, that this term was the highest enriched for both gene sets indicated that the development of these cell populations is differentially controlled by epigenetic mechanisms. *Lhx2*, *Emx2* together with their interaction partners *Otx1,* the antisense transcript *Emx2os*, *Nr2e1*, *Sfrp1, Fezf2* and *Sall1* were among the gene set related to nervous system development in the CTX. These genes were previously shown to play a crucial role in the early patterning and development of the forebrain (Bulchand et al. 2001; Kimura et al. 2005; Monaghan et al. 1997; Trevant et al. 2008; Harrison et al. 2012, 1) (Appendix **Figure 4-7**). Among the genes higher expressed and differentially regulated in CB astrocytes and related to nervous system development were midbrain-hindbrain specification factors such as *En1*, *En2*, *Irx1*, *Irx2*, *Irx3*, *Irx5*, *Zic1 and Zic4* (Cheng et al. 2010; Elsen et al. 2008; Lecaudey et al. 2005; Matsumoto et al. 2004) (Appendix **Figure 4-7**). As shown in **Figure 4-7**, these genes are highly expressed in the early embryonic brain in a region restricted fashion. However, the persistent expression of these genes in astrocytes has not been reported or investigated in detail so far. Of note, the evaluation of single cell expression data showed a heterogenous expression pattern in the respective astrocyte population (Appendix **Figure 4-5, Figure 4-6**), which might be due to a heterogeneous expression or technical reasons.

The functional analysis of the intersected genes revealed further a distinct regulation of transcription in both astrocytes. Interestingly, in CTX astrocytes negative transcription regulators were overrepresented, while in the CB astrocytes more positive regulators were observed (**Figure 3-20C, Figure 3-20D**). This is in line with the overall higher accessibility but comparable global gene expression in CTX astrocytes.



D

CTX		
GO:0007399	nervous system development	App, Bdnf, Camk1d, Dclk1, Dpf3, Enc1, Erbb4, Fezf2, Gfra2, Grip1, Hdac7, Hes5, Kalrn, Lhx2, Mef2c, Nav2, Ngef, Nr4a2, Nrn1, Nrp2, Prkd1, Robo1, Serpine2, Srgap2, Tcf4, Ttll7, Vegfa, Zc3h12a
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	Atf3, Bmp2, Btg2, Cdkn1c, Ctnnbip1, Dlg1, Epas1, Ezr, Fezf2, Fgfr3, Fzd8, Hdac7, Hes5, Hey1, Hey2, Lef1, Lmcd1, Lmo1, Lrp8, Mef2c, Msx1, Nfil3, Nog, Nr2e1, Nr2f2, Nr4a2, Pde2a, Prdm16, Sall1, Smad7, Smyd2, Tbx18, Tcf4, Tfcp2l1, Tle4, Usp2, Vegfa, Vldlr, Zfp217, Zfpm1

CB		
GO:0007399	nervous system development	Ascl1, Atoh8, Btbd3, Camk2g, Chl1, Dock7, Dok5, Efnb2, Gas7, Itm2b, Mtss1, Mylip, Nlgn1, Ntrk3, Pax3, Rbfox1, Sema6d, Sh2b2, Sox1, Zic1, Zic5
GO:0045944	positive regulation of transcription from RNA polymerase II promoter	Akirin2, Akna, Arnt2, Ascl1, Cbfb, Creb3l1, En1, En2, Fgfr2, Foxo1, Lpin1, Lpin2, Maf, Maml3, Meis1, Meox1, Nfix, Pax3, Pbx3, Pik3r1, Rfx3, Rora, Rxra, Smad3, Sox1, Sox5, Wnt5a, Zic1

Figure 3-20 Integrative analysis of epigenetics and transcriptomics. ***A****. Overlap of genes associated with DMR(s), genes with differential ATAC peak(s) and differential expression.* ***B****. Classification of the 661 intersected genes (A) into protein classes.* ***C****. Biological processes of the 661 intersected genes (A) higher expressed in CB astrocytes (n=255, left) or higher expressed in CTX astrocytes (n=405, right).* ***D****. Higher expressed and differentially regulated genes in CTX or CB astrocytes (related to C).*

The previous integration of epigenetic and transcriptome data was based on the premise, that a gene is differentially expressed and associated with at least one DMR and at least one DAR regardless of the relative position of the DMR and DAR. Hence, the next step of data integration was a stricter integration of the three differential analyses. To this end, differentially expressed genes were extracted that were associated with a DMR that overlapped a DAR (by at least 1bp). This led to 194 differential genes following this criterium.

The visualization of the relative differences in DNA methylation, open chromatin and expression revealed that the majority of the cases followed the classical mechanism of epigenetic gene regulation where low DNA methylation and open chromatin lead to gene expression (cluster 1, cluster 5 in **Figure 3-21**). Of note, within cluster 1 the TF gene *Lhx2* was found along with the TF gene *Otx1*, an interaction partner of Emx2 during the forebrain development, as well as the regulatory antisense transcript of Emx2 *Emx2os*. The previous findings suggested that Lhx2 and plays a crucial role in cortical astrocytes (see chapter 3.1.8). The comparative integrative analysis now showed that this factor is also epigenetically distinct regulated and revealed potential regulatory regions within the promoter and first intron of the short *Lhx2* isoform (**Figure 3-21** red box). Recent work based on ChIP-Seq experiments demonstrated that Lhx2 is regulated by the telencephalon specification marker Foxg1 in the developing brain (Godbole et al. 2018). The identified Foxg1 binding site at the *Lhx2* locus overlapped with an enriched open region in cortical astrocytes upstream of the promoter (**Figure 3-21**). However, the strong enrichment of open chromatin together with hypomethylation at the promoter and intron 1 suggested either the involvement of further TFs or an alternative binding site of Foxg1 in adult astrocytes. Cluster 5 represents the genes that were hypomethylated within a region that showed higher enrichment of open chromatin and elevated expression in CB astrocytes. Interestingly, *Zic1* was found among the genes in cluster 5. Similar to the findings of Lhx2 in cortical astrocytes, this result goes hand in hand with the previous results that suggested an important role of Zic1 in CB astrocytes (see chapter 3.1.8). The epigenetic profile at the *Zic1/Zic4* locus revealed potential regulatory regions downstream of *Zic1* (**Figure 3-21** red box). Of note, the epigenetic profiles of *Lhx2* and *Zic1/Zic4* showed

additional DARs that were hypomethylated in both astrocyte populations, indicating that the distinct activity of TFs at these elements was regulated only by DNA accessibility.

In addition, several other genes related to astrocyte-specific programs, including the $GABA_B$ receptors subunit gene *Gabbr1*, GABA transporter gene *Slc6a11*, membrane bound growth factor receptor gene *Fgfr3*, and the metabolic genes *Dagla* and *Hk1* (**Figure 3-21** cluster1) were identified to be hypomethylated, more accessible and higher expressed in cortical astrocytes. The NMDA receptor gene *Grin2b*, the metabolic genes *Hk2* and *Man1c1* were identified by this integrative analysis to be hypomethylated, more accessible and higher expressed in in CB astrocytes (**Figure 3-21** cluster5).

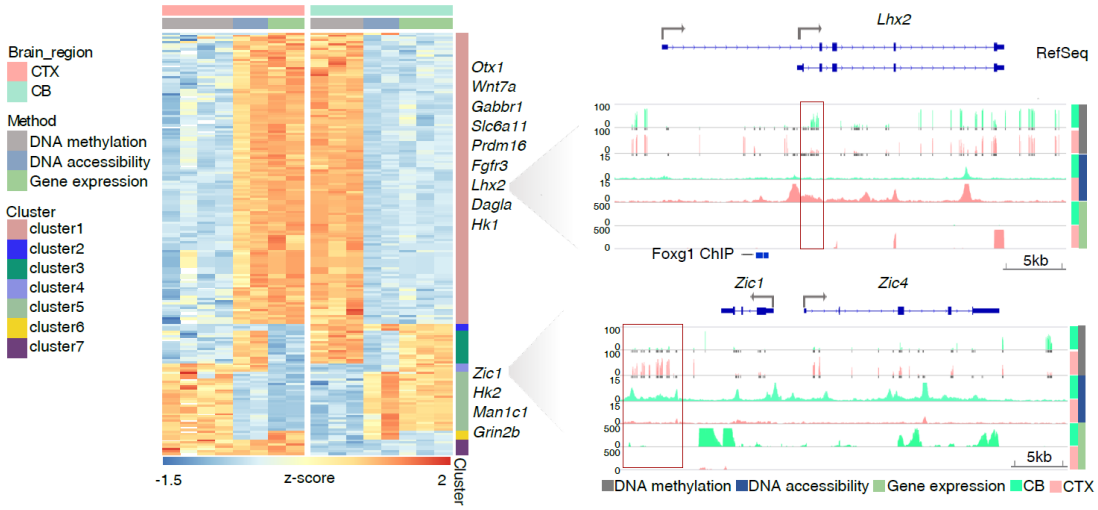


***Figure 3-21 Locus based integrative analysis of epigenetics and transcriptomics**. On the left: Heatmap displaying methylation changes, chromatin accessibility changes and expression changes of genes with DMR(s) that overlap DAR(s). Row-Z-scores were calculated for each data set separately using logRPKM, DNA methylation in %, and DNA accessibility in cpm. On the right: Genomic browser representation of data of the Lhx2 and Zic1/4 loci as representative genes found in cluster 1 or cluster 5. The red box indicates the epigenetically different positions. Blue bar marks a Foxg1 binding site in the Lhx2 locus. Bars and peaks of biological replicates are displayed as overlaid tracks for genomic browser representations. Black bars mark covered CpG position.*

The TF motif enrichment analyses of DMRs and DARs together with expression profiles and local epigenetic differences of *Lhx2* and *Zic1* strongly suggested the involvement of these factors in the differential molecular features of cortical and cerebellar astrocytes. Therefore, known or predicted interaction partners of Lhx2 and Zic family members were retrieved from the protein-protein interaction database STRING, and the epigenetic states of these genes were analyzed (**Figure 3-22**). Various transcription factors such as Foxn3, Otx1, Fezf2, Nr2e1, Emx2, Lef1, Wnt7a, Msx1 and Dbx2 form a network around Lhx2 and were predominantly expressed in CTX astrocytes. Together with higher expression, these genes were all associated with differential DNA accessibility and differential DNA methylation. In CB astrocytes, twelve TFs formed a network around Zic1. Among the connected genes/proteins the network comprises nearly all members of the Iroqois homeobox containing TF family Irx1, Irx2, Irx3 and Irx5. The TFs have been associated with brain development in fish, frog, and chicken, but were so far not characterized in the context of mouse brain function or development (Matsumoto et al. 2004; Lecaudey et al. 2005). Zic1 has also been shown to interact with the midbrain-hindbrain regionalization transcription factors Pax3 and En2 (Sato, Sasai, and Sasai 2005; Nakata et al. 1998). *En2* has an accessible region around its TSS in CB astrocytes, it was hypomethylated and higher expressed in CB. Published Zic1 ChIP-Seq data of adult cerebellum confirmed the binding of Zic1 at this particular DAR.

In conclusion, the comparative integrative analysis demonstrated that the gene expression of CTX astrocyte-, and CB astrocyte-specific genes were strongly controlled by the epigenetic state of transcription factors within networks formed around the developmentally important TFs Lhx2 and Zic1, which themselves were affected by cell type-specific epigenetic control (**Figure 3-21, Figure 3-22**).

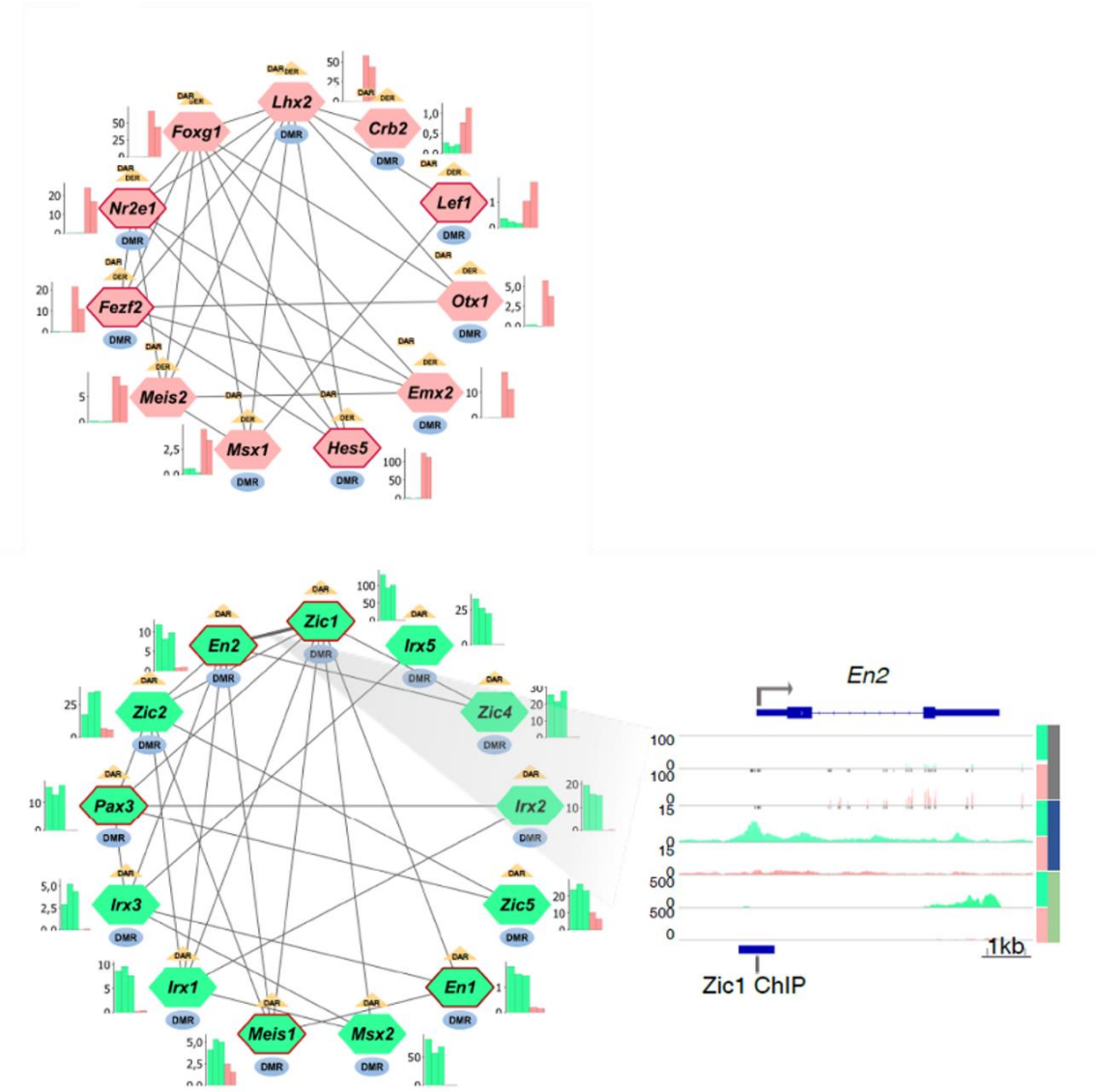


Figure 3-22 Network analysis of Lhx2 and Zic family members. *Bar diagrams display the expression values, blue circle indicates the presence of a DMR in the respective gene and a yellow triangle indicates the presence of a DAR for the respective gene. The red outlined genes mark the genes involved in transcription regulation (related to* ***Figure 3-20****D). The genome browser representation shows the En2 locus, the blue bar indicates a Zic1 binding position in P60 cerebellum.*

3.1.10 Discussion

Astrocytes were long regarded as a homogenous cell type but are now known to be a morphologically and functionally diverse cell population whose complexity is only beginning to be appreciated. Morphological differences of protoplasmic astrocytes in grey matter, fibrous astrocytes in white matter, radial Bergmann glia (BG) in cerebellum and Müller cells in the retina were extensively described since the early stages of neuroscience. However, the distinct morphologies were not assigned to functional differences mainly due to the lack of suitable techniques. Technical advancement and detailed physiological studies demonstrated common functionalities between astrocytes from distinct brain regions, such as K^+ homeostasis, glutamate uptake, gap junctional communication, as well as marker expression (GFAP, S100B, glutamine synthetase) (Reichenbach, Derouiche, and Kirchhoff 2010). In addition to the common astroglial features, physiological and molecular studies revealed region-specific functionality and specific molecular features of astrocytes from distinct brain regions (Boisvert et al. 2018; Chai et al. 2017; Morel et al. 2017; Zeisel et al. 2018; W. Todd Farmer and Murai 2017). Developmental and transcriptome approaches investigated the establishment of astrocyte diversity and, in particular by bulk and single cell transcriptomic studies, indicate that the molecular properties of astrocytes are established by cell-extrinsic cues. These include neural tube patterning during brain development, which leads to region-specific transcription factor expression in astrocytes, as well as signals derived from distinct neuronal environments, implying an instructive, neuronal activity dependent mechanism. To better understand the underlying cell-intrinsic mechanisms, that include a complex balance of transcriptional regulation, epigenomic and transcriptomic signatures of cortical protoplasmic astrocytes and astrocytes of the cerebellum were systematically compared in this thesis.

Astrocytes from cerebellum and cortex show commonalities and differences

The initial comparison of the data showed that CB and CTX astrocytes shared a broad spectrum of similarities. First, the global DNA methylation was similarly high and gene expression levels were similarly distributed in both astrocyte populations (chapter 3.1.3,

chapter 3.1.5). Second, the local DNA methylation and DNA accessibility was at comparable levels at genes typical for astrocytes, neurons, and oligodendrocytes. Finally, gene expression data confirmed various known core astrocyte functions, such as neuroprotection, metabolism of neurotransmitters, and synapse formation and elimination (chapter 3.1.3). Of note, the list of the 20 most abundant astrocyte transcripts in CB and CTX astrocytes revealed several genes whose functions are unclear.

The data also revealed differences between CTX and CB astrocytes. The epigenomic and transcriptomic data strongly suggested that astrocytes of the adult mouse brain retained specific signatures of epigenetic memory of early regional radial glia specification, a principle that also governs neuronal diversity (Greig et al. 2013). The memory of regional specification is in line with previous developmental approaches, such as fate mapping experiments, demonstrating pre-specified spatial diversification of astrocytes in the brain. The early regionalization memory was evident in nearly all results of the data analyses. For instance, the transcription factors *Lhx2*, *Emx2*, and *Foxg1,* which have been shown to play a crucial role in the regional specification of cortical neural progenitors, were strongly expressed in cortical astrocytes (Hanashima et al. 2004; Molyneaux et al. 2007; Monuki, Porter, and Walsh 2001). Moreover, hypomethylated and more open regions in CTX astrocytes were enriched for Lhx2/Emx2 binding motifs. As counterpart in CB astrocytes, the transcription factor Zic1 was found to be highly expressed and biding sites of Zic1 were strongly overrepresented in CB-specific open regions. Zic1 was shown to be involved in the development of the cerebellum (Aruga et al. 1998, 1). Of particular interest, it takes part in chromatin reorganization and establishment of gene expression profiles of mature neurons during development of granule cells, a neuronal cell population in the cerebellum (Frank et al. 2015). The high expression and the overrepresentation of binding sites for Lhx2 and Zic1 in cell type-specific open chromatin regions, and for Lhx2 in DMRs, argues for their important role in maintaining cell-specific expression of downstream target genes.

Gene expression programs of CTX and CB astrocytes are strongly controlled by transcriptional and epigenetic networks around Lhx2 and Zic1

The inspection of the interaction partners of Lhx2 and Zic1 further highlighted the programmatic differences manifested in the course of early brain development. Lhx2 was connected to transcription factors such as Foxn3, Otx1, Fezf2, Nr2e1, Emx2, Lef1, Wnt7a, Msx1 and Dbx2 that were exclusively or predominantly expressed in CTX astrocytes. These transcription factors are known to be important for brain patterning and neuronal specification (Rubenstein and Rakic 2013). Zic1 is known to interact with a series of homeobox genes, involved in early brain or neuronal development, such as *En2*, *Meis2*, *Pax3*, *Hopx*, and *Irx1-5* (Agoston et al. 2012, 2; Zweifel et al. 2018). These genes showed a strong and predominant expression in adult CB astrocytes. TF's within these networks have previously been shown to be important for regional specification of neurons, indicating that regional programs of common precursors remain manifested in epigenetic signatures of adult astrocytes. The expression of these TFs within the networks was furthermore accompanied by epigenetic signatures, such as the presence of DARs and DMRs. These epigenetic networks can be related to the functional networks like demonstrated for the connection of Foxg1 and Lhx2, Lhx2 and Hes5, or Zic1 and En2. A recent study showed that Foxg1 acts upstream of Lhx2 during the development of the cortical signaling center hem by regulating its expression (Godbole et al. 2018). The authors identified a Foxg1 binding site in the *Lhx2* locus in E14.5 cortical tissue by ChIP-Seq experiments which overlapped with an enriched open region in cortical astrocytes. de Melo et al., 2016b demonstrated that Lhx2 directly upregulates the Notch signaling pathway effector Hes5 by binding to a *cis*-regulatory region and by this promotes gliogenesis in the developing retina. Published Zic1 ChIP-Seq data of adult cerebellum identified a Zic1 binding site at the promoter of the midbrain-hindbrain regionalization transcription factor En2. This particular region was highly open and hypomethylated in CB astrocytes, indicating the potential activity of Zic1 in *En2* gene expression in CB astrocytes. These examples remarkably demonstrate the interplay of identified astrocyte diversity regulators and local epigenetic state of downstream target genes. These networks can be further used for future studies to

resolve the interplay of all these factors in the establishment and maintenance of astrocyte diversity.

Nfi is involved in the establishment of pan-astrocyte profiles as well as cell type-specific programs

In both adult astrocyte populations evidence for a strong role of the nuclear factor one (Nfi) family was observed. Studies on brain development and astrogliogenesis identified Nfi as a central factor for the onset of astrogliogenesis. It was recently demonstrated that Nfia is involved in DNA-demethylation of gliogenic gene promoters such as *Gfap* and *Olig1* and promotes the transition from neurogenesis to gliogenesis in radial glia (Deneen et al. 2006; Piper et al. 2010; Sanosaka et al. 2017; Tiwari et al. 2018; Namihira et al. 2009). In line with this, the ATAC-Seq data showed a very high enrichment of Nfi binding sites in about 25 % of open chromatin sites shared between CTX and CB astrocytes. This overrepresentation of binding sites was accompanied by a high and almost equal expression of *Nfia*, *Nfic* and *Nfix* in CTX and CB astrocytes. The enrichment of Nfi binding sites in shared open regions together with the pan-astrocyte expression of *Nfi* suggests a broad regulation of common astrocytic programs in both brain regions. Interestingly, within region-specific open regions Nfi binding sites were also overrepresented. Consequently, Nfi might also be of importance for cell type-specific programs through differences in epigenetic control. Of note, Nfi is so far regarded as a factor involved in early astrocyte differentiation. However, gene expression data as well as published single cell expression data (Appendix **Figure 4-6**) demonstrated the persistent expression of Nfi in adult astrocytes, suggesting the importance of Nfi family members not only for establishment but also maintenance of astrocyte identities.

Environmental influence on astrocyte diversity

Bergmann glia, the astrocytes of the cerebellum come with a unique morphology and a specialized gene expression profile optimized for dealing with the high concentrations of glutamate in the molecular layer (W. Todd Farmer and Murai 2017). Astrocytic adaption to the neural circuit environment was also reported for striatal and hippocampal astrocytes, where measured K^+-currents differed in both astrocyte

populations and were accompanied by distinct expression of K^+ channels (Chai et al. 2017). Indeed, in this thesis, a functional adaptation of astrocytes to the neuronal environment was reflected in expression of receptors, transmembrane receptors, ion channels, transporters, signaling molecules and cytoskeletal proteins in CTX and CB astrocytes. Many of these genes were linked to epigenetically controlled regulation such as *Gabbr1*, encoding the subunit 1 of the $GABA_B$ receptor, and *Slc6a11*, encoding the GABA transporter Gat3, which were hypomethylated, more accessible and higher expressed in cortical astrocytes. *Grin2b*, encoding the glutamate NMDA receptor subunit 2B, was higher expressed in CB astrocytes and was linked to hypomethylation and higher accessibility at its gene locus.

In the cerebellum, neuron-derived Shh protein controls the gene expression profile of Bergmann glia (W. T. Farmer et al. 2016). Indeed, a clear difference in expression and underlying epigenetic signatures were distinct between CB and CTX astrocytes. In CB astrocytes, the epigenetic states were permissive for gene expression of Shh receptor *Ptch2* and the effector *Gli2* which results in their higher expression (**Figure 3-23**).

The local epigenetic differences of genes related to signal perception and transmission leads to the hypothesis that extracellular signals are transduced to the nucleus where effectors impinge on "preset" regulatory landscapes of epigenetic control to readily activate or repress the transcription of respective neurotransmitter receptor and transporter genes.

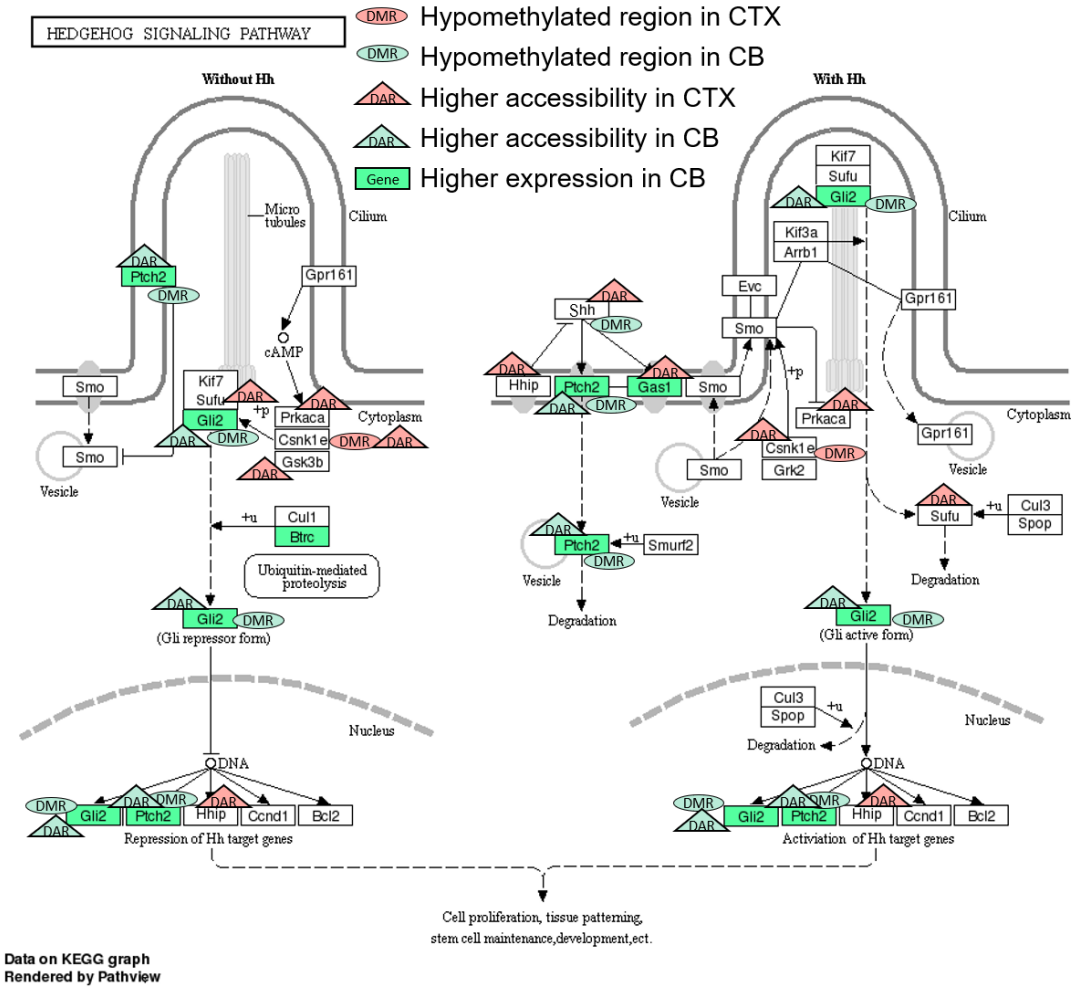


Figure 3-23 Sonic hedgehog signaling pathway. *Genes with differential expression, DNA methylation, and DNA accessibility between CTX and CB astrocytes are marked by color, circles, and triangles, respectively.*

Astrocytes from cerebellum and cortex show extensive differences in chromatin organization

One major observation was the striking difference of chromatin openness around DARs of astrocytes from CTX and CB. The overall higher DNA accessibility in CTX astrocytes may reflect a higher degree of heterogeneity. Cortical astrocytes are organized in layers and are interacting with layer-specific neurons which results in individual gene expression programs (Lanjakornsiripan et al. 2018). Furthermore, astrocytes of the cerebral cortex reside within functional restricted areas like somatosensory cortex, motor cortex and visual cortex, which was also shown to be reflected in slightly distinct

gene expression profiles (Boisvert et al. 2018). Finally, transcriptome analysis on single cell level demonstrated a broader spectrum of individual gene expression programs in cortical protoplasmic astrocytes than in CB astrocytes (Appendix **Figure 4-3**; Zeisel et al. 2018).

Besides the degree of heterogeneity, an overall distinct chromatin organization might explain the observed chromatin accessibility differences. In the cerebellum, the detection of a higher gene expression of core histones was accompanied by a higher expression of histone chaperones. Furthermore, higher signal intensities of H3K27ac (enhancer) and H3K27me3 (heterochromatin) marks were shown by immunohistochemistry. These results lead to the speculation that the chromatin of CB astrocytes may comprise an overall higher nucleosome density than cortical astrocytes which enhances epigenetic signals and affects the general DNA accessibility. This hypothesis has to be tested by e.g. mass spectrometry in the future.

Finally, transcriptional repressors were overrepresented among higher expressed genes in CTX astrocytes. The elevated activity of transcriptional repressors could explain the discrepancy between a strong difference in openness and a comparable expression rate in CTX and CB astrocytes. Intriguingly, Lhx2 was shown to regulate the subcortical neuronal subtype identity through interaction with the nucleosome remodeling and histone deacetylase (NuRD) complex at *Fezf2* and *Sox11* loci leading to a repression of these genes (Muralidharan et al. 2017). Lhx2 also regulates the chromatin accessibility in retinal progenitor cells (Zibetti et al. 2017). Lhx2 may therefore be involved in the higher chromatin accessibility in cortical astrocytes through binding or recruiting chromatin remodelers, or as has been shown by Lhx2 knockout in Müller glia, to regulate such effects indirectly through the Notch signaling pathway (de Melo et al. 2016; de Melo, Clark, and Blackshaw 2016). This would in turn mean, that CB and CTX astrocytes employ distinct basic transcriptional regulatory mechanism, that determine the global transcription rate by the nucleosome density/DNA accessibility (in the case of CB astrocytes), or by the active repression of gene expression (in the case of CTX astrocytes). This hypothesis can be explored in the future by knockout experiments of Lhx2 in astrocytes and subsequent monitoring of the global and target gene expression.

Overall conclusion

In summary, the findings provided insights into regionally transcription factor-controlled regulation of astrocyte identity. Key regulators, previously identified as regional specifiers in neurons, are also prominent actors in astrocytes. Their function seems to be embedded into layers of transcriptional and epigenomic control with regional specificities. The question remains how the clearly distinct functionalities are implemented. Genome wide epigenetic data along with the respective transcriptomes provides a good resource for prediction of regulatory networks which in turn can be targeted to clarify the control of regional, and cell (sub)type-specific gene expression. The observed regional specification might be a result from a combination of developmental and spatially determined epigenomic programming to match the interaction of neurons and astrocytes in distinct neural circuits. Future studies could explore the dynamics of epigenetic and transcriptional changes in astrocyte development in distinct brain regions to investigate whether regional TFs regulate the same target genes in neurons as they do in astrocytes or do these factors regulate genes with complementary functions to ensure regional matched interaction? Gradually disentangle these questions will extend the understanding of computational processes in distinct neural circuits and neural circuit related diseases.

3.2 LAG-3 Inhibitory Receptor Expression Identifies Immunosuppressive Natural Regulatory Plasma Cells.

3.2.1 Introduction

B-lymphocytes play multifaceted roles in immune responses such as infection, autoimmune reactions, and cancer. They execute stimulatory and inhibitory functions in order to protect the organism from infectious diseases and remove damaged cell components. The anti-inflammatory function is of great importance to regulate immune responses and to prevent disproportionate reactions against auto-antigens. The first reported evidence for the regulatory effect of B-cells was demonstrated in mice lacking B-cells in the context of experimental autoimmune encephalomyelitis (EAE), an animal model for autoimmune central nervous system disease (Wolf 1996). These mice suffered a severe progression of EAE and further dissection of the underlying mechanisms revealed the production of interleukin-10 (Il10) by B-cells as key factor in this autoimmune disease (Fillatreau et al. 2002). It was further demonstrated that Il10 expression in B-cells was activated in anti-microbial immune reaction driven by various types of microbes including bacteria, helminths and viruses (Shen and Fillatreau 2015). However, the identity of the B-cell subpopulation producing Il10 as a suppressive factor *in vivo* was yet incompletely defined. *In vivo* experiments using Il10-eGFP reporter mice infected with *Salmonella typhimurium* showed an accumulation of Il10 producing cells in the spleen within 24 hours after infection. Moreover, these cells expressed elevated levels of the plasma cell-specific surface marker CD138, identifying a subset of plasmocytes[8] as source of Il10 production after challenge (Neves et al. 2010).

The identification of plasmocytes as regulatory B-cells (Bregs) raises the questions about the progenitors of this B-cell subtype and the difference of these plasmocytes in comparison to non-regulatory plasmocytes. DNA methylation is a key mechanism to establish and maintain cell identity and functional competence, and contains information of the origin of the cell (see chapter 1.1.1 on page 10 and chapter 1.1.2 on page 15). Therefore, in a collaboration with Prof. Dr. Simon Fillatreau and his

[8] Collective term for plasmablasts and plasma cells.

group this part of the thesis addressed these questions by comprehensive epigenome and transcriptome analyses. Initial gene expression analysis and subsequent validation by flowcytometry of Il10^{+}CD138^{hi} cells in comparison to Il10^{-}CD138^{hi} cells by the collaborative group revealed discriminative expression of the inhibitory receptor lymphocyte-activation gene 3 (Lag3), an MHC-class-II-binding CD4 homolog. This surface protein is known to be expressed on activated CD4^{+} T cells, but also on MHC class II non-binding cells such as activated CD8^{+} T cells, natural killer cells and myeloid cells, suggesting multifaceted functions of Lag3 (Lui and Davis 2018).Interestingly, in contrast to the homologous CD4, inhibition of Lag3 prolongs T cell responses (Huard et al. 1994). However, the mechanism of action remains unclear. Using Lag3 as a tracer, the group around Prof. Dr. Simon Fillatreau discovered that Lag3^{+}CD138^{hi} plasma cells were present in the spleens of naïve mice. These naïve Lag3^{+}CD138^{hi} plasma cells rapidly upregulate Il10 upon *Salmonella typhimurium* infection. To delineate the molecular properties and possible origin of Lag3^{+}CD138^{hi} plasma cells, a comparative analysis of the methylomes of Lag3^{+}CD138^{hi} plasma cells and other B-cell subtypes was performed. For this comparison, Dr. Andreia Lino isolated various B-cell subsets including Lag3^{+}CD138^{hi} plasma cells and Lag3^{-}CD138^{hi} plasmocytes from the spleen of naïve mice and Lag3^{+}CD138^{hi}, Lag3^{-}CD138^{hi}, Il10^{+}CD138^{hi}, and Il10^{-}CD138^{hi} cells isolated at day 1 post infection (p.i.) with *Salmonella typhimurium.* Furthermore, B1a, B1b, and B2 from the peritoneal cavity as well as the immature transitional cells T1, and T2, follicular and marginal zone B-cells from the spleen of naïve mice were isolated and subjected to methylome analysis. In addition to the DNA methylation analysis, transcriptomes of Lag3^{+}CD138^{hi} plasma cells and Lag3^{-}CD138^{hi} plasmocytes of naïve and day1 p.i. mice were profiled and integrated with the DNA methylation analyses (**Table 18**). The results of these comparative analyses were partially published (Lino et al. 2018).

Table 18 Overview of analyzed B-cell subsets. *The number of biological replicates for the methylome and transcriptome approaches is given for the respective B-cell subset. "Day1" labels the Salmonella typhimurium infected cells 24 hours after infection.*

Origin	B-cell subtype	Methylomes (n)	Transcriptomes (n)
Spleen	T1	3	
Spleen	T2	3	
Spleen	Marginal zone (MZ)	3	
Spleen	Follicular (FO)	3	
Spleen	$Lag3^-$ plasma cells naïve ($Lag3^-$)	3	2
Spleen	$Lag3^-$ plasma cells day1 ($Lag3^-$ day1)	3	2
Spleen	$Lag3^+$ plasma cells naïve ($Lag3^+$)	3	2
Spleen	$Lag3^+$ plasma cells day1 ($Lag3^+$ day1)	3	3
Spleen	*Il10*eGFP$^-$ Plasma cells day1	3	
Spleen	*Il10*eGFP$^+$ Plasma cells day1	3	
Peritoneal cavity	B1a	3	
Peritoneal cavity	B1b	3	
Peritoneal cavity	B2	3	

3.2.2 Quality control of transcriptome data

For the molecular characterization of $Lag3^+$ plasma cells, sorted cells were used for the preparation of replicate libraries for sequencing on an Illumina HiSeq2500. The transcriptome was profiles by mRNA-Seq, and genome-wide DNA methylation by RRBS (see chapter 2.2.4 above, and chapter 2.2.5 on page 47 for more details).

The following chapters will first cover the transcriptome analyses by which the functional differences between naïve $Lag3^+$ plasma cells and naïve $Lag3^-$ plasma cells, as well as the difference between activated $Lag3^+$ plasma cells and $Lag3^-$ plasma cells will be determined. Moreover, the molecular processes after activation of $Lag3^+$ plasma cells will be analyzed. Following the transcriptome analyses, the methylome comparison will deal with both, the distinct features of $Lag3^+$ plasma cells and their origin. The subsequent integrated analysis will highlight the basic programmatical differences that define the identity of $Lag3^+$ plasma cells.

The data quality of the sequenced mRNA libraries was assessed by evaluation of various features of the mapped sequencing reads. The exonic rate was similarly high and the intronic rate similarly low among the samples, indicating a proper distribution of the reads across exons and reflecting a high-quality transcriptome profile. The number of detected genes was in the range of 13,000-18,000, which reflects a gene repertoire of highly specialized cells. Ribosomal RNA was nearly not detected in the samples. Taken together, the sequencing data represented transcriptomes with high quality (**Table 19**).

Table 19 Quality control of sequenced mRNA libraries.

mRNA-Sequencing						
Sample	**Reads sequenced**	**Transcripts detected**	**Genes detected**	**rRNA rate**	**Exonic Rate**	**Intronic Rate**
Lag3$^-$ R2	67709257	48006	16706	0.001	0.72	0.15
Lag3$^-$ R1	62957878	49819	17875	0.001	0.68	0.17
Lag3$^-$ day1 R1	61729857	46429	15819	0.001	0.74	0.16
Lag3$^-$ day1 R2	88672096	49235	16331	0.001	0.76	0.14
Lag3$^-$ day1 R3	51726966	43251	14647	0.001	0.75	0.15
Lag3$^+$ R1	30929565	37385	13481	0.001	0.71	0.16
Lag3$^+$R2	63315503	47683	18216	0.001	0.62	0.20
Lag3$^+$ day1 R1	30796942	38933	13411	0.001	0.74	0.16
Lag3$^+$ day1 R2	36308480	38656	13389	0.001	0.74	0.15

3.2.3 Transcriptome of Lag3$^+$ plasma cells shows an overall plasma cells conformity

In order to characterize the differences between Lag3$^+$ and Lag3$^-$ cells, pairwise comparisons of the expression of all detected genes in naïve and 1dpi Lag3$^+$ and Lag3$^-$ cells were performed. In fact, this first comparative analysis revealed a high correlation between both plasma cell types (**Figure 3-24A**; Pearsosn r=0.96 - 097), indicating an overall similar transcriptional profile. The scatterplots also showed subsets of genes that were higher expressed in naïve and 1dpi Lag3$^-$ cells compared to naïve and 1dpi Lag3$^+$ plasma cells. Inspection of these subsets disclosed distinct transcription levels of numerous immune globuline variable heavy and light chains between these plasma cell types (Appendix **Table 22**). The evaluation of the highest expressed genes in the respective cell subpopulations showed, as expected from plasma cells, a predominat and remarkably high expression of immune globuline components (**Figure 3-24B**). The majority of these genes were transcribed at a comparable level, however, some exceptions were observed. For instance, *Igkv14-126* was higher expressed in Lag3$^+$ (naïve and 1dpi) than in Lag3$^-$ (naïve and 1dpi). This is of particular interest, since *Igkv14-126* encodes a component of the VH11$^+$Vk14.126$^+$ B-cell receptor (BCR), typical for B1a cells (Hardy, Wei, and Hayakawa 2004). Notably, both naïve Lag3$^+$ and Lag3$^+$ isolated from infected mice on day 1 p.i. expressed mostly IgM, the first immunoglobulin class produced after antigen exposure, indicating that they had not undergone isotype switching. In contrast, naïve and day1 Lag3$^-$ plasma cells expressed IgA at highest levels as well as IgM, IgG, and IgE.

In summary, the examination and first comparisons of Lag3$^+$ and Lag3$^-$ plasma cell transcriptomes revealed an overall high similarity, while pointing to a distinct repertoire of immune globulins.

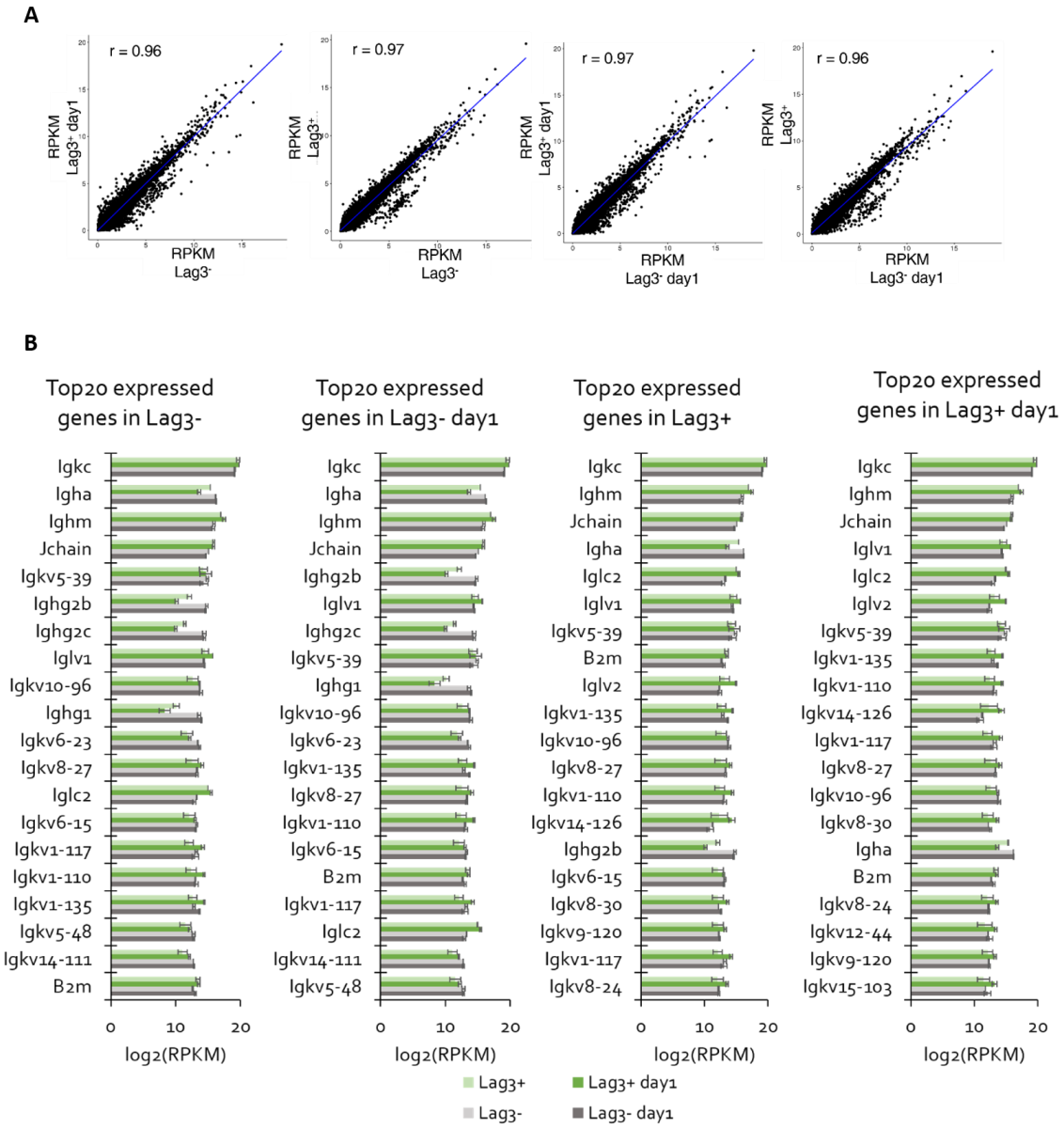


***Figure 3-24 Pairwise comparison of plasma cell transcriptome profiles. A**. Pairwise correlation of gene expression using RPKM values; Pearson correlation coefficient. **B**. Expression values of the highest expressed genes in the respective plasma cell subpopulation.*

3.2.4 Differential gene expression highlights cell cycle control and immune response as the major processes discriminating Lag3$^+$ and Lag3$^-$ plasma cells

To determine the transcriptional variance between Lag3$^+$ and Lag3$^-$ plasma cells, principal component analysis and hierarchical clustering were performed. Both methods revealed a clear separation of Lag3$^+$ plasma cells from Lag3$^-$ plasma cells (**Figure 3-25**). Moreover, the distance in PCA (PC2) and the height of the dendrogram branches indicated a stronger divergence between Lag3$^+$ naïve and Lag3$^+$ day1 than Lag3$^-$ naïve and Lag3$^-$ day1, indicating a stronger impact of the activation on the cell function of Lag3$^+$ cells.

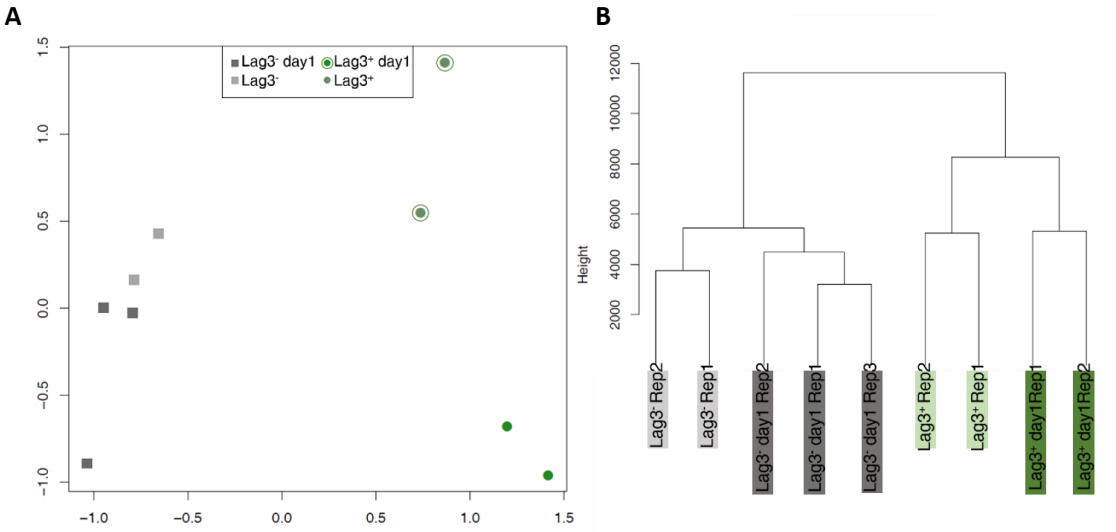


***Figure 3-25 Gene expression variance between Lag3$^+$ and Lag3$^-$ plasma cells (naive and 1dpi). A**. PCA based on the 1000 most variable genes detected in naïve and 1dpi Lag3$^+$ and Lag3$^-$ plasma cells. **B**. Hierarchical clustering of full transcriptomes; Manhattan distance coupled to ward method represented.*

To identify significant functional differences discriminating Lag3$^+$ from Lag3$^-$ plasma cells, a differential analysis on the mRNA-Seq data was performed using the R package edgeR (see chapter 2.2.11 on page 53 for more details). Using cutoffs for significance of differential expression (pval ≤ 0.01; FDR ≤ 0.05), 648 DEGs were found in the comparison of naïve Lag3$^+$ and naïve Lag3$^-$ plasma cells (**Figure 3-26**). A nearly four-fold higher number of DEGs (n= 2280) was found in the comparison of Lag3$^+$ and Lag3$^-$ plasma cells on day1 after *Salmonella* challenge, suggesting that their transcriptomes became increasingly different after challenge. The majority of the 648 genes found to be differentially expressed between naïve Lag3$^+$ and naïve Lag3$^-$ plasma cells were also found in the other comparisons (n= 379+11+60+6), indicating distinct ground states of Lag3$^+$ compared to Lag3$^-$ plasma cells that are maintained after activation. Among all DEGs, 11 genes were found to be differentially expressed between Lag3$^+$ and Lag3$^-$ and between the naïve and activated state (**Figure 3-27A** lower panel). The function of these genes is not well characterized for plasma cells but can be summarized into inflammatory (*Pycard, Slamf9, Ifi27l2a, Rsad2, Il2rg*), signaling (*Lpp, Anxa2*), and metabolic (*Tg, Ada, Ctsz*) functions.

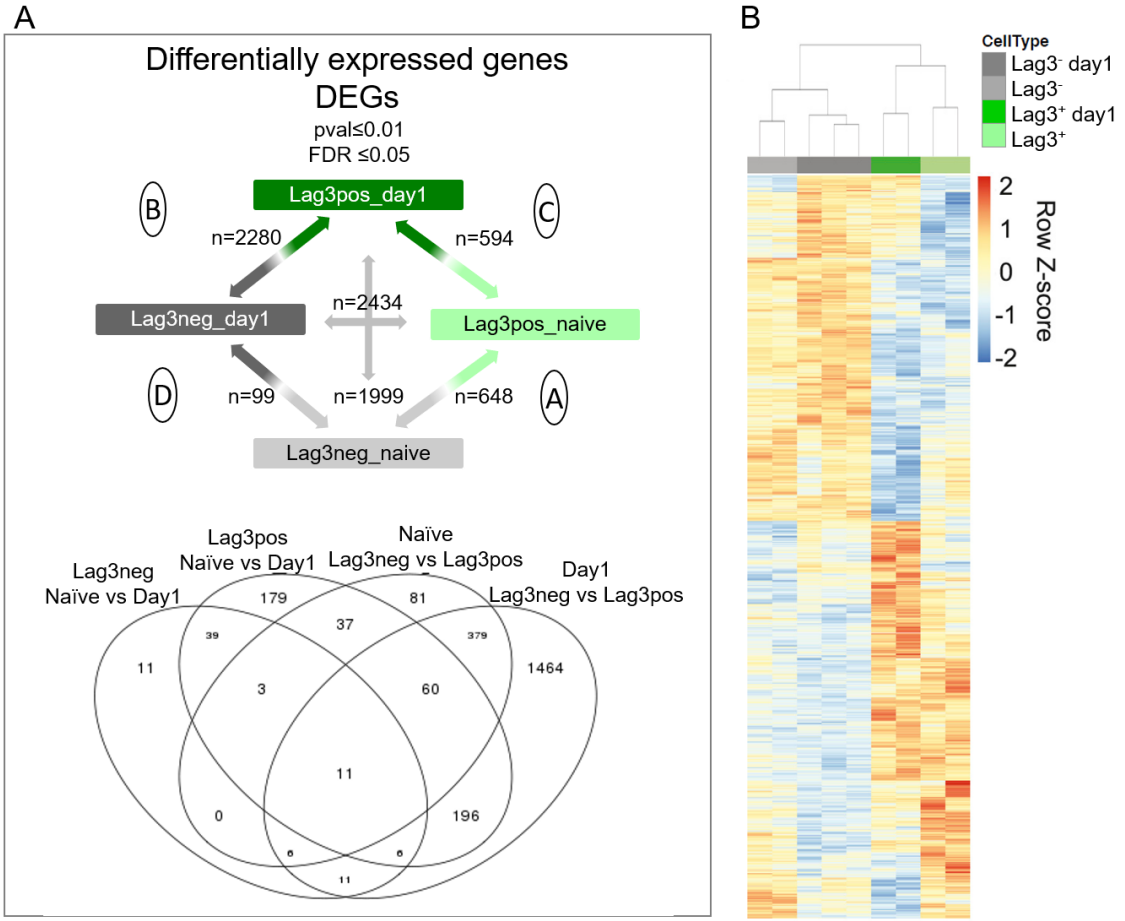


Figure 3-26 Differential gene expression in plasma cells. A*. Diagram showing the number of differentially expressed genes (DEGs) resulting from the respective comparison (upper panel). The venn diagram displays the number of intersected genes from each comparison (lower panel).* ***B****. Heatmap representing the gene expression difference of all detected DEGs from all comparisons (n= 3851). Row-Z-scores were calculated based on RPKM values.*

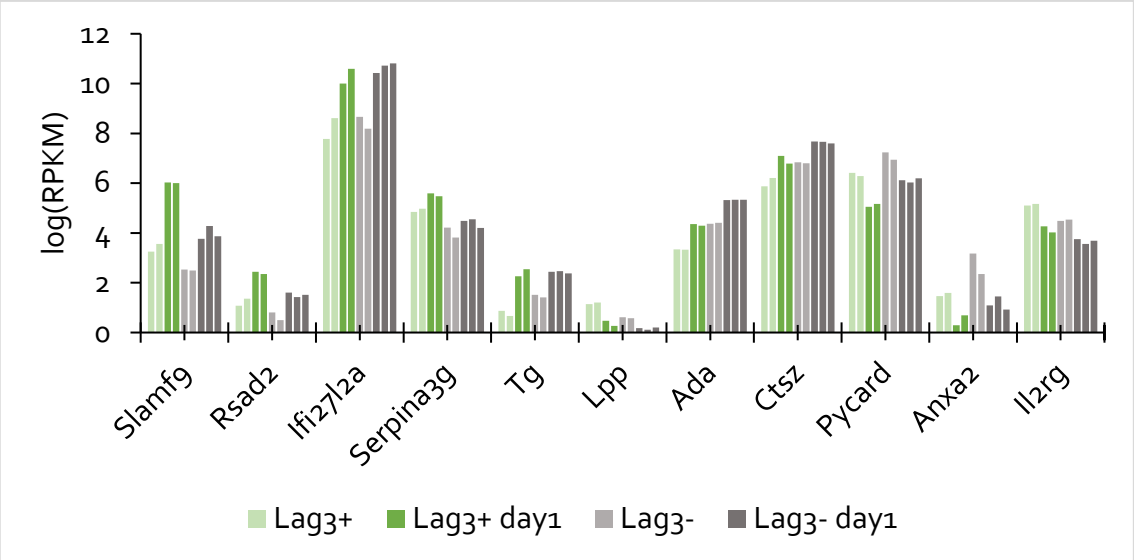


Figure 3-27 Gene expression of expressed genes found to differentially expressed in all analyzed plasma cells. *Related to* ***Figure 3-26****A.*

Among the top DEGs identified by the comparison of naïve Lag3$^+$ and Lag3$^-$ plasma cells and day1 Lag3$^+$ and day1 Lag3$^-$ plasma cells were genes coding for Il10, cell surface proteins such as Lag3, CD200, L-Selectin (*Sell*) and Cxcr3, transcription factors such as Egr1 and Egr2, and immune globulins such as IgG and IgE (**Table 23**, **Table 24**). Lag3 and CD200 are known receptors implicated in negative regulation of immunity (Mihrshahi, Barclay, and Brown 2009, 200; Lui and Davis 2018, 3). Thus, CD200 extends the mechanisms by which Lag3$^+$ plasma cells mediate regulatory function. Of particular interest, the transcription factor Egr2 was shown to regulate the expression of *Lag3* and *Il10* in regulatory T-cells (Tregs) (Zheng et al. 2013; Okamura et al. 2009, 4). This suggests a shared program between Tregs and Bregs by which they control the anti-inflammatory function. The distinct expression of *Sell* (higher expressed in naïve and day1 Lag3$^+$ plasma cells) and *Cxcr3* (higher expressed in naïve and day1 Lag3$^-$ plasma cells) suggests a distinct migratory program to ensure homing to inflammatory sites, which also points to potentially distinct regions of origin.

A GO-Term enrichment analysis of the 648 DEGs in naïve mice further elucidated the differences between $Lag3^+$ and $Lag3^-$ plasma cells. Genes higher expressed in $Lag3^-$ were mainly associated with cell cycle regulation which indicated a higher proliferation rate of these cells (**Figure 3-28**). In contrast, genes higher expressed in $Lag3^+$ were related to immune system regulation. Similarly, when comparing differentially expressed gene sets in infected mice, cell cycle regulation was the major process annotated to genes higher expressed in $Lag3^-$ plasma cells (**Figure 3-29**). Interestingly, genes higher expressed in day1 $Lag3^+$ plasma cells were associated with chromatin organization, suggesting a restructuring of the chromatin to adapt gene expression as a proliferation-independent activation response.

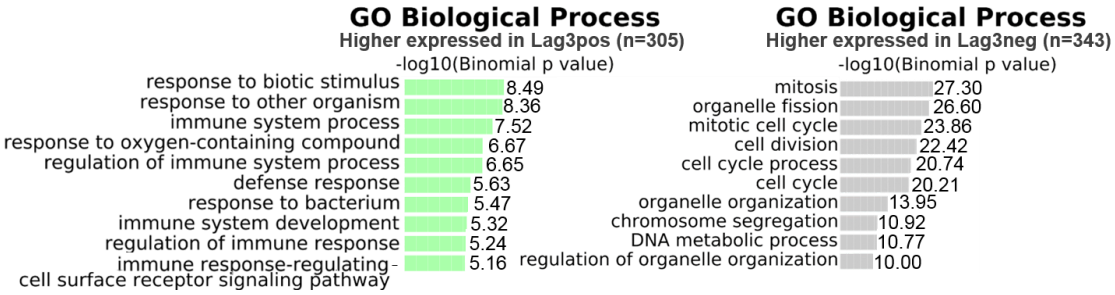


Figure 3-28 GO-Term enrichment of differentially expressed genes between naive Lag3⁺ and naive Lag3⁻ plasma cells.

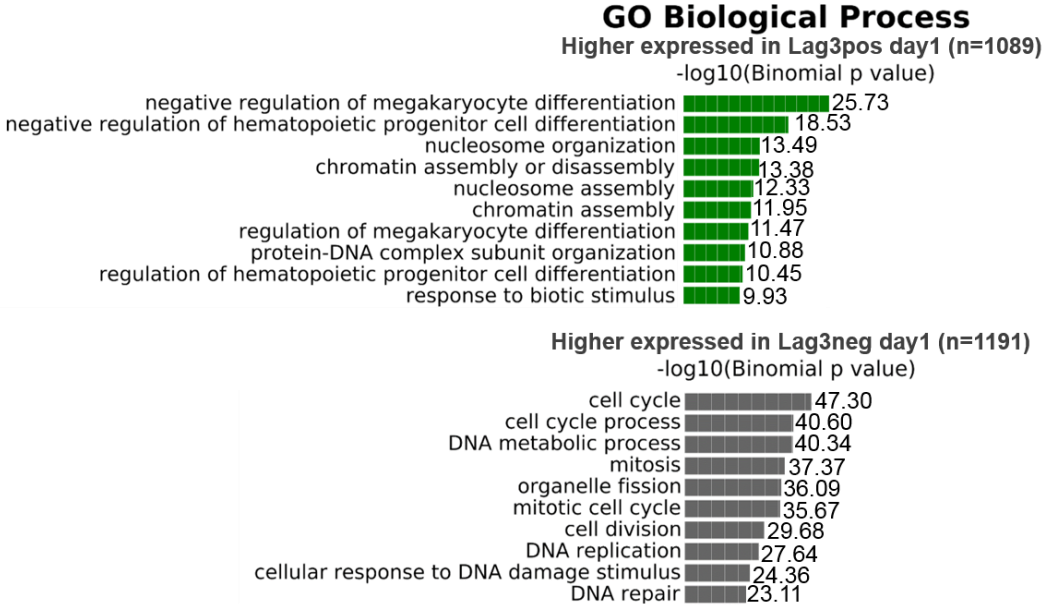


Figure 3-29 GO-Term enrichment of differentially expressed genes between Lag3⁺ day1 and Lag3⁻ day1 plasma cells.

3.2.5 Transcriptional changes in Lag3⁺ plasma cells upon *Salmonella* infection

The differential analysis of Lag3⁺ and Lag3⁻ plasma cells before and after *Salmonella* infection revealed an increasingly different transcriptome after challenge (**Figure 3-26A** (A),(B)). This raised the question how the cells change after activation. The comparison of naïve and day1 Lag3⁺ transcriptomes highlighted the changes in Lag3⁺ cells after antigen exposure (**Figure 3-26A** (C)). In total, 593 genes were differentially expressed after activation of Lag3⁺ plasma cells, of which 320 were higher expressed in the naïve state and 273 were higher expressed 24h after challenge. The genes higher expressed in naïve state were related to cell cycle regulation as well as immune response, while the higher expressed genes at day 1 were enriched for immune system processes only (**Figure 3-30**).

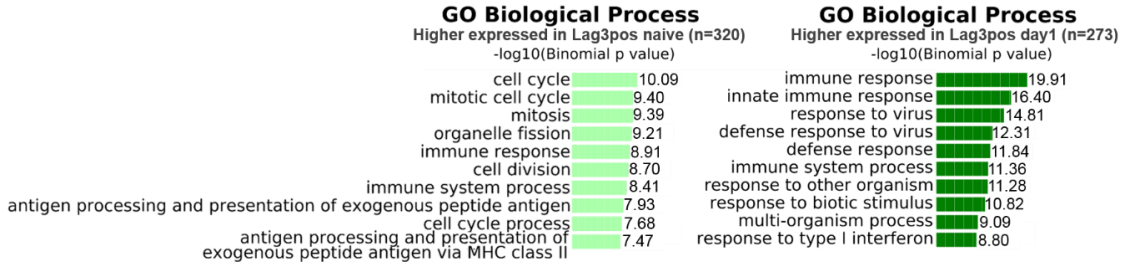


Figure 3-30 GO-Term enrichment of differentially expressed genes between naïve Lag3⁺ and Lag3⁺ day1 plasma cells.

Among the top 50 differential genes that were higher expressed in the activated cells were genes implicated in metabolic functions (e.g. *Tg, Ggta1, Usp18, Ada, Psmd14, Ube2l6, Npc2, Cox7a2l*), membrane bound proteins (e.g. *Lag3, Ly6c2, Ly6c1, Ly6a, Slamf9, Vmp1, Fcgr2b),* as well as the transcription factor genes *Irf7*, *Phf11b*, *Phf11d*, and *Stat1*, involved in the transcriptional activation of cytokines (Appendix **Table 25**). The closer inspection of the 50 most differentially expressed genes of day1 Lag3⁻ plasma cells actually showed that many of the upregulated genes found in Lag3⁺ plasma cells were also upregulated in Lag3⁻ plasma cells after activation (Appendix **Table 26**). This was also displayed by the high overlap shown in the venn diagram in **Figure 3-26A**. Consequently, another strategy was applied to identify genes that were up- or

downregulated solely in activated Lag3$^+$ plasma cells. From the 594 DEGs discriminating naïve and day1 Lag3$^+$ plasma cells, genes were excluded that showed the same direction of expression fold change when comparing the gene expression of Lag3$^-$ plasma cells before and after activation. This resulted in 178 genes that were downregulated only in Lag3$^+$ plasma cells after activation and 122 genes that were upregulated only in Lag3$^+$ plasma cells after activation (**Figure 3-31**). The majority of genes that were downregulated in Lag3$^+$ cells after infection were related to cell division. Of note, the genes in this group were expressed at very low levels, which indicated a low level of cell division in the naïve cells and an even lower rate in the activated state (**Table 20**, Appendix **Table 27**). Moreover, a reduction in the expression of IgG (*IgG2b, IgG2c*) could be observed in Lag3$^+$ plasma cells after infection, while the IgG expression in Lag3$^-$ plasma cells remained the same. Genes that were upregulated only in Lag3$^+$ plasma cells upon infection, were involved in processes related to cell activation and rapid immune response such as signal transduction, translation and protein transfer (**Table 20**, Appendix **Table 28**). In addition, various *Igkv* genes were identified to be highly upregulated after infection in Lag3$^+$ cells, representing unique BCR combinations.

Among the genes with unique expression profiles, *Enpp1* and *Cbfa2t3*, are of particular interest as they can be implicated in Il10 expression of day1 Lag3$^+$ plasma cells. *ENPP1* encodes the plasma cell alloantigen 1 (PC1) protein which is expressed by a subset of Il10 expressing B1a cells in the human (H. Wang et al. 2012). B1a cells that do not express PC1 also do not express Il10 which suggests a co-expression of these proteins. Moreover, PC1 expressing B1a cells were shown to have a protective function during pregnancy by participation in tolerance induction toward the fetus, further demonstrating the regulatory function of this co-expression (Schumacher et al. 2018). Cbfa2t3 is a transcriptional corepressor which facilitates transcriptional repression. This repressor was expressed at the same level in Lag3$^-$, day1 Lag3$^-$ and naïve Lag3$^+$ plasma cells and was strongly reduced in day1 Lag3$^+$ plasma cells (Appendix **Table 27**). The reduction of a repressor in turn may activate the transcription of its target genes. Hence, further studies on the target genes of this repressor would be helpful to identify the transcriptional regulation of *Il10*.

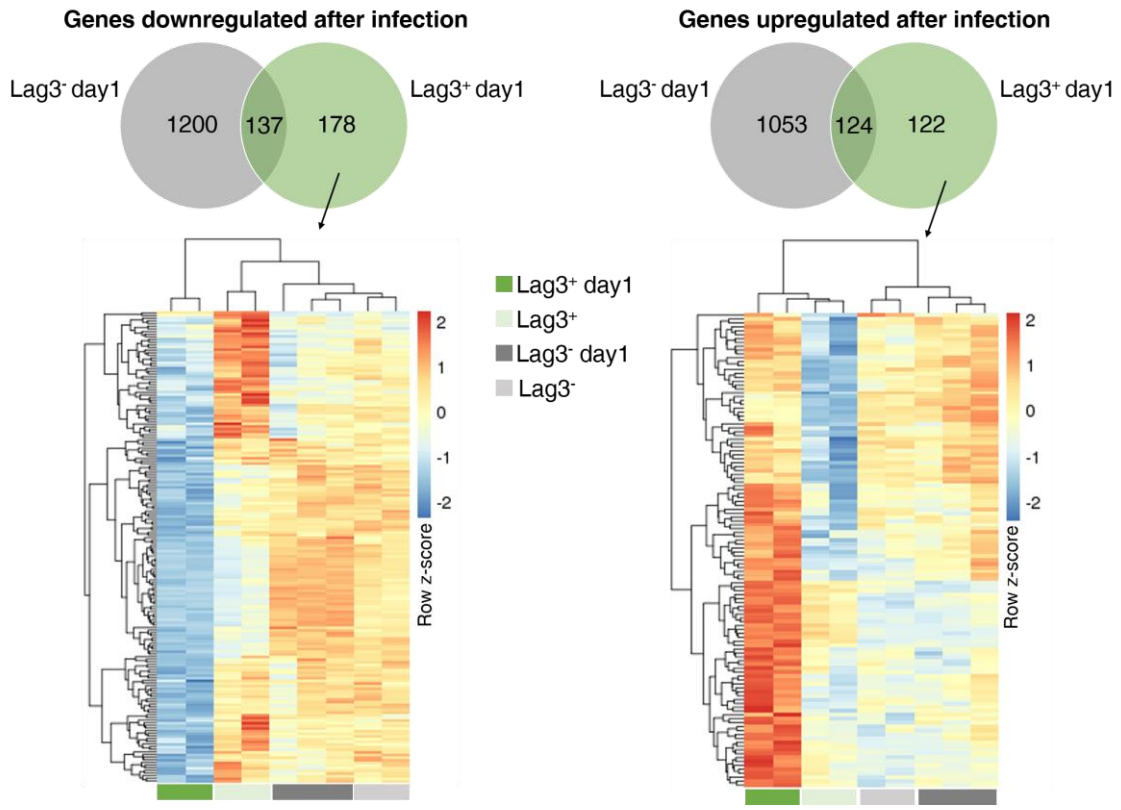


Figure 3-31 Influence of the Salmonella infection on the expression profile in Lag3⁺ plasma cells. *Venn diagram displays the number of genes that are downregulated (left) and upregulated (right) in Lag3⁻ or Lag3⁺ plasma cells after infection. The heatmaps represent the expression difference of the uniquely up- or downregulated genes. Row z-scores were calculated using RPKM.*

Table 20 GO-Term enrichment of genes affected by the Salmonella infection only in Lag3⁺ plasma cells.

Genes that are downregulated only in Lag3⁺ cells after infection		Genes that are upregulated only in Lag3⁺ cells after infection	
Term	PValue	Term	PValue
Cell cycle	3.16E-14	Positive regulation of cap-independent translational initiation	0.016993
Cell division	7.39E-13	Protein transport	0.019791
Cellular response to DNA damage stimulus	5.48E-11	Immune system process	0.022298
Mitotic nuclear division	1.86E-10	Positive regulation of B-cell proliferation	0.024934
DNA repair	1.14E-08	Regulation of translation	0.031878
Mitotic cytokinesis	4.73E-06	Negative regulation of cell growth	0.03461
Chromosome segregation	1.00E-05	Regulation of cell shape	0.045096
Regulation of attachment of spindle microtubules to kinetochore	0.001434	Positive regulation of cellular senescence	0.050125
Positive regulation of exit from mitosis	0.002431	Transport	0.052671
Protein ubiquitination	0.003577	Positive regulation of proteasomal protein catabolic process	0.055538

3.2.6 Transcriptional control of Il10 production in Lag3⁺ plasma cells

The previous analysis pointed to potential transcriptional regulators of *Il10* expression. Thus, a systematically analysis of the transcription factor repertoire in Lag3⁺ and Lag3⁻ plasma cells was performed to identify more potential *Il10* regulators. First, TFs and transcription co-factors, gene expression regulators, and general DNA binding proteins were identified that were higher expressed in Lag3⁺ plasma cells 24h after *Salmonella* infection compared to Lag3⁻ plasma cells after infection (**Figure 3-32A**). Next, the differential genes between naïve Lag3⁺ and naïve Lag3⁻ plasma cells were filtered for TFs that were higher expressed in Lag3⁺ plasma cells (**Figure 3-32B**). Finally, 22 factors were identified with higher expression in Lag3⁺ plasma cells (**Figure 3-32C**). Lag3⁺

plasma cells showed a higher expression of the cell cycle control genes *Klf4*, *Bhlhe40*, and *Bmyc*. The higher expression of *Hes1* indicated higher Notch signaling activity in Lag3$^+$ plasma cells and the higher expression of *Fos*, *Junb*, *Egr1*, *Egr2*, *Cebpb*, *Irf8*, and *Satb1* suggested a distinct regulatory program. In order to link the expression of these factors to the Il10 expression, the DNA sequence of the *Il10* locus (40 kb upstream of TSS and 30 kb downstream of TTS) was scanned for TF binding sites. In fact, 8 of the 22 factors had at least one binding site in the Il10 locus (**Figure 3-32D**, **E**). Notably, Fos, Jun-b, and Egr-2 are known drivers of Il10 expression in other immune cells (Iwasaki et al. 2013; Z.-Y. Wang et al. 2005; Yoshida et al. 2012). In addition, the early growth response factor 2 Egr2 was also shown to be involved in Lag3 expression in CD4$^+$ T cells, along with Blimp-1-mediated induction of Il10 in CD4$^+$ T cells (Iwasaki et al. 2013), and their regulatory activity (Okamura et al. 2009). Furthermore, putative repressors of *Il10* expression were identified by filtering the DEGs for lower expression of TFs and transcription co-factors, gene expression regulators, and general DNA binding proteins in Lag3$^+$ plasma cells (**Figure 3-33A**, **B**, **C**). By comparing the TF binding sites at the *Il10* locus with the resulted TF list, *Foxm1* was identified to be the factor that is reduced in Lag3$^+$ and has a binding site at the *Il10* locus, suggesting potentially negative control of *Il10* transcription (**Figure 3-33D**).

Previous results suggested a co-expression of *Il10* together with *Lag3*, *Cd200* (**Table 23**, **Table 24**) and *Pdcd1lg2* (CD273 or PD-L2; flow cytometry analysis by the group of Prof. Dr. Simon Fillatreau and validated by mRNA expression: RPKM Lag3$^-$ naïve 1.06; Lag3$^-$ day1 0.86; Lag3$^+$ naïve 2.85; Lag3$^+$ day1 2.01). Accordingly, the DNA sequences at these loci were scanned for TF binding motifs to identify factors that bind to the *Il10* locus and in addition to the loci of *Lag3*, *Cd200*, and *Cd273*. Notably, all 9 factors previously found to bind at the Il10 locus (*Aire*, *Fos*, *Junb*, *Egr1*, *Egr2*, *Cebpb*, *Irf8*, *Satb1, and Foxm1*) also had predicted binding sites in the *Lag3*, *Cd200* and *Cd273* loci (**Figure 3-33**).

Taken together, these expression analyses combined with binding motif analyses identified 9 factors as putative regulators of the regulatory activity of Lag3$^+$ plasma cells. The function of these factors has to be validated in further experiments, such as B-cell-specific knock outs or over expression experiments.

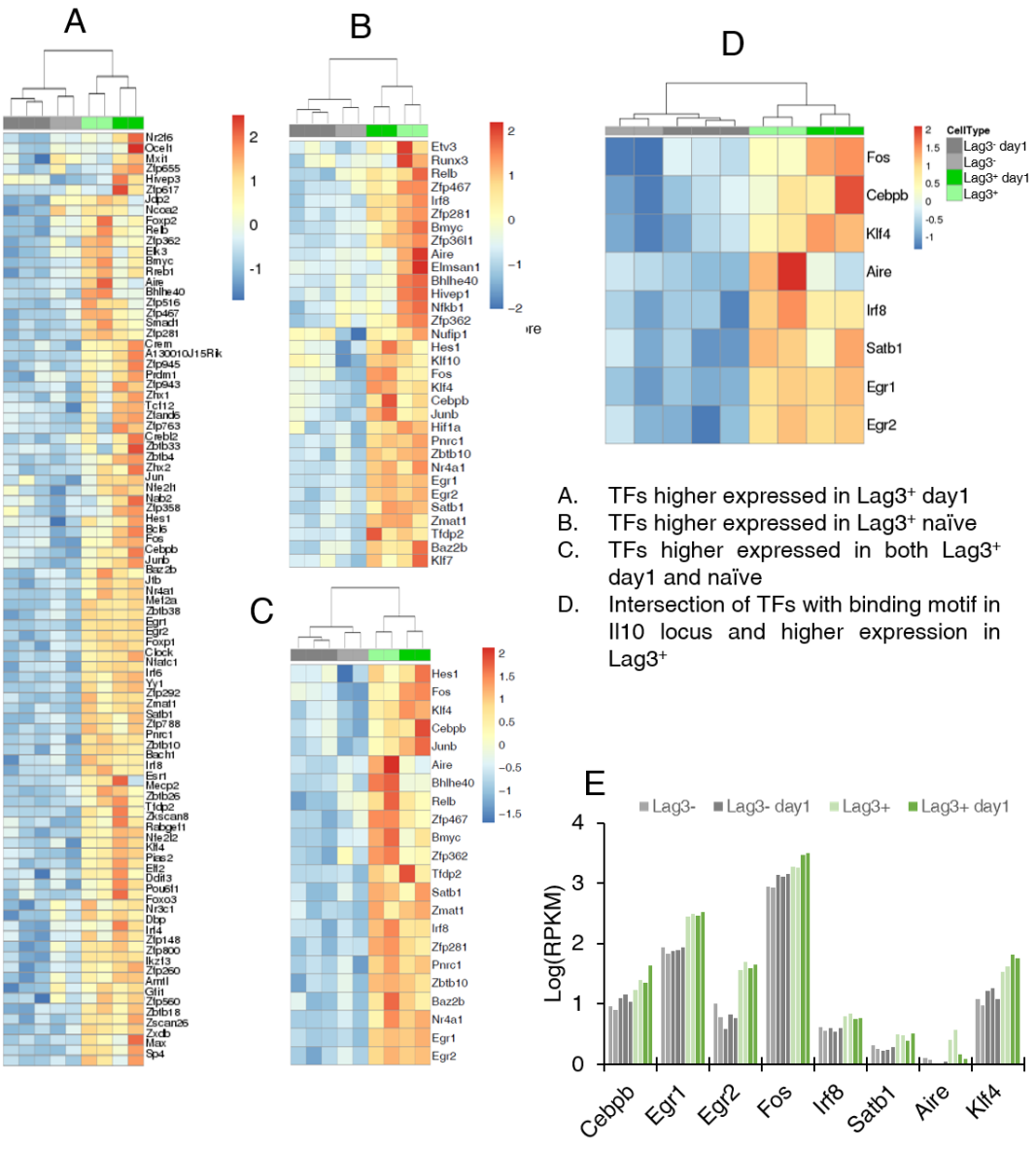


***Figure 3-32 Transcriptional control of programs in Lag3⁺ plasma cells. A**. Expression difference between Lag3⁺ and Lag3⁻ plasma cells (naïve and 1dpi) of transcription factors (TFs) that are higher expressed in Lag3⁺ plasma cells 1dpi. **B**. Expression difference of TFs that are higher expressed in naïve Lag3⁺ plasma cells. **C**. TFs that are higher expressed in naïve and 1dpi Lag3⁺ plasma cells compared to naïve and 1dpi Lag3⁻ plasma cells. **D**. Expression difference of TFs that are higher expressed in Lag3⁺ plasma cells (naïve and 1dpi) and with TF binding motifs in the Il10 locus. **E**. Expression levels of the TFs in D. Row z-scores were calculated using RPKM.*

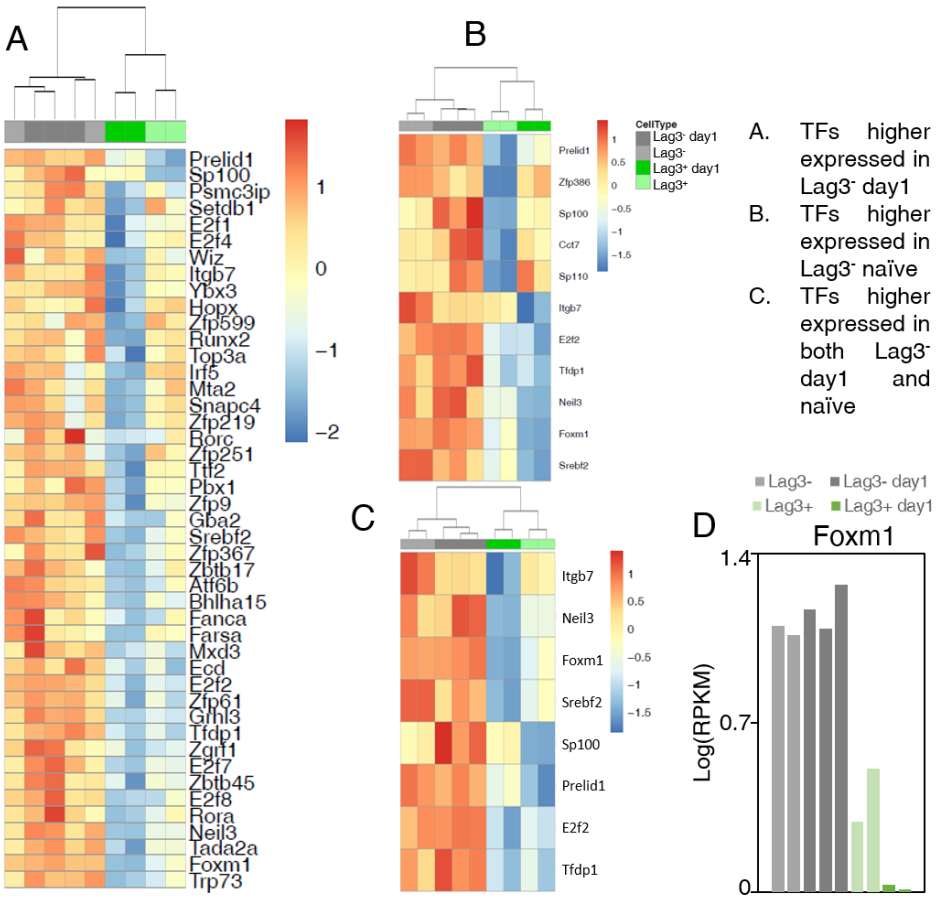


Figure 3-33 Transcriptional control of programs in Lag3⁺ plasma cells. ***A****. Expression difference between Lag3⁺ and Lag3⁻ plasma cells (naïve and 1dpi) of transcription factors (TFs) that are higher expressed in Lag3⁻ plasma cells 1dpi.* ***B****. Expression difference of TFs that are higher expressed in naïve Lag3-⁺ plasma cells.* ***C.*** *TFs that are higher expressed in naïve and 1dpi Lag3⁻ plasma cells compared to naïve and 1dpi Lag3⁺ plasma cells.* ***D****. Expression levels of Foxm1. Row z-scores were calculated using RPKM.*

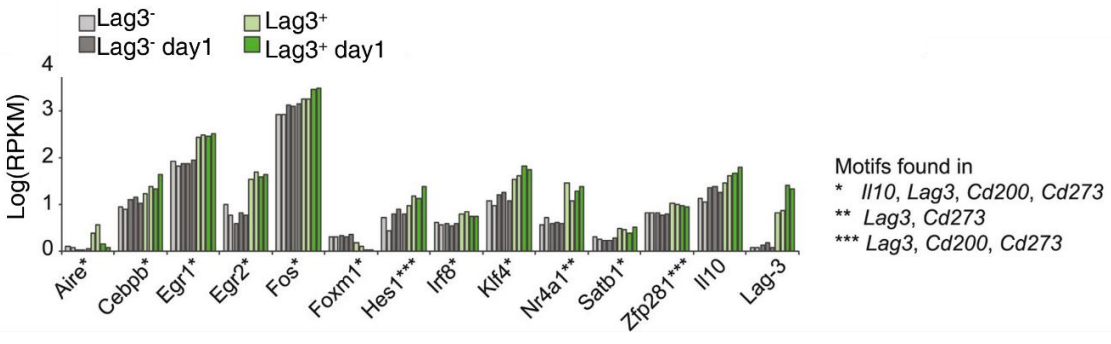


Figure 3-34 Expression levels of potential regulators of Lag3⁺ plasma cells. *Il10 and Lag3 are included as potential target genes of the represented transcription factors.*

3.2.7 Global DNA methylation profiles

To delineate the epigenetic status of Lag3$^+$ plasma cells and to relate them epigenetically to other B-cells and plasma cells the genome-wide DNA-methylomes of naïve Lag3$^+$ plasma cells, naïve Lag3$^-$ plasma cells, and peripheral B-cell subsets from naïve mice were compared. In addition, Lag3$^+$ plasma cells, Lag3$^-$ plasma cells, Il10$^+$CD138^{hi}, and Il10$^-$CD138^{hi} cells isolated at day 1 p.i. were included. RRBS on the sorted cells covered 3-4 million CpGs per sample and yielded a mean coverage above 10x. This assured a broad and representative DNA methylation profile (**Table 21**).

Table 21 Quality control of RRBS sequencing data of the analyzed B-cell subsets.

DNA methylation RRBS				
Cell type	**Number of replicates**	**Average read number per replicate**	**Average number of covered CpG sites**	**Mean coverage**
MZ	3	78539896	4307863	17
B2	3	56108154	3745166	17
T2	3	74758822	3920087	20
T1	3	81720486	4326240	18
FO	3	71351059	4325648	18
B1a	3	70952888	3630778	21
B1b	3	83341406	4223385	17
Lag3$^+$	3	68000054	4413438	14
Lag3$^-$	3	83785905	4562053	18
Lag3$^+$ day1	3	30233795	1658746	20
Lag3$^-$ day1	3	33263739	3033639	9
***Il10*eGFP$^+$ Plasma cells day1**	3	41321783	2743169	18
***Il10*eGFP$^-$ Plasma cells day1**	3	47658459	4063170	11

An unsupervised PCA separated B-cells from plasma cells on the major component PC1 (**Figure 3-35A**). Furthermore, the position of B1 cells on PC1 suggested similarity of B1 cells to plasma cells (B1a > B1b). PC2 split $Lag3^+$ and $Il10^+$ plasma cells from $Lag3^-$ and $Il10^-$ plasma cells, as well as from other B-cell subsets, and displayed the closest proximity to B1 cells. The relative cluster position of $Lag3^+$ and $Lag3^-$ plasma cells did not change for cells derived from day 1 p.i. as compared to naïve mice. Of note, $Il10^+$ and $Il10^-$ plasma cells co-localized with $Lag3^+$ and $Lag3^-$, respectively, underlining the expression of Lag3 as valid marker for Il10 producing plasma cells.

A striking difference in the global DNA methylation level was observed across the B-cell subsets and plasma cells. In line with a previous study in human, plasma cells displayed the lowest degree of genome wide methylation among all analyzed B-cell populations (Kulis et al. 2015) (**Figure 3-35B**). Loss of methylation predominantly occurred in partially methylated domains (PMDs, see chapter 1.1.2 on page 15) (**Figure 3-35C**). PMDs usually cover up to 75 % of the entire genome (Salhab et al. 2018). Hence, PMDs were responsible for the observed low global DNA methylation in plasma cells. Progressive loss of DNA methylation in PMDs is a hallmark of cell differentiation and is linked to an increase in heterochromatic marks along with a decrease of gene expression (Salhab et al. 2018). This is reflected in the DNA methylation loss in the analyzed plasma cells and the relatively low number of detected genes (**Table 19**). Durek et al. 2016 observed a progressive and proliferation-associated global loss of DNA methylation in PMDs during the differentiation of memory T-cells, indicating a shared mechanism of epigenome and genome organization during lymphocyte development.

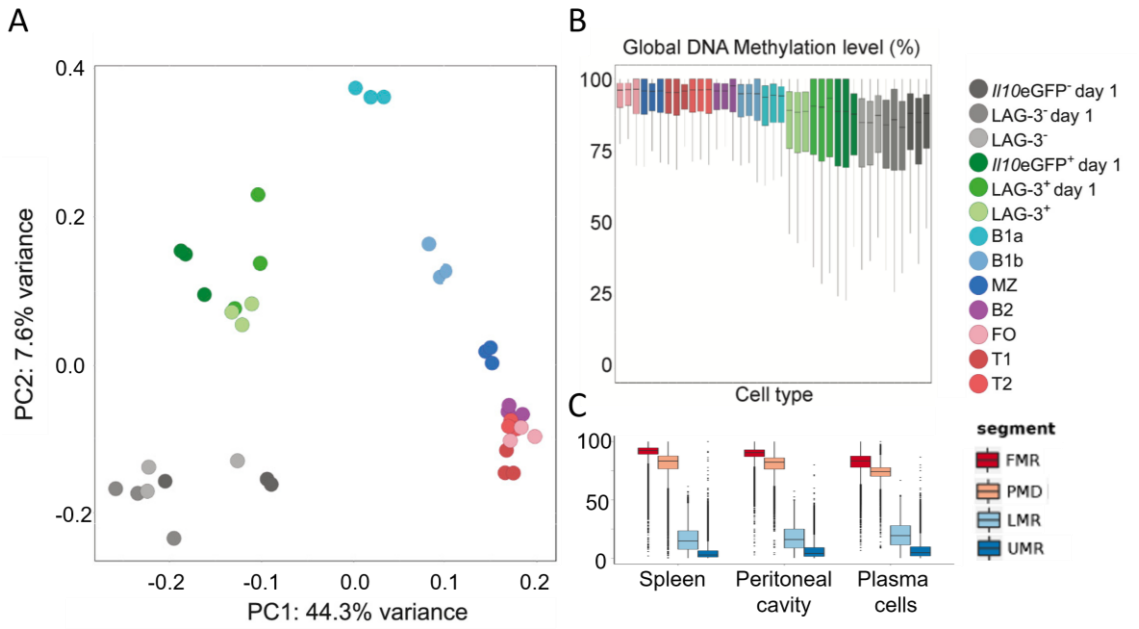


Figure 3-35 Overview of the global DNA methylation in distinct B-cell subsets. A*. Unsupervised PCA of genome-wide DNA methylation data based on single CpG levels.* ***B****. Average DNA methylation level in the analyzed B-cell subsets.* ***C****. Average DNA methylation level of genomic segments identified by the hidden markov model for B-cells isolated from the spleen, peritoneal cavity and for plasma cells.*

3.2.8 Differentially methylated regions reveal intimate relationship of Lag3$^+$ plasma cells to B1a cells

In search for epigenetic states contributing to the separation in PC2 (**Figure 3-35A**) the local DNA methylation differences between Lag3$^+$ and Lag3$^-$ plasma cells from naïve mice were determined. Using stringent filtering methods (see chapter 2.2.10 on page 53) 469 predominantly gene-associated DMRs distributed across the genome were identified (**Figure 3-36A, B**). These DMRs not only discriminated naïve Lag3$^+$ plasma cells from naïve Lag3$^-$ plasma cells, but also separated them in a hierarchical clustering along with B1 cells from the other B-cell subsets (**Figure 3-36A**). Intriguingly, the methylation state of a set of these DMRs showed the highest similarity of Lag3$^+$ to B1a cells (**Figure 3-36A, D**). This similarity was the strongest in hypomethylated regions, however, was also observed in hypermethylated regions (**Figure 3-36C**). Including the data on plasma cells derived from infected mice revealed the intimate relationship of naïve Lag3$^+$, day1 Lag3$^+$, Il10$^+$, and B1a cells (**Figure 3-36D**).

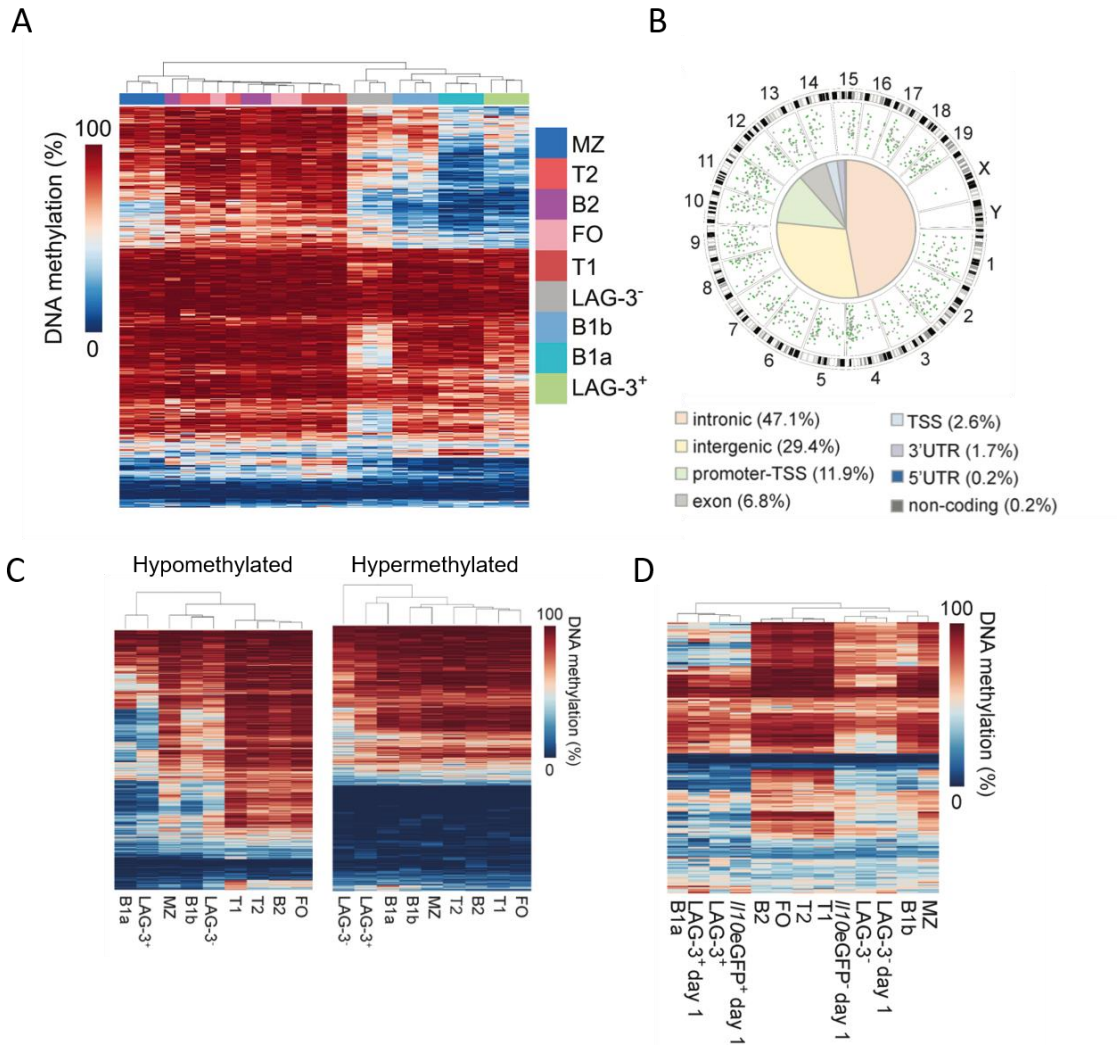


Figure 3-36 Differentially methylated regions discriminating naive Lag3⁺ and naive Lag3⁻ plasma cells. ***A***. *Heatmap representing 469 DMRs identified between naïve Lag3⁺ and naïve Lag3⁻ plasma cells.* ***B***. *Genomic distribution of the 469 DMRs with green dots representing the genomic position of hypomethylated regions in naïve Lag3⁺ plasma cells and grey dots representing hypomethylated region in naïve Lag3⁻ plasma cells.* ***C***. *Heatmaps showing the 469 DMRs separated into hypomethylated and hypermethylated regions in naïve Lag3⁺ plasma cells.* ***D***. *Heatmap representing the mean methylation status of the 469 DMRs including the methylation state of infected mice.*

Next, the discriminators between Lag3$^+$ and Lag3$^-$ of infected mice were identified by a pairwise DNA methylation comparison. The search for DMRs between day1 Lag3$^+$ and day1 Lag3$^-$ resulted in 359 regions (**Figure 3-37A**). The majority of these DMRs (n=291) were higher methylated in day1 Lag3$^+$ plasma cells compared to day1 Lag3$^-$ plasma cells. Inspecting the methylation level of these regions, a generally higher methylation was observed in day1 Lag3$^+$ compared to naïve Lag3$^+$ and Il10$^+$ plasma cells, while B1a showed the same methylation hypermethylation tendency. The search for DMRs between Il10$^+$ and Il10$^-$ plasma cells resulted in 369 regions (**Figure 3-37B**). In this case, the majority of the DMRs (n=335) were hypomethylated in Il10$^+$ plasma cells compared to Il10$^-$ plasma cells. Similar to the previously identified DMRs of naïve Lag3$^+$ and naïve Lag3$^-$ plasma cells (**Figure 3-36D**, **Figure 3-37C**) the methylation level of those regions correlated well with naïve Lag3$^+$, day1 Lag3$^+$ and B1a cells.

These comparisons revealed a preferential epigenetic relationship of Lag3$^+$ plasma cells, Il10$^+$ plasma cells to B1a cells. Interestingly, the majority of the identified DMRs were exclusively found in the respective comparison (**Figure 3-37D**), indicating activation state dependent local changes. Only three regions were found to be differentially methylated in all three comparisons. These DMRs were annotated to Biotinidase gene *Btd* (last exon), Sialyltransferase gene *St3gal4* (intron 1), and *Tbc1d23* (30 kb upstream of TSS). *Tbc1d23* was shown to be a general inhibitor of innate immunity signaling, affecting toll-like-receptor signaling (De Arras et al. 2012). Along with the methylation difference, mRNA-Seq data showed a slightly higher expression of *Tbc1d23* in Lag3$^+$ plasma cell (mean RPKM: day1 Lag3$^-$ 3.49; naive Lag3$^-$ 3.90; day1 Lag3$^+$ 4.84, naive Lag3$^+$ 4.55).

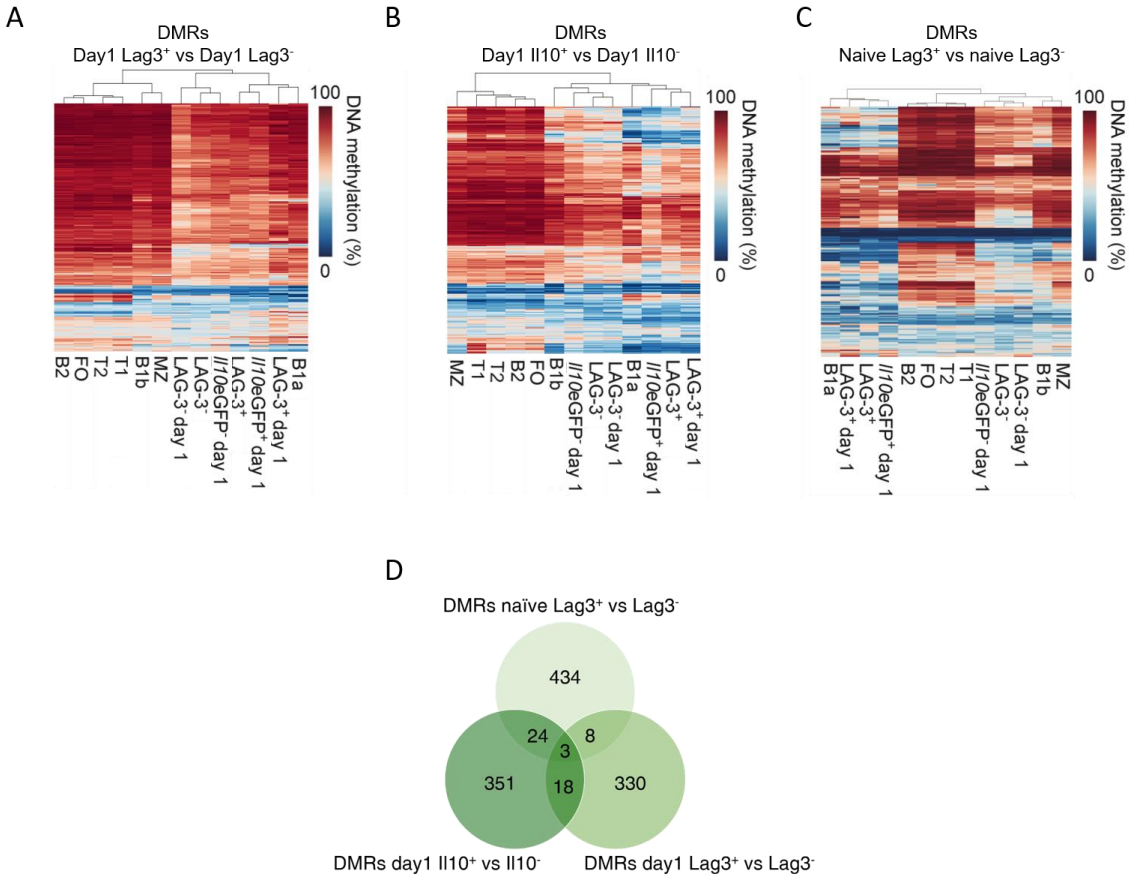


***Figure 3-37 Differentially methylated regions in Lag3⁺ and Il10⁺ plasma cells derived from infected mice. A**. Heatmap representing 359 DMRs identified between day1 Lag3⁺ and day1 Lag3⁻ plasma cells. **B**. Heatmap representing 369 DMRs identified between Il10+ and Il10⁻ plasma cells. **C**. Heatmap representing 469 DMRs identified between naïve Lag3⁺ and naïve Lag3⁻ plasma cells. **D**. Overlap of DMRs resulted from the respective comparisons. Mean of the methylation is presented in A-C.*

3.2.9 Epigenetic state of Il10 locus

A particular local DNA methylation difference was found at the *Il10* locus (40 kb upstream of TSS and 30 kb downstream of TTS) and further highlighted the closer epigenetic relationship of Lag3$^+$ plasma cells (naïve and day1) and B1a cells. The overall degree of DNA methylation at this locus was lowest in Il10$^+$ plasma cells, Lag3$^+$ plasma cells (naïve and day1), and B1a cells (**Figure 3-38A**), a hierarchy not following the genome-wide level where Lag3$^-$ and Il10$^-$ plasma cells displayed the lowest degree of DNA methylation (**Figure 3-35B**). Furthermore, Lag3$^+$ and Il10$^+$ plasma cells shared distinct patterns of hypomethylated regions around the *Il10* gene that overlapped with DNase I hypersensitivity sites indicating open chromatin in B-cells (**Figure 3-38B**). These hypomethylated and open regions flanked the *Il10* locus up to 50 kb, and were thus likely to demarcate regulatory regions. Taken together, the epigenetic signatures at the *Il10* locus indicated an epigenetic "preprogrammed" permissive state of *Il10* expression in Il10$^+$ plasma cells, Lag3$^+$ plasma cells (naïve and day1), and B1a cells.

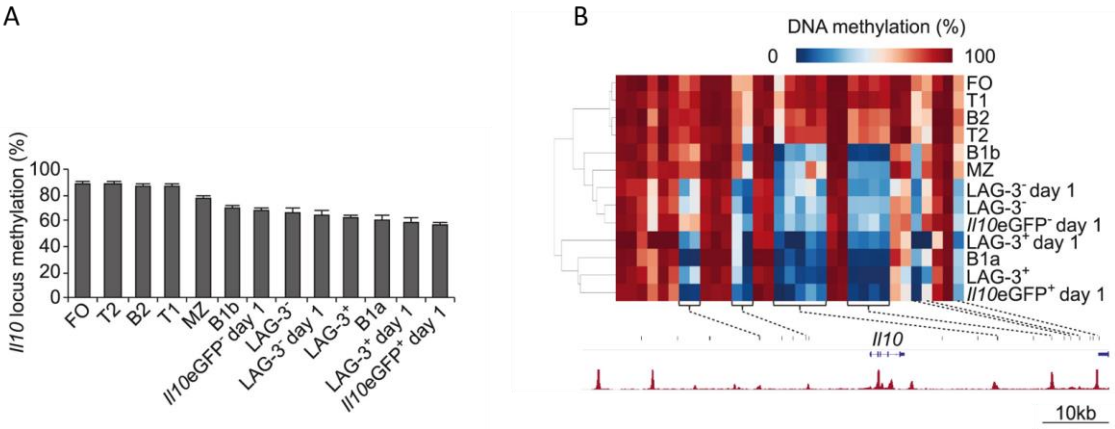


***Figure 3-38 Epigenetic state of Il10 locus. A**. DNA methylation of Il10 locus (mean ± SEM). **B**. Methylation for covered CpG in the Il10 locus. Coverage weighted average methylation of three replicates is represented. The positions of selected CpG are indicated by vertical black bars, with the Il10 gene locus depicted. ENCODE B-cell DNase I data is represented in the red track.*

3.2.10 Integrated analysis highlights epigenetic control of crucial factors

When selecting for genes exhibiting both differential expression and methylation between naïve Lag3$^+$ and naïve Lag3$^-$ plasma cells, 23 genes were identified including *Il10*, *Bcl2*, *Irf8* and other known B-cell regulators, such as the NF-κB inhibitor *Nfkbiz*, as well as the apoptosis related nuclear factor *Nr4a1* (**Figure 3-39A**). Of particular interest, an epigenetic control of mRNA binding factors such as *Zfp36l1*, *Pdcd4*, and *Ptbp2* was observed. The posttranscriptional regulation of Il10 is focus of current studies since *Il10* mRNA is constitutively transcribed in many cells but the protein is not detected in all of these cells (as the case for naïve Lag3$^+$ plasma cells and day1/naïve Lag3$^-$ plasma cells; **Figure 3-39B**). It has been shown that AU-rich elements (ARE) in the 3' UTR of *Il10* lead to the degradation of its mRNA and in turn to variation of the protein levels (Powell et al. 2000). Zfp36l1 has been shown to be a factor that destabilize several cytoplasmic ARE-containing mRNA transcripts (Blackshear et al. 2003; Hudson et al. 2004). The expression of *Zfpl36l1* decreased in Lag3$^+$ plasma cells after infection, which in turn could lead to translation of Il10 and emphasizes Zfp36l1 as an interesting candidate in posttranscriptional Il10 regulation (**Figure 3-39B**). However, the lower expression level of *Zfp36l1*, as well as the other mRNA binding factors *Pdcd4*, and *Ptbp2* in Lag3$^-$ plasma cells points to a complex regulation of Il10 that yet has to be unraveled.

The integration of DNA methylation and expression of the associated gene further displayed an epigenetic control of chromatin modifiers *Elmsan1, Jade2*, and the Hdac1 down-regulator *Kctd6*. The function of these chromatin organizers is yet not clearly defined, and the implicated distinct chromatin modifications have to be investigated in the future.

In conclusion, the observed relationship between gene hypomethylation and elevated expression for *Il10* and other factors argued for an epigenetically primed control of the expression of these genes.

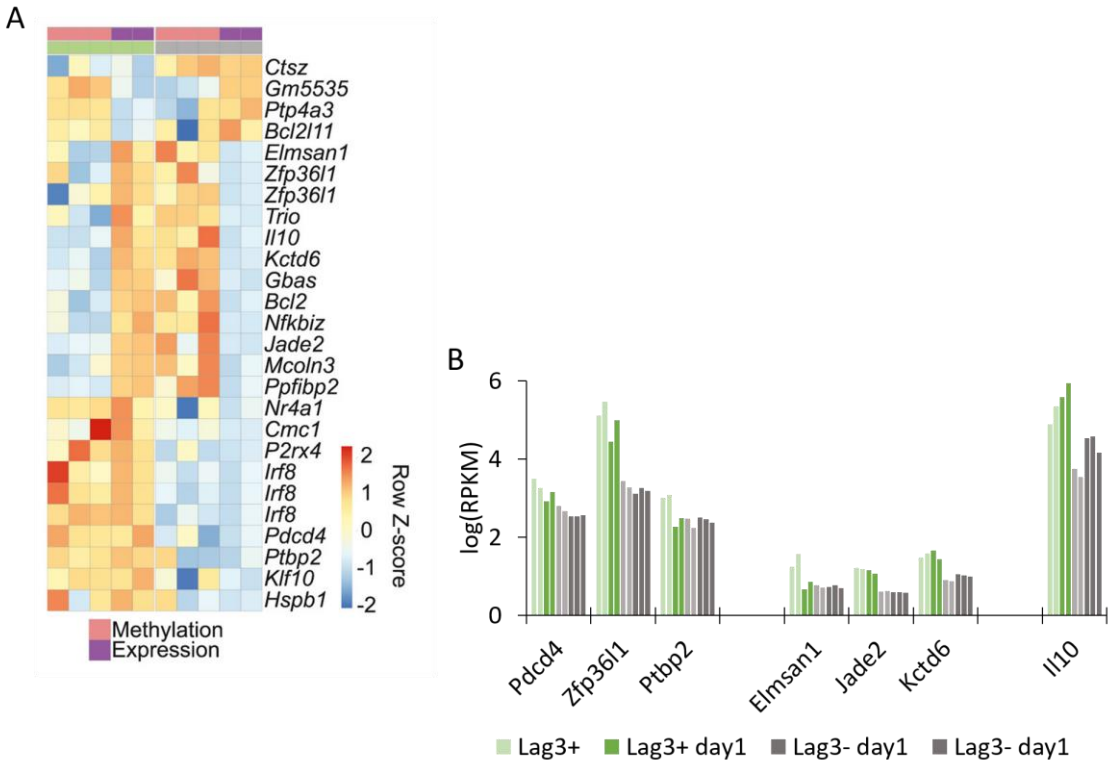


Figure 3-39 Integrated analysis of differential expression and differential DNA methylation in naïve plasma cell subtypes. A*. Relative changes in DNA methylation and expression of respective gene. Row z-scores were calculated using RPKM and DNA methylation in %.* ***B****. Gene expression levels of mRNA binding factors, chromatin binding factors and Il10.*

3.2.11 Discussion

B-cells are classically known as one of the mediators of humoral immunity. It is commonly considered that B-cells differentiate into antibody-secreting cells, known as plasma cells, and thereby exert immune responses against pathogens and maintenance of immune homeostasis. Their broad implication in cellular immunity was discovered and investigated intensively in the past decades. B-cells contribute to T-cell activation and expansion by antigen-presentation (Townsend and Goodnow 1998; Bouaziz et al. 2007). They release a variety of cytokines modulating $CD4^+$ T-cell functions (F. E. Lund 2008) and are capable to expand $Foxp3^+$ Tregs by releasing TGF-β (Shah and Qiao 2008). The activation and expansion of Tregs is only one mechanism by which B-cells participate in the regulation of cellular and humoral immune responses to pathogens or self-antigens. Over the past 10 years, the regulatory function of B-cells has been studied intensively and was demonstrated in animal models of e.g. rheumatoid arthritis, and systemic lupus erythematosus (SLE) (Carter et al. 2011; Fillatreau et al. 2002). These studies led to a wide acceptance of the concept of Bregs as negative regulators of the immune system to prevent a pathological autoreactive reaction and protecting from uncontrolled inflammation. Further investigation on the mechanisms of action revealed that Bregs exert the immunosuppressive function primarily via the production of the anti-inflammatory cytokine Interleukin-10 (Il10) (Fillatreau et al. 2002; Mauri 2010). The Breg-derived Il10 inhibits the proliferation of T-helper-cells, affects the expression of pro-inflammatory cytokines by innate immune cells, and alters expression of molecules involved in antigen presentation (Couper, Blount, and Riley 2008). Il10 expressing B-cells can be detected in the spleen of *Salmonella typhimurium* infected mice within a few hours after challenge (Neves et al. 2010). Interestingly, these Il10 producing B-cells were identified among $CD138^{hi}$ plasmocytes. Further characterization of $Il10^+CD138^{hi}$ cells revealed the expression of the inhibitory receptor lymphocyte-activation gene 3 (Lag3) as a discriminative surface marker on these cells (Lino et al. 2018). However, their precursor in the naïve immune system and mechanisms of action are still not completely identified. In order to delineate the molecular properties of $Lag3^+$ plasma cells, epigenetic and transcriptomic signatures of $Lag3^+$ and $Lag3^-$ plasma cells isolated from naïve, as well as *Salmonella* infected mice

(1dpi) were systematically compared in this thesis. Furthermore, DNA methylomes of Lag3$^+$, Lag3$^+$ 1dpi, Il10$^+$ 1dpi, Lag3$^-$, Lag3$^-$ 1dpi, Il10$^-$ 1dpi plasma cells, as well as immature and mature B-cells were generated and compared in order to assess the possible origin of Lag3$^+$ plasma cells.

Epigenetic similitude between Lag3$^+$ plasma cells and Il10$^+$ plasma cells reflects their close developmental relation

Similar expression profiles of B-cell receptor/antibody components between Lag3$^+$ cells from naïve mice and Lag3$^+$ cells from infected mice corroborated their close relation. Moreover, DNA methylation as the hallmark of cell identity and developmental memory, was highly similar between Lag3$^+$ cells from naïve mice, Lag3$^+$ cells 1dpi, and Il10$^+$ cells 1dpi and differed from the methylomes of Lag3$^-$, Lag3$^-$ 1dpi, Il10$^-$ 1dpi plasma cells and other B-cell subsets. Noteworthy, naïve Lag3$^+$, Lag3$^+$ 1dpi, and Il10$^+$ 1dpi displayed the lowest DNA methylation of putative regulatory regions at the *Il10* locus, reflecting a local regulatory adaptation to rapidly activate *Il10* expression. These results validated Lag3 as marker of Il10 producing Bregs.

Transcriptome profiles reveal non-proliferating status of Lag3$^+$ plasma cells

The differentiation of antibody-secreting plasma cells is associated with B-cell proliferation, epigenome remodeling, the expression of plasma cell-specific transcription factors, and the expression of cell division related proteins, and usually takes up to 3 days (Nutt et al. 2015). The transcriptome analysis of Lag3$^+$ and Lag3$^-$ plasma cells revealed a decreased expression of genes involved in cell cycle regulation in Lag3$^+$ plasma cells. Moreover, the lowest global DNA methylation level was observed in Lag3$^-$ plasma cells, pointing to a high proliferation rate of these cells. These results suggest a non-proliferating state of Lag3$^+$ plasma cells, whereas Lag3$^-$ cell population comprise proliferating plasmablasts. Intriguingly, the non-proliferating state of Lag3$^+$ cells 24h after infection implicates the pre-existence of these cells in the naïve mice rather than the activation-dependent conversion of B-cells into plasma cells. This characterizes Lag3$^+$ plasma cells as natural regulatory plasma cells. The higher expression of genes involved in chromatin organization in Lag3$^+$ cells 1dpi compared to

Lag3$^-$ 1dpi suggests a remodeling of the chromatin as mechanism to adapt gene expression as a proliferation-independent activation response.

Lag3$^+$ plasma cells display features reminiscent of B1a cells and may be involved in the removal of damaged cells

The principal component analysis of the global DNA methylation of all analyzed cell types revealed B1a cells to have the closest relation to Lag3$^+$ (naïve and 1di) and Il10$^+$ 1dpi plasma cells. Moreover, the signatures of local DNA methylation differences that distinguished Lag3$^+$ and Il10$^+$ from Lag3$^-$ and Il10$^-$ plasma cells showed striking similarities to B1a cells. Of all analyzed B-cell subsets, B1a cells showed the lowest degree of DNA methylation at regulatory regions at the *Il10* locus, suggesting a primed state for *Il10* expression in these cells.

The transcriptome analysis revealed a distinct BCR repertoire of Lag3$^+$ plasma cells compared to Lag3$^-$ plasma cells. Of particular interest, Lag3$^+$ plasma cells (naïve and 1dpi) expressed *Igkv14-126* at significantly higher level than Lag3$^-$ plasma cells (naïve and 1dpi). This gene encodes the immunoglobulin kappa chain variable 14-126 which is a component of the BCR typically found on B1a cells and which recognizes the phosphatidylcholine (PtC) that is present in plasma membranes of host cells (Hardy, Wei, and Hayakawa 2004). The expression of this BCR by B1a cells is part of their innate-like function to provide first line of defense against pathogens and auto-antigens (Baumgarth 2010). The expression of antibodies of this type can be detected in germfree organisms and are so-called natural antibodies which are produced by plasma cells derived from activated B1a cells (Bos et al. 1989; Haury et al. 1997; Berland and Wortis 2002). Accordingly, the expression of this BCR/antibody component by Lag3$^+$ plasma cells links their development to the recognition of auto-antigens such as damaged and apoptotic cells, as well as aged red blood cells. This hypothesis is further extended by the observation that the apoptosis related nuclear factor *Nr4a1* is differentially methylated in Lag3$^+$ and Lag3$^-$ plasma cells and higher expressed in Lag3$^+$ plasma cells. Nr4a1 was shown to mediate anti-inflammatory properties of macrophages after the phagocytosis of apoptotic cells to ensures the nonimmunogenic clearance of dying cells (Ipseiz et al. 2014). This Nr4a1-dependent process was

mediated by the inhibition of NF-κB signaling. In line with this, the NF-κB inhibitor *Nfkbiz* was also differentially methylated and higher expressed in Lag3$^+$ plasma cells. In addition, Lag3$^+$ plasma cells expressed *Sirpa*, encoding the immunoglobulin-like cell surface receptor for the so called "don't eat me'" molecule CD47, at a higher level than Lag3$^-$ plasma cells (Appendix **Table 23**) (Bian et al. 2016, 47). This enables Lag3$^+$ plasma cells to distinguish between apoptotic and healthy cells and associates them with the clearance of apoptotic cells.

The expression of *Enpp1* by Lag3$^+$ plasma cells after *Salmonella* infection furthermore underlines the colse relation to B1a cells. This gene encodes the plasma cell alloantigen 1 (PC1) protein and was shown to be co-expressed with Il10 by a subset of B1a cells which mediated fetus-tolerance during pregnancy (H. Wang et al. 2012; Schumacher et al. 2018).

Lag3$^+$ plasma cells exert immunosuppressive function by the co-expression of multiple inhibitory molecules

Lag3$^+$ plasma cells distinctly expressed genes coding for several molecules implicated in the negative regulation of immunity such as Lag3, PD-L2, CD200, and Tbc1d23. Lag3 was shown to play immunoregulatory roles on activated CD4$^+$ T-cells (Maruhashi et al. 2018). The PD-1 ligand PD-L2 dramatically inhibits cytokine production by CD4+ T-cells and their proliferation (Latchman et al. 2001). CD200 is a membrane protein expressed by a broad range of cell types, including B-cells, T-cells, endothelial cells, and neurons. Its interaction with the CD200 receptor, expressed by myeloid cells, mediates an immunological inhibitory signal of myeloid cells (Barclay et al. 2002). Tbc1d23 affects toll-like-receptor signaling leading to an inhibitory effect of innate immune cells (De Arras et al. 2012). The synergistically activity of a variety of molecules involved in immunoregulatory mechanisms define more precisely Lag3$^+$ as a regulatory plasma cell subset.

Interplay of multiple mechanisms result in the regulation of Il10 and Lag3 expression in Lag3$^+$ plasma cells

The co-expression of multiple inhibitory molecules together with Il10 raised the question for the transcriptional control of these genes. The transcription factor genes

Aire, *Fos*, *Junb*, *Egr1*, *Egr2*, *Cebpb*, *Irf8*, *and Satb1* were found to be higher expressed by Lag3$^+$ (naïve and 1dpi) plasma cells compared to Lag3$^-$ plasma cells. In addition, these factors had binding sites at the *Il10*, *Lag3*, *CD200*, and *Pdcd1lg2* (PD-L2) loci, emphasizing an important role in the expression of these genes. The transcriptional repressor gene *Cbfa2t3* was found to be downregulated in Lag3$^+$ after infection which in turn activates the transcription of the target genes. However, the specific function of these factors in Lag3$^+$ plasma cells has to be studied by performing B-cell-specific knock outs or over expression experiments in order to validate the predicted transcriptional control.

Noteworthy, the transcription factor Egr2 activates the co-expression of *Lag3* and *Il10* in CD4$^+$CD25$^-$Foxp3$^-$regulatory T-cells (Okamura et al. 2009, 4). This indicates a shared program between Tregs and Bregs which controls the anti-inflammatory function.

Il10 mRNA is expressed by a broad range of cell types. However, the protein is produced only by some cell types. The regulation of *Il10* expression in different immune cells is achieved by a complex interplay of signal transduction, epigenetics, transcription factor binding and posttranscriptional modifications (Saraiva et al. 2005). The DNA methylation data and transcriptome data on Lag3$^+$ and Lag3$^-$ plasma cells revealed a differential regulation of the mRNA decay activator *Zfp36l1*. Together with the methylation signatures of the Il10 locus, and the predicted binding of certain transcription factors, the observed downregulation of *Zfp36l1* in Lag3$^+$ plasma cells after infection extends the potential mechanisms of Il10 regulation in Lag3$^+$ plasma cells.

Overall conclusion

In conclusion, Lag3$^+$ plasma cells possess an epigenome status of mature plasma cells yet display a distinct epigenetic landscape with features reminiscent of B1a cells. Furthermore, the methylome analyses supports the hypothesis that Lag3$^+$ plasma cells in naïve mice are the primed precursors of Il10-expressing Lag3$^+$ plasma cells found on day 1 p.i., supporting the hypothesis of pre-existing natural Bregs. The combinatorial analysis of DNA methylation and gene expression revealed epigenetic programs that

controlled expression of key effector genes (such as *Il10*) and regulatory transcription factors (*Egr2*, *Fos*, *Foxm1*). Further studies shall aim at the function of the identified potential regulators of Il10 as well as other checkpoint inhibitors.

Chapter 4.

Overall Discussion and Outlook

One single fertilized egg can give rise to a multitude of distinct cell types forming a complex mammalian organism. How this fascinating process is achieved is a question of high interest and its answer is encoded in the epigenome. Precise temporal and spatial orchestration of gene expression by epigenetic mechanisms is required to establish functional cellular diversity (see chapter 1.1.1 on page 3). Transcriptional regulation shows the highest dependency on the binding of TFs to regulatory elements such as promoters, enhancers, and repressors which together form complex regulatory landscapes in the non-coding part of the genome (Vaquerizas et al. 2009). Through the plethora of epigenetic studies, especially in the case of DNA methylation studies, the promoter-centric dogma assuming an anti-correlation with gene expression shifted towards a more regulatory perspective in which epigenetic state of distal elements, such as enhancers, are closely linked to gene expression in development and disease (Carullo and Day 2019; Schübeler 2015; Lister et al. 2009).

This thesis dissected the molecular diversity of two highly distinct cell type compartments. On the on hand, astrocytes from two distinct adult brain regions (cerebral cortex and cerebellum) were analyzed by an integrative genome wide approach using comprehensive epigenomic and gene expression profiling. Astrocytes represent a tissue resident cell type that derives from the same progenitor cell which undergoes spatial specification by extrinsic factors (see chapter 3.1.10 on page 110) (Bayraktar et al. 2015). Furthermore, mature astrocytes differ in their morphology and functionality depending on the distinct neuronal environment in the respective brain region (see chapter 3.1.10 on page 111) (W. Todd Farmer and Murai 2017). B-cells and plasma cells on the other hand, represent morphologically homogenous, non-resident,

circulating cell types that exert their pre-specified and specialized function upon various extrinsic signals, and derive from distinct progenitors (LeBien and Tedder 2008).

Epigenetic signatures provide information on developmental origins

Despite the distinct characteristics of these cell types, similar conclusions and interpretations can be drawn from the epigenomic analyses. For instance, local DNA methylation provided information on the developmental origin of the respective cell (sub)type and a potential future response. In the case of $Lag3^+$ and $Il10^+$ plasma cells, methylation levels of DMRs discriminating these cells from $Lag3^-$ and $Il10^-$ plasma cells were highly similar in B1a cells. In addition, the methylation pattern at the *Il10* locus was highly similar between Il10-producing plasma cells and B1a cells. Together with other observation, these results revealed an intimate relation of $Lag3^+$/$Il10^+$ plasma cells to B1a cells and supports the hypothesis that B1a cells contribute as progenitor cells to the $Lag3^+$/$Il10^+$ plasma cell population. Moreover, the strong hypomethylation at the *Il10* locus suggests a priming of B1a cells and $Lag3^+$ for *Il10* expression. The DNA methylation profiles of astrocytes also shed light on the developmental history of these cells even despite the lack of data on potential progenitors. A high number of transcription factors involved in brain cell development were differentially methylated. Noteworthy, differential methylation was not predominantly associated with differential expression (**Figure 3-17**), but rather represented long-lasting epigenomic signature of transient expression during cell development (Mo et al. 2015). Mo et al. identified hypermethylated regions flanking TF genes that are transiently expressed and shape neuronal subtype identity, while the respective gene body remained unmethylated (Mo et al. 2015). Such a signature can be observed at the *Zic1/Zic4* locus, where the gene bodies are unmethylated in both astroglia populations and the flanking regions show hypermethylation in the *Zic1/Zic4* non-expressing CTX astrocytes (**Figure 3-21**). This methylation pattern is accompanied by closed chromatin in CTX astrocytes ensuring inactivated expression despite the absence of DNA methylation. On the on hand, this demonstrates one possible approach how DNA methylation of adult cells captures the developmental history, and on the other hand, this example remarkably demonstrates the benefit of integrating chromatin accessibility to infer gene expression. Furthermore, hypomethylated regions in the cortical astrocytes were

enriched for binding sites of Lhx2, a TF broadly expressed in the cortex and involved in cortical neurogenesis (Godbole et al. 2018). This observation is indirectly supporting the hypothesis, that neurons and astrocytes maintain the transcriptional program established in common regional pre-specified progenitors in order to imprint a region-specific functionality. This interpretation is further supported by the region-specific open chromatin regions and expression of respective region-specific TFs and was discussed in chapter 3.1.10. This demonstrates how detangling epigenomic data generated from terminally differentiated cells can provide information on cell development without the direct comparison with epigenomic data of the progenitor cell.

Distal regulatory regions determine cell identity

In both cases, astrocytes and B-cells, DNA methylation differences that delineated the respective cell subtypes were predominantly identified at intergenic and intronic regions. This extends the current discussions that epigenetic states at distal *cis* elements rather than promoter regions determine the cell identity, as was demonstrated in e.g. stem cells and neuronal progenitors, with evidence in astroglia subpopulations and plasma cell subpopulations (Stadler et al. 2011). However, caution has to be taken since the distribution of DMRs identified by using RRBS is biased due to the usage of a sequence-specific endonuclease and a WGBS approach would provide more robust conclusions (**Figure 1-7**). Nevertheless, *HaeIII*-RRBS provides a valid approximation to WGBS and the observed predominant enrichment of astrocyte DARs at intergenic and intronic regions affirms the conclusion of enhancer-driven cell diversity.

Concluding remarks on integrative analyses

Integrating multi-omics data paves the way for a more refined understanding of transcriptional regulation. However, the integration is challenging due to various factors that will be discussed in the following. Firstly, regulatory elements need to be defined and identified and secondly, they need to be associated with the right target gene. The identification of regulatory regions can be addressed in different ways. For instance, by defining enhancers via the co-occurrence of H3K27ac, H3K4me1 and

absence of H3K4me3, via binding of enhancer associated protein p300, or via TF binding prediction using DNA accessibility data (Schmidt et al. 2017; Visel et al. 2009; Ernst et al. 2011). ChIP Seq data as well as DNA accessibility data are usually processed using peak calling algorithms to determine the presence or absence of the respective enrichment. Peak calling, however, is not trivial and has limitations by setting thresholds for signal-to-noise ratios (Koohy et al. 2014). Furthermore, technical reasons such as cell number can limit an accurate detection of enriched regions (Gilfillan et al. 2012). The binary output from the peak caller can further be analyzed by ChromHMM using a multivariate hidden Markov model to identify chromatin-state signatures (Ernst and Kellis 2017). However, the focus on the qualitative signal limits the data interpretation by excluding quantitative differences.

Various distinct approaches have been proposed to assign putative regulatory regions to their target genes. These different approaches are based on biological evidence that regulatory elements affect either the gene with the lowest genomic distance or influence the expression of a gene located at great genomic distances of up to hundreds of kilobases by loop formation (Rada-Iglesias et al. 2011; G. Li et al. 2012). Methods based on physical interactions of genomic regions such as Hi-C and ChIA-PET enable the mapping of direct interactions between enhancers and promoters of the target gene (Fullwood and Ruan 2009; Yao, Berman, and Farnham 2015). A different approach is to use correlation-based methods to link enhancer state, defined by DNase I-Seq or histone modifications, with the expression of nearby genes (Ernst and Kellis 2017; Yao, Berman, and Farnham 2015).

As pointed out above, there is no golden standard in integrative multi-omics analyses. Several factors will influence the results and consequently the interpretations. It is therefore all the more important to clearly describe the approaches used to integrate epigenomic data with gene expression data in order to recognize potential limitations of the analysis. In this work, putative regulatory regions were defined by differential DNA methylation or differential DNA accessibility. This definition is prone to result in the identification of false positives since DNA methylation differences are not always associated with distinct TF binding, but e.g. can reflect distinct proliferation states. In addition, a distinction of 5-mC and 5-hmC is not possible

with the standard procedure of RRBS. Hence, hypermethylation might partially represent an enrichment of 5-hmC at enhancer regions which is currently discussed to be an intermediate during active DNA demethylation and in turn leads to an activation of the enhancer (Stroud et al. 2011; Hon et al. 2014). Further, a subset of regulatory regions was defined by the overlap of at least 1bp of DMRs and DARs. This reduces the accurate prediction of all regulatory regions since the DMR detection is based on RRBS. Nevertheless, one ends up with a good approximation of regulatory regions by using these simple approaches. Finally, the putative regulatory regions were assigned to the closest gene which does not take long-range interactions into account. Ideally, the predicted involvement of certain loci in transcriptional regulation of genes should be validated by assays such as luciferase reporter assay, or genome editing using the CRISPR/Cas9 system (Sander and Joung 2014).

In conclusion, comprehensive integrated analyses are just beginning to be systematically approached. A great effort and contribution to elucidate the regulatory roles of the epigenome and its influence on cell identity was done by the epigenomic consortia like ENCODE, Roadmap, DEEP and IHEC that provided a deluge of epigenomic data and standardized methods and analyses (The ENCODE Project Consortium 2012; Roadmap Epigenomics Consortium et al. 2015; http://www.deutsches-epigenom-programm.de/; http://ihec-epigenomes.org/). The data produced for this thesis extends the currently available murine epigenomic data and provide a valuable resource for future hypothesis-driven experiments.

4.1 Outlook

As pointed out in the previous discussion, the CRISPR/Cas9 editing enables highly specific targeting of regulators which holds the key to validation of the derived hypotheses on regulatory regions. In addition, the predicted activity of Lhx2 and Zic1/Zic4 needs to be validated. One possibility would be to determine functional consequences of an astrocyte-specific loss of function of Lhx2 or Zic1/Zic4 in the cortex or cerebellum, respectively. Another possible validation of their activity would be a Lhx2/Zic1/Zic4 immunoprecipitation to verify binding sites and identify target genes. Besides the validation of the results, it would also be interesting to perform similar

assays and analyses using different mouse models. The hGFAP-EGFP mouse model used in this thesis is known to have a higher recombination rate in perivascular astrocytes of the cortex and Bergmann glia of the cerebellum (Nolte et al. 2001). The use of Aldh1l1 reporter mice was shown to label astrocytes of the cortex more homogenously and would result in a slightly distinct subset of astrocytes (Winchenbach et al. 2016; Srinivasan et al. 2016). On the one hand, the comparison of various mouse models could reveal a common epigenetic network-specific for each brain region, and on the other it would emphasize a strong heterogeneity of astrocytes in the cortex.

The bulk astrocyte epigenomes and transcriptomes shed light on the basic differences of regional astrocytes and provided hypotheses on how they are established. Further disentanglement of the cell heterogeneity by single cell transcriptomic and epigenomic assays will pave the way to accurately characterize molecular properties of astrocytes residing in distinct brain regions.

In regard of the observed high expression of Egr2 in $Lag3^{+}$ plasma cells, and the link to the known activation of *Lag3* and *Il10* co-expression in $Foxp3^{-}$regulatory T-cells, it is necessary to investigate the question whether Il10 producing regulatory B-cells share common programs with regulatory T-cells. It will be exciting to test whether comparative methylome and transcriptome analyses of $Lag3^{+}$ Tregs and $Lag3^{+}$ Bregs will reveal a common epigenetic signature for immuno-regulatory function.

A. List of Abbreviations

5-caC	5-Carboxycytosine
5-fC	5-Formylcytosine
5-hmC	5-Hydroxymethylcytosine
5-mC	5-Methylcytosine
ATAC-Seq	Assay for Transposase Accessible Chromatin Sequencing
BCR	B-Cell Receptor
BG	Bergmann Glia
Bregs	Regulatory B-Cells
CB	Cerebellum
cDNA	Copy DNA
CGIs	CpG Islands
ChIP-Seq	Chromatin Immunoprecipitation Sequencing
CNS	Central Nervous System
cov	Coverage
cpm	Counts Per Million
CREs	Cis-Regulatory Elements
CTX	Cerebral Cortex
DARs	Differentially Accessible Regions
DEGs	Differentially Expressed Genes
DMRs	Differentially Methylated Regions
DNMTs	DNA-Methyltransferases
dpi	Days Post Infection
ES cell	Embryonic Stem Cell
FACS	Fluorescence Activated Cell Sorting
FC	Fold Change
FMRs	Fully Methylated Regions
FO	Follicular B-Cells
GO	Gene Ontology
H3K27ac	Histone 3 Lysine 27 acetylation

H3K27me3	Histone 3 Lysine 27 trimethylation
H3K4me1	Histone 3 Lysine 4 monomethylation
H3K4me2	Histone 3 Lysine 4 dimethylation
H3K4me3	Histone 3 Lysine 4 trimethylation
HAT	Histone Acetyltransferase
HDACs	Histone Deacetylases
HDMs	Histone Demethylases
HMTs	Histone Methyltransferases
HSC	Hematopoetic Stem Cells
Il	Interleukin
Il10	Interleukin-10
K	Lysine
Lag3	Lymphocyte-Activation Gene 3
LMRs	Lowly Methylated Regions
lncRNA	Long Non-Coding RNA
miRNA	Micro RNA
mRNA	Messenger RNA
MZ	Marginal Zone B-Cells
ncRNAs	Non-Coding RNAs
Nfi	Nuclear Factor 1
NGS	Next Generation Sequencing
NPCs	Neural Progenitor Cells
NSCs	Neural Stem Cells
OPCs	Oligodendrocyte Precursor Cells
PBs	Plasmablasts
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PGCs	Primordial Germ Cells
piRNA	Piwi-Interacting RNA
PMDs	Partially Methylated Regions
polyA	Poly Adenylated

RNAPII	RNA Polymerase II
RPKM	Reads Per Kilobase Per Million
RRBS	Reduced Representation Bisulfite Sequencing
SAM	S-Andenosyl Methionine
Shh	Sonic Hedgehog
siRNA	Short Interfering RNA
T1	Transitional 1 B-Cells
T2	Transitional 2 B-Cells
TADs	Topological Associating Domains
TET	Ten-Eleven Translocation
TFs	Transcription Factors
Tregs	Regulatory T-Cells
TSS	Transcription Start Site
TTS	Transcription Termination Site
UMRs	Unmethylated Regions
WGBS	Whole Genome Bisulfite Sequencing

B. List of Figures

C. List of Tables

D. References

Afgan, Enis, Dannon Baker, Bérénice Batut, Marius van den Beek, Dave Bouvier, Martin Čech, John Chilton, et al. 2018. 'The Galaxy Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2018 Update'. *Nucleic Acids Research* 46 (W1): W537–44. https://doi.org/10.1093/nar/gky379.

Agoston, Zsuzsa, Naixin Li, Anja Haslinger, Andrea Wizenmann, and Dorothea Schulte. 2012. 'Genetic and Physical Interaction of Meis2, Pax3 and Pax7 during Dorsal Midbrain Development'. *BMC Developmental Biology* 12 (1): 10. https://doi.org/10.1186/1471-213X-12-10.

Akalin, Altuna, Matthias Kormaksson, Sheng Li, Francine E Garrett-Bakelman, Maria E Figueroa, Ari Melnick, and Christopher E Mason. 2012. 'MethylKit: A Comprehensive R Package for the Analysis of Genome-Wide DNA Methylation Profiles'. *Genome Biology* 13 (10): R87. https://doi.org/10.1186/gb-2012-13-10-r87.

Andriezen, W. Lloyd. 1893. 'The Neuroglia Elements in the Human Brain'. *British Medical Journal* 2 (1700): 227–30.

Armañanzas, Rubén, and Giorgio A. Ascoli. 2015. 'Towards the Automatic Classification of Neurons'. *Trends in Neurosciences* 38 (5): 307–18. https://doi.org/10.1016/j.tins.2015.02.004.

Aruga, Jun, Osamu Minowa, Hiroyuki Yaginuma, Junko Kuno, Takeharu Nagai, Tetsuo Noda, and Katsuhiko Mikoshiba. 1998. 'Mouse Zic1 Is Involved in Cerebellar Development'. *Journal of Neuroscience* 18 (1): 284–93. https://doi.org/10.1523/JNEUROSCI.18-01-00284.1998.

Barclay, A. Neil, Gavin J. Wright, Gary Brooke, and Marion H. Brown. 2002. 'CD200 and Membrane Protein Interactions in the Control of Myeloid Cells'. *Trends in Immunology* 23 (6): 285–90. https://doi.org/10.1016/S1471-4906(02)02223-8.

Barski, Artem, Suresh Cuddapah, Kairong Cui, Tae-Young Roh, Dustin E. Schones, Zhibin Wang, Gang Wei, Iouri Chepelev, and Keji Zhao. 2007. 'High-Resolution Profiling of Histone Methylations in the Human Genome'. *Cell* 129 (4): 823–37. https://doi.org/10.1016/j.cell.2007.05.009.

Baumgarth, Nicole. 2010. 'The Double Life of a B-1 Cell: Self-Reactivity Selects for Protective Effector Functions'. *Nature Reviews Immunology* 11 (1): 34–46. https://doi.org/10.1038/nri2901.

Bayraktar, Omer Ali, Luis C. Fuentealba, Arturo Alvarez-Buylla, and David H. Rowitch. 2015. 'Astrocyte Development and Heterogeneity'. *Cold Spring Harbor Perspectives in Biology* 7 (1): a020362. https://doi.org/10.1101/cshperspect.a020362.

Berland, Robert, and Henry H. Wortis. 2002. 'Origins and Functions of B-1 Cells with Notes on the Role of CD5'. *Annual Review of Immunology* 20: 253–300. https://doi.org/10.1146/annurev.immunol.20.100301.064833.

Berlivet, Soizik, Denis Paquette, Annie Dumouchel, David Langlais, Josée Dostie, and Marie Kmita. 2013. 'Clustering of Tissue-Specific Sub-TADs Accompanies the Regulation of HoxA Genes in Developing Limbs'. Edited by Bing Ren. *PLoS Genetics* 9 (12): e1004018. https://doi.org/10.1371/journal.pgen.1004018.

Bian, Zhen, Lei Shi, Ya-Lan Guo, Zhiyuan Lv, Cong Tang, Shuo Niu, Alexandra Tremblay, et al. 2016. 'Cd47-Sirpα Interaction and IL-10 Constrain

Inflammation-Induced Macrophage Phagocytosis of Healthy Self-Cells'. *Proceedings of the National Academy of Sciences of the United States of America* 113 (37): E5434–43. https://doi.org/10.1073/pnas.1521069113.

Bird, A. P. 1980. 'DNA Methylation and the Frequency of CpG in Animal DNA'. *Nucleic Acids Research* 8 (7): 1499–1504.

Bird, Adrian, Mary Taggart, Marianne Frommer, Orlando J. Miller, and Donald Macleod. 1985. 'A Fraction of the Mouse Genome That Is Derived from Islands of Nonmethylated, CpG-Rich DNA'. *Cell* 40 (1): 91–99. https://doi.org/10.1016/0092-8674(85)90312-5.

Blackshear, Perry J., Wi S. Lai, Elizabeth A. Kennington, Gary Brewer, Gerald M. Wilson, Xiaoju Guan, and Pei Zhou. 2003. 'Characteristics of the Interaction of a Synthetic Human Tristetraprolin Tandem Zinc Finger Peptide with AU-Rich Element-Containing RNA Substrates'. *The Journal of Biological Chemistry* 278 (22): 19947–55. https://doi.org/10.1074/jbc.M301290200.

Boisvert, Matthew M., Galina A. Erikson, Maxim N. Shokhirev, and Nicola J. Allen. 2018. 'The Aging Astrocyte Transcriptome from Multiple Regions of the Mouse Brain'. *Cell Reports* 22 (1): 269–85. https://doi.org/10.1016/j.celrep.2017.12.039.

Bos, N. A., H. Kimura, C. G. Meeuwsen, H. De Visser, M. P. Hazenberg, B. S. Wostmann, J. R. Pleasants, R. Benner, and D. M. Marcus. 1989. 'Serum Immunoglobulin Levels and Naturally Occurring Antibodies against Carbohydrate Antigens in Germ-Free BALB/c Mice Fed Chemically Defined Ultrafiltered Diet'. *European Journal of Immunology* 19 (12): 2335–39. https://doi.org/10.1002/eji.1830191223.

Bossard, P., and K. S. Zaret. 1998. 'GATA Transcription Factors as Potentiators of Gut Endoderm Differentiation'. *Development (Cambridge, England)* 125 (24): 4909–17.

Bouaziz, Jean-David, Koichi Yanaba, Guglielmo M. Venturi, Yaming Wang, Roland M. Tisch, Jonathan C. Poe, and Thomas F. Tedder. 2007. 'Therapeutic B Cell Depletion Impairs Adaptive and Autoreactive CD4+ T Cell Activation in Mice'. *Proceedings of the National Academy of Sciences of the United States of America* 104 (52): 20878–83. https://doi.org/10.1073/pnas.0709205105.

Boyle, Alan P., Sean Davis, Hennady P. Shulha, Paul Meltzer, Elliott H. Margulies, Zhiping Weng, Terrence S. Furey, and Gregory E. Crawford. 2008. 'High-Resolution Mapping and Characterization of Open Chromatin across the Genome'. *Cell* 132 (2): 311–22. https://doi.org/10.1016/j.cell.2007.12.014.

Boyles, J. K., R. E. Pitas, E. Wilson, R. W. Mahley, and J. M. Taylor. 1985. 'Apolipoprotein E Associated with Astrocytic Glia of the Central Nervous System and with Nonmyelinating Glia of the Peripheral Nervous System.' *The Journal of Clinical Investigation* 76 (4): 1501–13. https://doi.org/10.1172/JCI112130.

Brown, Angus M., and Bruce R. Ransom. 2007. 'Astrocyte Glycogen and Brain Energy Metabolism'. *Glia* 55 (12): 1263–71. https://doi.org/10.1002/glia.20557.

Buenrostro, Jason D., Paul G. Giresi, Lisa C. Zaba, Howard Y. Chang, and William J. Greenleaf. 2013. 'Transposition of Native Chromatin for Fast and Sensitive Epigenomic Profiling of Open Chromatin, DNA-Binding Proteins and Nucleosome Position'. *Nature Methods* 10 (12): 1213–18. https://doi.org/10.1038/nmeth.2688.

Bulchand, S., E. A. Grove, F. D. Porter, and S. Tole. 2001. 'LIM-Homeodomain Gene Lhx2 Regulates the Formation of the Cortical Hem'. *Mechanisms of Development* 100 (2): 165–75. https://doi.org/10.1016/S0925-4773(00)00515-3.

Burger, Lukas, Dimos Gaidatzis, Dirk Schübeler, and Michael B. Stadler. 2013. 'Identification of Active Regulatory Regions from DNA Methylation Data'. *Nucleic Acids Research* 41 (16): e155–e155. https://doi.org/10.1093/nar/gkt599.

Bushong, Eric A., Maryann E. Martone, Ying Z. Jones, and Mark H. Ellisman. 2002. 'Protoplasmic Astrocytes in CA1 Stratum Radiatum Occupy Separate Anatomical Domains'. *Journal of Neuroscience* 22 (1): 183–92. https://doi.org/10.1523/JNEUROSCI.22-01-00183.2002.

Campbell, Kenneth. 2003. 'Dorsal-Ventral Patterning in the Mammalian Telencephalon'. *Current Opinion in Neurobiology* 13 (1): 50–56. https://doi.org/10.1016/S0959-4388(03)00009-6.

Carter, Natalie A., Rita Vasconcellos, Elizabeth C. Rosser, Calogero Tulone, Alba Muñoz-Suano, Masahito Kamanaka, Michael R. Ehrenstein, Richard A. Flavell, and Claudia Mauri. 2011. 'Mice Lacking Endogenous IL-10-Producing Regulatory B Cells Develop Exacerbated Disease and Present with an Increased Frequency of Th1/Th17 but a Decrease in Regulatory T Cells'. *Journal of Immunology (Baltimore, Md.: 1950)* 186 (10): 5569–79. https://doi.org/10.4049/jimmunol.1100284.

Carullo, Nancy V. N., and Jeremy J. Day. 2019. 'Genomic Enhancers in Brain Health and Disease'. *Genes* 10 (1). https://doi.org/10.3390/genes10010043.

Chai, Hua, Blanca Diaz-Castro, Eiji Shigetomi, Emma Monte, J. Christopher Octeau, Xinzhu Yu, Whitaker Cohn, et al. 2017. 'Neural Circuit-Specialized Astrocytes: Transcriptomic, Proteomic, Morphological, and Functional Evidence'. *Neuron* 95 (3): 531-549.e9. https://doi.org/10.1016/j.neuron.2017.06.029.

Chen, Ting-Wen, Hsin-Pai Li, Chi-Ching Lee, Ruei-Chi Gan, Po-Jung Huang, Timothy H Wu, Cheng-Yang Lee, Yi-Feng Chang, and Petrus Tang. 2014. 'ChIPseek, a Web-Based Analysis Tool for ChIP Data'. *BMC Genomics* 15 (1): 539. https://doi.org/10.1186/1471-2164-15-539.

Cheng, Yulan, Anamaria Sudarov, Kamila U. Szulc, Sema K. Sgaier, Daniel Stephen, Daniel H. Turnbull, and Alexandra L. Joyner. 2010. 'The Engrailed Homeobox Genes Determine the Different Foliation Patterns in the Vermis and Hemispheres of the Mammalian Cerebellum'. *Development* 137 (3): 519–29. https://doi.org/10.1242/dev.027045.

Chung, Won-Suk, Laura E. Clarke, Gordon X. Wang, Benjamin K. Stafford, Alexander Sher, Chandrani Chakraborty, Julia Joung, et al. 2013. 'Astrocytes Mediate Synapse Elimination through MEGF10 and MERTK Pathways'. *Nature* 504 (7480): 394–400. https://doi.org/10.1038/nature12776.

Chung, Won-Suk, Philip B. Verghese, Chandrani Chakraborty, Julia Joung, Bradley T. Hyman, Jason D. Ulrich, David M. Holtzman, and Ben A. Barres. 2016. 'Novel Allele-Dependent Role for APOE in Controlling the Rate of Synapse Pruning by Astrocytes'. *Proceedings of the National Academy of Sciences of the United States of America* 113 (36): 10186–91. https://doi.org/10.1073/pnas.1609896113.

Clevers, Hans, Susanne Rafelski, Michael Elowitz, Allon Klein, Cole Trapnell, Jay Shendure, Ed Lein, et al. 2017. 'What Is Your Conceptual Definition of "Cell

Type" in the Context of a Mature Organism?' *Cell Systems* 4 (3): 255–59. https://doi.org/10.1016/j.cels.2017.03.006.

Couper, Kevin N., Daniel G. Blount, and Eleanor M. Riley. 2008. 'IL-10: The Master Regulator of Immunity to Infection'. *Journal of Immunology (Baltimore, Md.: 1950)* 180 (9): 5771–77.

Cremer, Thomas, and Marion Cremer. 2010. 'Chromosome Territories'. *Cold Spring Harbor Perspectives in Biology* 2 (3): a003889. https://doi.org/10.1101/cshperspect.a003889.

Dann, Geoffrey P., Glen P. Liszczak, John D. Bagert, Manuel M. Müller, Uyen T. T. Nguyen, Felix Wojcik, Zachary Z. Brown, et al. 2017. 'ISWI Chromatin Remodellers Sense Nucleosome Modifications to Determine Substrate Preference'. *Nature* 548 (7669): 607–11. https://doi.org/10.1038/nature23671.

De Arras, Lesly, Ivana V. Yang, Brad Lackford, David W. H. Riches, Rytis Prekeris, Jonathan H. Freedman, David A. Schwartz, and Scott Alper. 2012. 'SPATIOTEMPORAL INHIBITION OF INNATE IMMUNITY SIGNALING BY THE TBC1D23 RAB-GAP'. *Journal of Immunology (Baltimore, Md. : 1950)* 188 (6): 2905–13. https://doi.org/10.4049/jimmunol.1102595.

DeLuca, David S., Joshua Z. Levin, Andrey Sivachenko, Timothy Fennell, Marc-Danie Nazaire, Chris Williams, Michael Reich, Wendy Winckler, and Gad Getz. 2012. 'RNA-SeQC: RNA-Seq Metrics for Quality Control and Process Optimization'. *Bioinformatics* 28 (11): 1530–32. https://doi.org/10.1093/bioinformatics/bts196.

Deneen, Benjamin, Ritchie Ho, Agnes Lukaszewicz, Christian J. Hochstim, Richard M. Gronostajski, and David J. Anderson. 2006. 'The Transcription Factor NFIA Controls the Onset of Gliogenesis in the Developing Spinal Cord'. *Neuron* 52 (6): 953–68. https://doi.org/10.1016/j.neuron.2006.11.019.

Dixon, Jesse R., Siddarth Selvaraj, Feng Yue, Audrey Kim, Yan Li, Yin Shen, Ming Hu, Jun S. Liu, and Bing Ren. 2012. 'Topological Domains in Mammalian Genomes Identified by Analysis of Chromatin Interactions'. *Nature* 485 (7398): 376–80. https://doi.org/10.1038/nature11082.

Dobin, Alexander, and Thomas R. Gingeras. 2015. 'Mapping RNA-Seq Reads with STAR: Mapping RNA-Seq Reads with STAR'. In *Current Protocols in Bioinformatics*, edited by Alex Bateman, William R. Pearson, Lincoln D. Stein, Gary D. Stormo, and John R. Yates, 11.14.1-11.14.19. Hoboken, NJ, USA: John Wiley & Sons, Inc. https://doi.org/10.1002/0471250953.bi1114s51.

Durek, Pawel, Karl Nordström, Gilles Gasparoni, Abdulrahman Salhab, Christopher Kressler, Melanie de Almeida, Kevin Bassler, et al. 2016. 'Epigenomic Profiling of Human CD4+ T Cells Supports a Linear Differentiation Model and Highlights Molecular Regulators of Memory Development'. *Immunity* 45 (5): 1148–61. https://doi.org/10.1016/j.immuni.2016.10.022.

Elsen, Gina E., Louis Y. Choi, Kathleen J. Millen, Yevgenya Grinblat, and Victoria E. Prince. 2008. 'Zic1 and Zic4 Regulate Zebrafish Roof Plate Specification and Hindbrain Ventricle Morphogenesis'. *Developmental Biology* 314 (2): 376–92. https://doi.org/10.1016/j.ydbio.2007.12.006.

Erdel, Fabian, and Karsten Rippe. 2011. 'Chromatin Remodelling in Mammalian Cells by ISWI-Type Complexes--Where, When and Why?' *The FEBS Journal* 278 (19): 3608–18. https://doi.org/10.1111/j.1742-4658.2011.08282.x.

Ernst, Jason, and Manolis Kellis. 2017. 'Chromatin-State Discovery and Genome Annotation with ChromHMM'. *Nature Protocols* 12 (12): 2478–92. https://doi.org/10.1038/nprot.2017.124.

Ernst, Jason, Pouya Kheradpour, Tarjei S. Mikkelsen, Noam Shoresh, Lucas D. Ward, Charles B. Epstein, Xiaolan Zhang, et al. 2011. 'Mapping and Analysis of Chromatin State Dynamics in Nine Human Cell Types'. *Nature* 473 (7345): 43–49. https://doi.org/10.1038/nature09906.

Fan, Guoping, Keri Martinowich, Mark H. Chin, Fei He, Shaun D. Fouse, Leah Hutnick, Daisuke Hattori, et al. 2005. 'DNA Methylation Controls the Timing of Astrogliogenesis through Regulation of JAK-STAT Signaling'. *Development* 132 (15): 3345–56. https://doi.org/10.1242/dev.01912.

Farmer, W. T., T. Abrahamsson, S. Chierzi, C. Lui, C. Zaelzer, E. V. Jones, B. P. Bally, et al. 2016. 'Neurons Diversify Astrocytes in the Adult Brain through Sonic Hedgehog Signaling'. *Science* 351 (6275): 849–54. https://doi.org/10.1126/science.aab3103.

Farmer, W. Todd, and Keith Murai. 2017. 'Resolving Astrocyte Heterogeneity in the CNS'. *Frontiers in Cellular Neuroscience* 11 (September). https://doi.org/10.3389/fncel.2017.00300.

Fillatreau, Simon. 2013. 'Cytokine-Producing B Cells as Regulators of Pathogenic and Protective Immune Responses'. *Annals of the Rheumatic Diseases* 72 (suppl 2): ii80–84. https://doi.org/10.1136/annrheumdis-2012-202253.

Fillatreau, Simon, Claire H. Sweenie, Mandy J. McGeachy, David Gray, and Stephen M. Anderton. 2002. 'B Cells Regulate Autoimmunity by Provision of IL-10'. *Nature Immunology* 3 (10): 944–50. https://doi.org/10.1038/ni833.

Frank, Christopher L, Fang Liu, Ranjula Wijayatunge, Lingyun Song, Matthew T Biegler, Marty G Yang, Christopher M Vockley, et al. 2015. 'Regulation of Chromatin Accessibility and Zic Binding at Enhancers in the Developing Cerebellum'. *Nature Neuroscience* 18 (5): 647–56. https://doi.org/10.1038/nn.3995.

Fudenberg, Geoffrey, Maxim Imakaev, Carolyn Lu, Anton Goloborodko, Nezar Abdennur, and Leonid A. Mirny. 2016. 'Formation of Chromosomal Domains by Loop Extrusion'. *Cell Reports* 15 (9): 2038–49. https://doi.org/10.1016/j.celrep.2016.04.085.

Fuks, F., W. A. Burgers, A. Brehm, L. Hughes-Davies, and T. Kouzarides. 2000. 'DNA Methyltransferase Dnmt1 Associates with Histone Deacetylase Activity'. *Nature Genetics* 24 (1): 88–91. https://doi.org/10.1038/71750.

Fuks, F., Paul J. Hurd, Rachel Deplus, and Tony Kouzarides. 2003. 'The DNA Methyltransferases Associate with HP1 and the SUV39H1 Histone Methyltransferase'. *Nucleic Acids Research* 31 (9): 2305–12.

Fullwood, Melissa J., and Yijun Ruan. 2009. 'ChIP-Based Methods for the Identification of Long-Range Chromatin Interactions'. *Journal of Cellular Biochemistry* 107 (1): 30–39. https://doi.org/10.1002/jcb.22116.

García-Cáceres, Cristina, Carmelo Quarta, Luis Varela, Yuanqing Gao, Tim Gruber, Beata Legutko, Martin Jastroch, et al. 2016. 'Astrocytic Insulin Signaling Couples Brain Glucose Uptake with Nutrient Availability'. *Cell* 166 (4): 867–80. https://doi.org/10.1016/j.cell.2016.07.028.

García-Marín, Virginia, Pablo García-López, and Miguel Freire. 2007. 'Cajal's Contributions to Glia Research'. *Trends in Neurosciences* 30 (9): 479–87. https://doi.org/10.1016/j.tins.2007.06.008.

Gardiner-Garden, M., and M. Frommer. 1987. 'CpG Islands in Vertebrate Genomes'. *Journal of Molecular Biology* 196 (2): 261–82. https://doi.org/10.1016/0022-2836(87)90689-9.

Geiman, Theresa M., Umesh T. Sankpal, Andrea K. Robertson, Yingxin Zhao, Yingming Zhao, and Keith D. Robertson. 2004. 'DNMT3B Interacts with HSNF2H Chromatin Remodeling Enzyme, HDACs 1 and 2, and Components of the Histone Methylation System'. *Biochemical and Biophysical Research Communications* 318 (2): 544–55. https://doi.org/10.1016/j.bbrc.2004.04.058.

Gibcus, Johan H., and Job Dekker. 2013. 'The Hierarchy of the 3D Genome'. *Molecular Cell* 49 (5): 773–82. https://doi.org/10.1016/j.molcel.2013.02.011.

Gilfillan, Gregor D., Timothy Hughes, Ying Sheng, Hanne S. Hjorthaug, Tobias Straub, Kristina Gervin, Jennifer R. Harris, Dag E. Undlien, and Robert Lyle. 2012. 'Limitations and Possibilities of Low Cell Number ChIP-Seq'. *BMC Genomics* 13 (1): 645. https://doi.org/10.1186/1471-2164-13-645.

Giresi, Paul G., Jonghwan Kim, Ryan M. McDaniell, Vishwanath R. Iyer, and Jason D. Lieb. 2007. 'FAIRE (Formaldehyde-Assisted Isolation of Regulatory Elements) Isolates Active Regulatory Elements from Human Chromatin'. *Genome Research* 17 (6): 877–85. https://doi.org/10.1101/gr.5533506.

Globisch, Daniel, Martin Münzel, Markus Müller, Stylianos Michalakis, Mirko Wagner, Susanne Koch, Tobias Brückl, Martin Biel, and Thomas Carell. 2010. 'Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates'. *PLoS ONE* 5 (12). https://doi.org/10.1371/journal.pone.0015367.

Godbole, Geeta, Ashwin S. Shetty, Achira Roy, Leora D'Souza, Bin Chen, Goichi Miyoshi, Gordon Fishell, and Shubha Tole. 2018. 'Hierarchical Genetic Interactions between FOXG1 and LHX2 Regulate the Formation of the Cortical Hem in the Developing Telencephalon'. *Development* 145 (1): dev154583. https://doi.org/10.1242/dev.154583.

Goldberg, Aaron D., C. David Allis, and Emily Bernstein. 2007. 'Epigenetics: A Landscape Takes Shape'. *Cell* 128 (4): 635–38. https://doi.org/10.1016/j.cell.2007.02.006.

Gourine, Alexander V., Vitaliy Kasymov, Nephtali Marina, Feige Tang, Melina F. Figueiredo, Samantha Lane, Anja G. Teschemacher, K. Michael Spyer, Karl Deisseroth, and Sergey Kasparov. 2010. 'Astrocytes Control Breathing through PH-Dependent Release of ATP'. *Science (New York, N.Y.)* 329 (5991): 571–75. https://doi.org/10.1126/science.1190721.

Gu, Zuguang, Lei Gu, Roland Eils, Matthias Schlesner, and Benedikt Brors. 2014. 'Circlize Implements and Enhances Circular Visualization in R'. *Bioinformatics (Oxford, England)* 30 (19): 2811–12. https://doi.org/10.1093/bioinformatics/btu393.

Gualdi, R., P. Bossard, M. Zheng, Y. Hamada, J. R. Coleman, and K. S. Zaret. 1996. 'Hepatic Specification of the Gut Endoderm in Vitro: Cell Signaling and Transcriptional Control'. *Genes & Development* 10 (13): 1670–82.

Hajkova, Petra, Sylvia Erhardt, Natasha Lane, Thomas Haaf, Osman El-Maarri, Wolf Reik, Jörn Walter, and M. Azim Surani. 2002. 'Epigenetic Reprogramming in

Mouse Primordial Germ Cells'. *Mechanisms of Development* 117 (1): 15–23. https://doi.org/10.1016/S0925-4773(02)00181-8.

Hanashima, Carina, Suzanne C. Li, Lijian Shen, Eseng Lai, and Gord Fishell. 2004. 'Foxg1 Suppresses Early Cortical Cell Fate'. *Science* 303 (5654): 56–59. https://doi.org/10.1126/science.1090674.

Hardy, Richard R., Chi-Ju Wei, and Kyoko Hayakawa. 2004. 'Selection during Development of VH11+ B Cells: A Model for Natural Autoantibody-Producing CD5+ B Cells'. *Immunological Reviews* 197 (February): 60–74.

Harrison, Susan J., Ryuichi Nishinakamura, Kevin R. Jones, and A. Paula Monaghan. 2012. 'Sall1 Regulates Cortical Neurogenesis and Laminar Fate Specification in Mice: Implications for Neural Abnormalities in Townes-Brocks Syndrome'. *Disease Models & Mechanisms* 5 (3): 351–65. https://doi.org/10.1242/dmm.002873.

Haury, M., A. Sundblad, A. Grandien, C. Barreau, A. Coutinho, and A. Nobrega. 1997. 'The Repertoire of Serum IgM in Normal Mice Is Largely Independent of External Antigenic Contact'. *European Journal of Immunology* 27 (6): 1557–63. https://doi.org/10.1002/eji.1830270635.

Haydon, Philip G., and Vladimir Parpura, eds. 2009. *Astrocytes in (Patho)Physiology of the Nervous System*. Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-79492-1.

He, Yu-Fei, Bin-Zhong Li, Zheng Li, Peng Liu, Yang Wang, Qingyu Tang, Jianping Ding, et al. 2011. 'Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA'. *Science (New York, N.Y.)* 333 (6047): 1303–7. https://doi.org/10.1126/science.1210944.

Heintzman, Nathaniel D., Gary C. Hon, R. David Hawkins, Pouya Kheradpour, Alexander Stark, Lindsey F. Harp, Zhen Ye, et al. 2009. 'Histone Modifications at Human Enhancers Reflect Global Cell Type-Specific Gene Expression'. *Nature* 459 (7243): 108–12. https://doi.org/10.1038/nature07829.

Hellman, Asaf, and Andrew Chess. 2007. 'Gene Body-Specific Methylation on the Active X Chromosome'. *Science (New York, N.Y.)* 315 (5815): 1141–43. https://doi.org/10.1126/science.1136352.

Hermann, Andrea, Rachna Goyal, and Albert Jeltsch. 2004. 'The Dnmt1 DNA-(Cytosine-C5)-Methyltransferase Methylates DNA Processively with High Preference for Hemimethylated Target Sites'. *The Journal of Biological Chemistry* 279 (46): 48350–59. https://doi.org/10.1074/jbc.M403427200.

Hnisz, Denes, Daniel S. Day, and Richard A. Young. 2016. 'Insulated Neighborhoods: Structural and Functional Units of Mammalian Gene Control'. *Cell* 167 (5): 1188–1200. https://doi.org/10.1016/j.cell.2016.10.024.

Hochstim, Christian, Benjamin Deneen, Agnès Lukaszewicz, Qiao Zhou, and David J. Anderson. 2008. 'Identification of Positionally Distinct Astrocyte Subtypes Whose Identities Are Specified by a Homeodomain Code'. *Cell* 133 (3): 510–22. https://doi.org/10.1016/j.cell.2008.02.046.

Hoffman, William, Fadi G. Lakkis, and Geetha Chalasani. 2016. 'B Cells, Antibodies, and More'. *Clinical Journal of the American Society of Nephrology : CJASN* 11 (1): 137–54. https://doi.org/10.2215/CJN.09430915.

Holliday, R., and J. E. Pugh. 1975. 'DNA Modification Mechanisms and Gene Activity during Development'. *Science (New York, N.Y.)* 187 (4173): 226–32.

Hon, Gary C., Chun-Xiao Song, Tingting Du, Fulai Jin, Siddarth Selvaraj, Ah Young Lee, Chia-an Yen, et al. 2014. '5mC Oxidation by Tet2 Modulates Enhancer Activity and Timing of Transcriptome Reprogramming during Differentiation'. *Molecular Cell* 56 (2): 286–97. https://doi.org/10.1016/j.molcel.2014.08.026.

Hovestadt, Volker, David T. W. Jones, Simone Picelli, Wei Wang, Marcel Kool, Paul A. Northcott, Marc Sultan, et al. 2014. 'Decoding the Regulatory Landscape of Medulloblastoma Using DNA Methylation Sequencing'. *Nature* 510 (7506): 537–41. https://doi.org/10.1038/nature13268.

Huang, Da Wei, Brad T. Sherman, and Richard A. Lempicki. 2009a. 'Systematic and Integrative Analysis of Large Gene Lists Using DAVID Bioinformatics Resources'. *Nature Protocols* 4 (1): 44–57. https://doi.org/10.1038/nprot.2008.211.

———. 2009b. 'Bioinformatics Enrichment Tools: Paths toward the Comprehensive Functional Analysis of Large Gene Lists'. *Nucleic Acids Research* 37 (1): 1–13. https://doi.org/10.1093/nar/gkn923.

Huard, Bertrand, Muriel Tournier, Thierry Hercend, Frédéric Triebel, and Florence Faure. 1994. 'Lymphocyte-Activation Gene 3/Major Histocompatibility Complex Class II Interaction Modulates the Antigenic Response of CD4+ T Lymphocytes'. *European Journal of Immunology* 24 (12): 3216–21. https://doi.org/10.1002/eji.1830241246.

Hudson, Brian P., Maria A. Martinez-Yamout, H. Jane Dyson, and Peter E. Wright. 2004. 'Recognition of the MRNA AU-Rich Element by the Zinc Finger Domain of TIS11d'. *Nature Structural & Molecular Biology* 11 (3): 257–64. https://doi.org/10.1038/nsmb738.

Ipseiz, Natacha, Stefan Uderhardt, Carina Scholtysek, Martin Steffen, Gernot Schabbauer, Aline Bozec, Georg Schett, and Gerhard Krönke. 2014. 'The Nuclear Receptor Nr4a1 Mediates Anti-Inflammatory Effects of Apoptotic Cells'. *The Journal of Immunology* 192 (10): 4852–58. https://doi.org/10.4049/jimmunol.1303377.

Ito, Shinsuke, Ana C. D'Alessio, Olena V. Taranova, Kwonho Hong, Lawrence C. Sowers, and Yi Zhang. 2010. 'Role of Tet Proteins in 5mC to 5hmC Conversion, ES Cell Self-Renewal, and ICM Specification'. *Nature* 466 (7310): 1129–33. https://doi.org/10.1038/nature09303.

Ito, Shinsuke, Li Shen, Qing Dai, Susan C. Wu, Leonard B. Collins, James A. Swenberg, Chuan He, and Yi Zhang. 2011. 'Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine'. *Science (New York, N.Y.)* 333 (6047): 1300–1303. https://doi.org/10.1126/science.1210597.

Iwasaki, Yukiko, Keishi Fujio, Tomohisa Okamura, Atsushi Yanai, Shuji Sumitomo, Hirofumi Shoda, Tomohiko Tamura, Hiroki Yoshida, Patrick Charnay, and Kazuhiko Yamamoto. 2013. 'Egr-2 Transcription Factor Is Required for Blimp-1-Mediated IL-10 Production in IL-27-Stimulated CD4+ T Cells'. *European Journal of Immunology* 43 (4): 1063–73. https://doi.org/10.1002/eji.201242942.

Jenuwein, T., and C. D. Allis. 2001. 'Translating the Histone Code'. *Science (New York, N.Y.)* 293 (5532): 1074–80. https://doi.org/10.1126/science.1063127.

John Lin, Chia-Ching, Kwanha Yu, Asante Hatcher, Teng-Wei Huang, Hyun Kyoung Lee, Jeffrey Carlson, Matthew C Weston, et al. 2017. 'Identification of Diverse

Astrocyte Populations and Their Malignant Analogs'. *Nature Neuroscience* 20 (3): 396–405. https://doi.org/10.1038/nn.4493.

Johnson, David S., Ali Mortazavi, Richard M. Myers, and Barbara Wold. 2007. 'Genome-Wide Mapping of in Vivo Protein-DNA Interactions'. *Science (New York, N.Y.)* 316 (5830): 1497–1502. https://doi.org/10.1126/science.1141319.

Jones, P. L., G. J. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis, and A. P. Wolffe. 1998. 'Methylated DNA and MeCP2 Recruit Histone Deacetylase to Repress Transcription'. *Nature Genetics* 19 (2): 187–91. https://doi.org/10.1038/561.

Jost, K. Laurence, Bianca Bertulat, and M. Cristina Cardoso. 2012. 'Heterochromatin and Gene Positioning: Inside, Outside, Any Side?' *Chromosoma* 121 (6): 555–63. https://doi.org/10.1007/s00412-012-0389-2.

Kamakaka, Rohinton T., and Sue Biggins. 2005. 'Histone Variants: Deviants?' *Genes & Development* 19 (3): 295–310. https://doi.org/10.1101/gad.1272805.

Kaplan, Noam, Irene K. Moore, Yvonne Fondufe-Mittendorf, Andrea J. Gossett, Desiree Tillo, Yair Field, Emily M. LeProust, et al. 2009. 'The DNA-Encoded Nucleosome Organization of a Eukaryotic Genome'. *Nature* 458 (7236): 362–66. https://doi.org/10.1038/nature07667.

Khan, Aziz, Oriol Fornes, Arnaud Stigliani, Marius Gheorghe, Jaime A Castro-Mondragon, Robin van der Lee, Adrien Bessy, et al. 2018. 'JASPAR 2018: Update of the Open-Access Database of Transcription Factor Binding Profiles and Its Web Framework'. *Nucleic Acids Research* 46 (D1): D260–66. https://doi.org/10.1093/nar/gkx1126.

Kimura, Jun, Yoko Suda, Daisuke Kurokawa, Zakir M. Hossain, Miwa Nakamura, Maiko Takahashi, Akemi Hara, and Shinichi Aizawa. 2005. 'Emx2 and Pax6 Function in Cooperation with Otx2 and Otx1 to Develop Caudal Forebrain Primordium That Includes Future Archipallium'. *Journal of Neuroscience* 25 (21): 5097–5108. https://doi.org/10.1523/JNEUROSCI.0239-05.2005.

Koohy, Hashem, Thomas A. Down, Mikhail Spivakov, and Tim Hubbard. 2014. 'A Comparison of Peak Callers Used for DNase-Seq Data'. *PLOS ONE* 9 (5): e96303. https://doi.org/10.1371/journal.pone.0096303.

Kucukdereli, Hakan, Nicola J. Allen, Anthony T. Lee, Ava Feng, M. Ilcim Ozlu, Laura M. Conatser, Chandrani Chakraborty, et al. 2011. 'Control of Excitatory CNS Synaptogenesis by Astrocyte-Secreted Proteins Hevin and SPARC'. *Proceedings of the National Academy of Sciences of the United States of America* 108 (32): E440-449. https://doi.org/10.1073/pnas.1104977108.

Kulis, Marta, Angelika Merkel, Simon Heath, Ana C Queirós, Ronald P Schuyler, Giancarlo Castellano, Renée Beekman, et al. 2015. 'Whole-Genome Fingerprint of the DNA Methylome during Human B Cell Differentiation'. *Nature Genetics* 47 (7): 746–56. https://doi.org/10.1038/ng.3291.

Lanjakornsiripan, Darin, Baek-Jun Pior, Daichi Kawaguchi, Shohei Furutachi, Tomoaki Tahara, Yu Katsuyama, Yutaka Suzuki, Yugo Fukazawa, and Yukiko Gotoh. 2018. 'Layer-Specific Morphological and Molecular Differences in Neocortical Astrocytes and Their Dependence on Neuronal Layers'. *Nature Communications* 9 (1): 1623. https://doi.org/10.1038/s41467-018-03940-3.

Latchman, Yvette, Clive R. Wood, Tatyana Chernova, Divya Chaudhary, Madhuri Borde, Irene Chernova, Yoshiko Iwai, et al. 2001. 'PD-L2 Is a Second Ligand for

PD-1 and Inhibits T Cell Activation'. *Nature Immunology* 2 (3): 261. https://doi.org/10.1038/85330.

Le Gall, Antoine, Alessandro Valeri, and Marcelo Nollmann. 2015. 'Roles of Chromatin Insulators in the Formation of Long-Range Contacts'. *Nucleus* 6 (2): 118–22. https://doi.org/10.1080/19491034.2015.1010962.

LeBien, Tucker W., and Thomas F. Tedder. 2008. 'B Lymphocytes: How They Develop and Function'. *Blood* 112 (5): 1570–80. https://doi.org/10.1182/blood-2008-02-078071.

Lecaudey, Virginie, Isabelle Anselme, Renate Dildrop, Ulrich Rüther, and Sylvie Schneider-Maunoury. 2005. 'Expression of the Zebrafish Iroquois Genes during Early Nervous System Formation and Patterning'. *Journal of Comparative Neurology* 492 (3): 289–302. https://doi.org/10.1002/cne.20765.

Lee, William, Desiree Tillo, Nicolas Bray, Randall H. Morse, Ronald W. Davis, Timothy R. Hughes, and Corey Nislow. 2007. 'A High-Resolution Atlas of Nucleosome Occupancy in Yeast'. *Nature Genetics* 39 (10): 1235–44. https://doi.org/10.1038/ng2117.

Li, Guoliang, Xiaoan Ruan, Raymond K. Auerbach, Kuljeet Singh Sandhu, Meizhen Zheng, Ping Wang, Huay Mei Poh, et al. 2012. 'Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation'. *Cell* 148 (1–2): 84–98. https://doi.org/10.1016/j.cell.2011.12.014.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Subgroup. 2009. 'The Sequence Alignment/Map Format and SAMtools'. *Bioinformatics* 25 (16): 2078–79. https://doi.org/10.1093/bioinformatics/btp352.

Li, Heng, and Richard Durbin. 2010. 'Fast and Accurate Long-Read Alignment with Burrows-Wheeler Transform'. *Bioinformatics (Oxford, England)* 26 (5): 589–95. https://doi.org/10.1093/bioinformatics/btp698.

Liao, Y., G. K. Smyth, and W. Shi. 2014. 'FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features'. *Bioinformatics* 30 (7): 923–30. https://doi.org/10.1093/bioinformatics/btt656.

Lino, Andreia C., Van Duc Dang, Vicky Lampropoulou, Anna Welle, Jara Joedicke, Jelka Pohar, Quentin Simon, et al. 2018. 'LAG-3 Inhibitory Receptor Expression Identifies Immunosuppressive Natural Regulatory Plasma Cells'. *Immunity* 49 (1): 120-133.e9. https://doi.org/10.1016/j.immuni.2018.06.007.

Lister, Ryan, Eran A. Mukamel, Joseph R. Nery, Mark Urich, Clare A. Puddifoot, Nicholas D. Johnson, Jacinta Lucero, et al. 2013. 'Global Epigenomic Reconfiguration during Mammalian Brain Development'. *Science (New York, N.Y.)* 341 (6146): 1237905. https://doi.org/10.1126/science.1237905.

Lister, Ryan, Mattia Pelizzola, Robert H. Dowen, R. David Hawkins, Gary Hon, Julian Tonti-Filippini, Joseph R. Nery, et al. 2009. 'Human DNA Methylomes at Base Resolution Show Widespread Epigenomic Differences'. *Nature* 462 (7271): 315–22. https://doi.org/10.1038/nature08514.

Liu, X., D. L. Yao, C. A. Bondy, M. Brenner, L. D. Hudson, J. Zhou, and H. D. Webster. 1994. 'Astrocytes Express Insulin-like Growth Factor-I (IGF-I) and Its Binding Protein, IGFBP-2, during Demyelination Induced by Experimental Autoimmune Encephalomyelitis'. *Molecular and Cellular Neurosciences* 5 (5): 418–30. https://doi.org/10.1006/mcne.1994.1052.

Liu, Xiaosong, Jinyan Huang, Taotao Chen, Ying Wang, Shunmei Xin, Jian Li, Gang Pei, and Jiuhong Kang. 2008. 'Yamanaka Factors Critically Regulate the Developmental Signaling Network in Mouse Embryonic Stem Cells'. *Cell Research* 18 (12): 1177–89. https://doi.org/10.1038/cr.2008.309.

Lo, W. S., R. C. Trievel, J. R. Rojas, L. Duggan, J. Y. Hsu, C. D. Allis, R. Marmorstein, and S. L. Berger. 2000. 'Phosphorylation of Serine 10 in Histone H3 Is Functionally Linked in Vitro and in Vivo to Gcn5-Mediated Acetylation at Lysine 14'. *Molecular Cell* 5 (6): 917–26.

Lui, Yuan, and Simon J. Davis. 2018. 'LAG-3: A Very Singular Immune Checkpoint'. *Nature Immunology* 19 (12): 1278. https://doi.org/10.1038/s41590-018-0257-1.

Lun, Aaron T. L., and Gordon K. Smyth. 2016. 'Csaw: A Bioconductor Package for Differential Binding Analysis of ChIP-Seq Data Using Sliding Windows'. *Nucleic Acids Research* 44 (5): e45. https://doi.org/10.1093/nar/gkv1191.

Lund, Frances E. 2008. 'Cytokine-Producing B Lymphocytes-Key Regulators of Immunity'. *Current Opinion in Immunology* 20 (3): 332–38. https://doi.org/10.1016/j.coi.2008.03.003.

Lund, Steven P., Dan Nettleton, Davis J. McCarthy, and Gordon K. Smyth. 2012. 'Detecting Differential Expression in RNA-Sequence Data Using Quasi-Likelihood with Shrunken Dispersion Estimates'. *Statistical Applications in Genetics and Molecular Biology* 11 (5). https://doi.org/10.1515/1544-6115.1826.

Lyons, Shawn M., Clark H. Cunningham, Joshua D. Welch, Beezly Groh, Andrew Y. Guo, Bruce Wei, Michael L. Whitfield, Yue Xiong, and William F. Marzluff. 2016. 'A Subset of Replication-Dependent Histone MRNAs Are Expressed as Polyadenylated RNAs in Terminally Differentiated Tissues'. *Nucleic Acids Research*, July, gkw620. https://doi.org/10.1093/nar/gkw620.

Maclean, Norman, and Brian Keith Hall. 1987. *Cell Commitment and Differentiation*. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

Magavi, Sanjay, Drew Friedmann, Garrett Banks, Alberto Stolfi, and Carlos Lois. 2012. 'Coincident Generation of Pyramidal Neurons and Protoplasmic Astrocytes in Neocortical Columns'. *Journal of Neuroscience* 32 (14): 4762–72. https://doi.org/10.1523/JNEUROSCI.3560-11.2012.

Magnani, Luca, Jérôme Eeckhoute, and Mathieu Lupien. 2011. 'Pioneer Factors: Directing Transcriptional Regulators within the Chromatin Environment'. *Trends in Genetics* 27 (11): 465–74. https://doi.org/10.1016/j.tig.2011.07.002.

Maiti, Atanu, and Alexander C. Drohat. 2011. 'Thymine DNA Glycosylase Can Rapidly Excise 5-Formylcytosine and 5-Carboxylcytosine: Potential Implications for Active Demethylation of CpG Sites'. *The Journal of Biological Chemistry* 286 (41): 35334–38. https://doi.org/10.1074/jbc.C111.284620.

Malatesta, Paolo, Irene Appolloni, and Filippo Calzolari. 2008. 'Radial Glia and Neural Stem Cells'. *Cell and Tissue Research* 331 (1): 165–78. https://doi.org/10.1007/s00441-007-0481-8.

Mallamaci, A., S. Mercurio, L. Muzio, C. Cecchi, C. L. Pardini, P. Gruss, and E. Boncinelli. 2000. 'The Lack of Emx2 Causes Impairment of Reelin Signaling and Defects of Neuronal Migration in the Developing Cerebral Cortex'. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 20 (3): 1109–18.

Maor, Galit Lev, Ahuvi Yearim, and Gil Ast. 2015. 'The Alternative Role of DNA Methylation in Splicing Regulation'. *Trends in Genetics* 31 (5): 274–80. https://doi.org/10.1016/j.tig.2015.03.002.

Marco-Sola, Santiago, Michael Sammeth, Roderic Guigó, and Paolo Ribeca. 2012. 'The GEM Mapper: Fast, Accurate and Versatile Alignment by Filtration'. *Nature Methods* 9 (12): 1185–88. https://doi.org/10.1038/nmeth.2221.

Martin, Marcel. 2011. 'Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads'. *EMBnet.Journal* 17 (1): 10–12. https://doi.org/10.14806/ej.17.1.200.

Martynoga, Ben, Daniela Drechsel, and François Guillemot. 2012. 'Molecular Control of Neurogenesis: A View from the Mammalian Cerebral Cortex'. *Cold Spring Harbor Perspectives in Biology* 4 (10): a008359. https://doi.org/10.1101/cshperspect.a008359.

Maruhashi, Takumi, Il-mi Okazaki, Daisuke Sugiura, Suzuka Takahashi, Takeo K. Maeda, Kenji Shimizu, and Taku Okazaki. 2018. 'LAG-3 Inhibits the Activation of CD4 + T Cells That Recognize Stable PMHCII through Its Conformation-Dependent Recognition of PMHCII'. *Nature Immunology* 19 (12): 1415. https://doi.org/10.1038/s41590-018-0217-9.

Matsumoto, Ken, Shigeki Nishihara, Mika Kamimura, Tomoki Shiraishi, Takao Otoguro, Masayuki Uehara, Yukiko Maeda, Keiko Ogura, Andrew Lumsden, and Toshihiko Ogura. 2004. 'The Prepattern Transcription Factor Irx2, a Target of the FGF8/MAP Kinase Cascade, Is Involved in Cerebellum Formation'. *Nature Neuroscience* 7 (6): 605–12. https://doi.org/10.1038/nn1249.

Mattick, John S. 2009. 'The Genetic Signatures of Noncoding RNAs'. *PLoS Genetics* 5 (4): e1000459. https://doi.org/10.1371/journal.pgen.1000459.

Mauri, Claudia. 2010. 'Regulation of Immunity and Autoimmunity by B Cells'. *Current Opinion in Immunology*, Autoimmunity ● Allergy and hypersensitivity, 22 (6): 761–67. https://doi.org/10.1016/j.coi.2010.10.009.

Mayer, Wolfgang, Alain Niveleau, Jörn Walter, Reinald Fundele, and Thomas Haaf. 2000. 'Embryogenesis: Demethylation of the Zygotic Paternal Genome'. *Nature* 403 (6769): 501–2. https://doi.org/10.1038/35000656.

McLean, Cory Y., Dave Bristor, Michael Hiller, Shoa L. Clarke, Bruce T. Schaar, Craig B. Lowe, Aaron M. Wenger, and Gill Bejerano. 2010. 'GREAT Improves Functional Interpretation of Cis-Regulatory Regions'. *Nature Biotechnology* 28 (5): 495–501. https://doi.org/10.1038/nbt.1630.

Mecklenburg, Nora, Jesus E. Martinez-Lopez, Juan Antonio Moreno-Bravo, Ariadna Perez-Balaguer, Eduardo Puelles, and Salvador Martinez. 2014. 'Growth and Differentiation Factor 10 (Gdf10) Is Involved in Bergmann Glial Cell Development under Shh Regulation'. *Glia* 62 (10): 1713–23. https://doi.org/10.1002/glia.22710.

Meissner, Alexander, Andreas Gnirke, George W. Bell, Bernard Ramsahoye, Eric S. Lander, and Rudolf Jaenisch. 2005. 'Reduced Representation Bisulfite Sequencing for Comparative High-Resolution DNA Methylation Analysis'. *Nucleic Acids Research* 33 (18): 5868–77. https://doi.org/10.1093/nar/gki901.

Melo, Jimmy de, Brian S. Clark, and Seth Blackshaw. 2016. 'Multiple Intrinsic Factors Act in Concert with Lhx2 to Direct Retinal Gliogenesis'. *Scientific Reports* 6: 32757. https://doi.org/10.1038/srep32757.

Melo, Jimmy de, Cristina Zibetti, Brian S. Clark, Woochang Hwang, Ana L. Miranda-Angulo, Jiang Qian, and Seth Blackshaw. 2016. 'Lhx2 Is an Essential Factor for Retinal Gliogenesis and Notch Signaling'. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 36 (8): 2391–2405. https://doi.org/10.1523/JNEUROSCI.3145-15.2016.

Mi, Huaiyu, Xiaosong Huang, Anushya Muruganujan, Haiming Tang, Caitlin Mills, Diane Kang, and Paul D. Thomas. 2017. 'PANTHER Version 11: Expanded Annotation Data from Gene Ontology and Reactome Pathways, and Data Analysis Tool Enhancements'. *Nucleic Acids Research* 45 (D1): D183–89. https://doi.org/10.1093/nar/gkw1138.

Mi, Huaiyu, Anushya Muruganujan, John T. Casagrande, and Paul D. Thomas. 2013. 'Large-Scale Gene Function Analysis with the PANTHER Classification System'. *Nature Protocols* 8 (8): 1551–66. https://doi.org/10.1038/nprot.2013.092.

Mihrshahi, Robin, A. Neil Barclay, and Marion H. Brown. 2009. 'Essential Roles for Dok2 and RasGAP in CD200 Receptor-Mediated Regulation of Human Myeloid Cells'. *The Journal of Immunology* 183 (8): 4879–86. https://doi.org/10.4049/jimmunol.0901531.

Mikkelsen, Tarjei S., Manching Ku, David B. Jaffe, Biju Issac, Erez Lieberman, Georgia Giannoukos, Pablo Alvarez, et al. 2007. 'Genome-Wide Maps of Chromatin State in Pluripotent and Lineage-Committed Cells'. *Nature* 448 (7153): 553–60. https://doi.org/10.1038/nature06008.

Mo, Alisa, Eran A. Mukamel, Fred P. Davis, Chongyuan Luo, Gilbert L. Henry, Serge Picard, Mark A. Urich, et al. 2015. 'Epigenomic Signatures of Neuronal Diversity in the Mammalian Brain'. *Neuron* 86 (6): 1369–84. https://doi.org/10.1016/j.neuron.2015.05.018.

Molyneaux, Bradley J., Paola Arlotta, Joao R. L. Menezes, and Jeffrey D. Macklis. 2007. 'Neuronal Subtype Specification in the Cerebral Cortex'. *Nature Reviews Neuroscience* 8 (6): 427–37. https://doi.org/10.1038/nrn2151.

Monaghan, A. P., D. Bock, P. Gass, A. Schwäger, D. P. Wolfer, H. P. Lipp, and G. Schütz. 1997. 'Defective Limbic System in Mice Lacking the Tailless Gene'. *Nature* 390 (6659): 515–17. https://doi.org/10.1038/37364.

Monuki, Edwin S, Forbes D Porter, and Christopher A Walsh. 2001. 'Patterning of the Dorsal Telencephalon and Cerebral Cortex by a Roof Plate-Lhx2 Pathway'. *Neuron* 32 (4): 591–604. https://doi.org/10.1016/S0896-6273(01)00504-9.

Morel, Lydie, Ming Sum R. Chiang, Haruki Higashimori, Temitope Shoneye, Lakshmanan K. Iyer, Julia Yelick, Albert Tai, and Yongjie Yang. 2017. 'Molecular and Functional Properties of Regional Astrocytes in the Adult Brain'. *The Journal of Neuroscience* 37 (36): 8706–17. https://doi.org/10.1523/JNEUROSCI.3956-16.2017.

Morris, K.V., and J.S. Mattick. 2014. 'The Rise of Regulatory RNA'. *Nature Reviews. Genetics* 15 (6): 423–37. https://doi.org/10.1038/nrg3722.

Moser, Muriel, and Oberdan Leo. 2010. 'Key Concepts in Immunology'. *Vaccine*, Vaccines Educational Supplement, 28 (August): C2–13. https://doi.org/10.1016/j.vaccine.2010.07.022.

Muralidharan, Bhavana, Zeba Khatri, Upasana Maheshwari, Ritika Gupta, Basabdatta Roy, Saurabh J. Pradhan, Krishanpal Karmodiya, et al. 2017. 'LHX2 Interacts with the NuRD Complex and Regulates Cortical Neuron Subtype Determinants

Fezf2 and *Sox11*'. *The Journal of Neuroscience* 37 (1): 194–203. https://doi.org/10.1523/JNEUROSCI.2836-16.2016.

Nakata, Katsunori, Takeharu Nagai, Jun Aruga, and Katsuhiko Mikoshiba. 1998. 'Xenopus Zic Family and Its Role in Neural and Neural Crest Development1During Submission of This Paper, Mizuseki et Al., Reported the Xenopus Zic-Related-1 Gene Which Was Highly Homologous to Mouse Zic1 Gene (Mizuseki et Al., 1998). Accession No. Zic1, AB009564; Zic2, AB009565.1'. *Mechanisms of Development* 75 (1): 43–51. https://doi.org/10.1016/S0925-4773(98)00073-2.

Namihira, Masakazu, Jun Kohyama, Katsunori Semi, Tsukasa Sanosaka, Benjamin Deneen, Tetsuya Taga, and Kinichi Nakashima. 2009. 'Committed Neuronal Precursors Confer Astrocytic Potential on Residual Neural Precursor Cells'. *Developmental Cell* 16 (2): 245–55. https://doi.org/10.1016/j.devcel.2008.12.014.

Namihira, Masakazu, Kinichi Nakashima, and Tetsuya Taga. 2004. 'Developmental Stage Dependent Regulation of DNA Methylation and Chromatin Modification in a Immature Astrocyte Specific Gene Promoter'. *FEBS Letters* 572 (1): 184–88. https://doi.org/10.1016/j.febslet.2004.07.029.

Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman, and A. Bird. 1998. 'Transcriptional Repression by the Methyl-CpG-Binding Protein MeCP2 Involves a Histone Deacetylase Complex'. *Nature* 393 (6683): 386–89. https://doi.org/10.1038/30764.

Neves, Patricia, Vicky Lampropoulou, Elisabeth Calderon-Gomez, Toralf Roch, Ulrik Stervbo, Ping Shen, Anja A. Kühl, et al. 2010. 'Signaling via the MyD88 Adaptor Protein in B Cells Suppresses Protective Immunity during Salmonella Typhimurium Infection'. *Immunity* 33 (5): 777–90. https://doi.org/10.1016/j.immuni.2010.10.016.

Nguyen, Nga Thi Thuy, Bruno Contreras-Moreira, Jaime A. Castro-Mondragon, Walter Santana-Garcia, Raul Ossio, Carla Daniela Robles-Espinoza, Mathieu Bahin, et al. 2018. 'RSAT 2018: Regulatory Sequence Analysis Tools 20th Anniversary'. *Nucleic Acids Research* 46 (W1): W209–14. https://doi.org/10.1093/nar/gky317.

Nolte, Christiane, Marina Matyash, Tatjyana Pivneva, Carola G. Schipke, Carsten Ohlemeyer, Uwe-Karsten Hanisch, Frank Kirchhoff, and Helmut Kettenmann. 2001. 'GFAP Promoter-Controlled EGFP-Expressing Transgenic Mice: A Tool to Visualize Astrocytes and Astrogliosis in Living Brain Tissue'. *Glia* 33 (1): 72–86. https://doi.org/10.1002/1098-1136(20010101)33:1<72::AID-GLIA1007>3.0.CO;2-A.

Nora, Elphège P., Bryan R. Lajoie, Edda G. Schulz, Luca Giorgetti, Ikuhiro Okamoto, Nicolas Servant, Tristan Piolot, et al. 2012. 'Spatial Partitioning of the Regulatory Landscape of the X-Inactivation Centre'. *Nature* 485 (7398): 381–85. https://doi.org/10.1038/nature11049.

Nutt, Stephen L., Philip D. Hodgkin, David M. Tarlinton, and Lynn M. Corcoran. 2015. 'The Generation of Antibody-Secreting Plasma Cells'. *Nature Reviews Immunology* 15 (3): 160–71. https://doi.org/10.1038/nri3795.

Oberheim, Nancy Ann, Steven A. Goldman, and Maiken Nedergaard. 2012. 'Heterogeneity of Astrocytic Form and Function'. In *Astrocytes*, edited by Richard Milner, 814:23–45. Totowa, NJ: Humana Press. https://doi.org/10.1007/978-1-61779-452-0_3.

Okamura, T., K. Fujio, M. Shibuya, S. Sumitomo, H. Shoda, S. Sakaguchi, and K. Yamamoto. 2009. 'CD4+CD25-LAG3+ Regulatory T Cells Controlled by the Transcription Factor Egr-2'. *Proceedings of the National Academy of Sciences* 106 (33): 13974–79. https://doi.org/10.1073/pnas.0906872106.

Okano, M., D. W. Bell, D. A. Haber, and E. Li. 1999. 'DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for de Novo Methylation and Mammalian Development'. *Cell* 99 (3): 247–57.

Ooi, Steen K. T., Chen Qiu, Emily Bernstein, Keqin Li, Da Jia, Zhe Yang, Hediye Erdjument-Bromage, et al. 2007. 'DNMT3L Connects Unmethylated Lysine 4 of Histone H3 to de Novo Methylation of DNA'. *Nature* 448 (7154): 714–17. https://doi.org/10.1038/nature05987.

Oswald, J, S Engemann, N Lane, W Mayer, A Olek, R Fundele, W Dean, W Reik, and J Walter. 2000. 'Active Demethylation of the Paternal Genome in the Mouse Zygote'. *Current Biology* 10 (8): 475–78. https://doi.org/10.1016/S0960-9822(00)00448-6.

Perea, Gertrudis, Marta Navarrete, and Alfonso Araque. 2009. 'Tripartite Synapses: Astrocytes Process and Control Synaptic Information'. *Trends in Neurosciences* 32 (8): 421–31. https://doi.org/10.1016/j.tins.2009.05.001.

Phillips-Cremins, Jennifer E., Michael E. G. Sauria, Amartya Sanyal, Tatiana I. Gerasimova, Bryan R. Lajoie, Joshua S. K. Bell, Chin-Tong Ong, et al. 2013. 'Architectural Protein Subclasses Shape 3D Organization of Genomes during Lineage Commitment'. *Cell* 153 (6): 1281–95. https://doi.org/10.1016/j.cell.2013.04.053.

Piper, Michael, Guy Barry, John Hawkins, Sharon Mason, Charlotta Lindwall, Erica Little, Anindita Sarkar, et al. 2010. 'NFIA Controls Telencephalic Progenitor Cell Differentiation through Repression of the Notch Effector Hes1'. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 30 (27): 9127–39. https://doi.org/10.1523/JNEUROSCI.6167-09.2010.

Plath, Kathrin, Jia Fang, Susanna K. Mlynarczyk-Evans, Ru Cao, Kathleen A. Worringer, Hengbin Wang, Cecile C. de la Cruz, Arie P. Otte, Barbara Panning, and Yi Zhang. 2003. 'Role of Histone H3 Lysine 27 Methylation in X Inactivation'. *Science* 300 (5616): 131–35. https://doi.org/10.1126/science.1084274.

Poulin, Jean-Francois, Bosiljka Tasic, Jens Hjerling-Leffler, Jeffrey M. Trimarchi, and Rajeshwar Awatramani. 2016. 'Disentangling Neural Cell Diversity Using Single-Cell Transcriptomics'. *Nature Neuroscience* 19 (9): 1131–41. https://doi.org/10.1038/nn.4366.

Powell, M. J., S. A. Thompson, Y. Tone, H. Waldmann, and M. Tone. 2000. 'Posttranscriptional Regulation of IL-10 Gene Expression through Sequences in the 3'-Untranslated Region'. *Journal of Immunology (Baltimore, Md.: 1950)* 165 (1): 292–96.

R Development Core Team. 2008. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.

Rada-Iglesias, Alvaro, Ruchi Bajpai, Tomek Swigut, Samantha A. Brugmann, Ryan A. Flynn, and Joanna Wysocka. 2011. 'A Unique Chromatin Signature Uncovers Early Developmental Enhancers in Humans'. *Nature* 470 (7333): 279–83. https://doi.org/10.1038/nature09692.

Ramírez, Fidel, Friederike Dündar, Sarah Diehl, Björn A. Grüning, and Thomas Manke. 2014. 'DeepTools: A Flexible Platform for Exploring Deep-Sequencing Data'. *Nucleic Acids Research* 42 (Web Server issue): W187-191. https://doi.org/10.1093/nar/gku365.

Reeber, Stacey L., Marife Arancillo, and Roy V. Sillitoe. 2018. 'Bergmann Glia Are Patterned into Topographic Molecular Zones in the Developing and Adult Mouse Cerebellum'. *Cerebellum (London, England)* 17 (4): 392–403. https://doi.org/10.1007/s12311-014-0571-6.

Reichenbach, Andreas, Amin Derouiche, and Frank Kirchhoff. 2010. 'Morphology and Dynamics of Perisynaptic Glia'. *Brain Research Reviews* 63 (1–2): 11–25. https://doi.org/10.1016/j.brainresrev.2010.02.003.

Reik, Wolf, and Jörn Walter. 2001. 'Genomic Imprinting: Parental Influence on the Genome'. *Nature Reviews Genetics* 2 (1): 21–32. https://doi.org/10.1038/35047554.

Remak, Robert. 1852. *Ueber Extracellulare Entstehung Thierischer Zellen Und Über Vermehrung Derselben Durch Theilung*. Archiv Für Anatomie, Physiologie Und Wissenschaftliche Medicin. Berlin: Veit.

Riggs, A. D. 1975. 'X Inactivation, Differentiation, and DNA Methylation'. *Cytogenetics and Cell Genetics* 14 (1): 9–25. https://doi.org/10.1159/000130315.

Roadmap Epigenomics Consortium, Anshul Kundaje, Wouter Meuleman, Jason Ernst, Misha Bilenky, Angela Yen, Alireza Heravi-Moussavi, et al. 2015. 'Integrative Analysis of 111 Reference Human Epigenomes'. *Nature* 518 (7539): 317–30. https://doi.org/10.1038/nature14248.

Robertson, Gordon, Martin Hirst, Matthew Bainbridge, Misha Bilenky, Yongjun Zhao, Thomas Zeng, Ghia Euskirchen, et al. 2007. 'Genome-Wide Profiles of STAT1 DNA Association Using Chromatin Immunoprecipitation and Massively Parallel Sequencing'. *Nature Methods* 4 (8): 651–57. https://doi.org/10.1038/nmeth1068.

Robinson, James T., Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S. Lander, Gad Getz, and Jill P. Mesirov. 2011. 'Integrative Genomics Viewer'. *Nature Biotechnology* 29 (January): 24–26. https://doi.org/10.1038/nbt.1754.

Robinson, M. D., D. J. McCarthy, and G. K. Smyth. 2010. 'EdgeR: A Bioconductor Package for Differential Expression Analysis of Digital Gene Expression Data'. *Bioinformatics* 26 (1): 139–40. https://doi.org/10.1093/bioinformatics/btp616.

Robinson, Mark D, and Alicia Oshlack. 2010. 'A Scaling Normalization Method for Differential Expression Analysis of RNA-Seq Data'. *Genome Biology* 11 (3): R25. https://doi.org/10.1186/gb-2010-11-3-r25.

Rosa, Stefanie, and Peter Shaw. 2013. 'Insights into Chromatin Structure and Dynamics in Plants'. *Biology* 2 (4): 1378–1410. https://doi.org/10.3390/biology2041378.

Rubenstein, John L. R, and Pasko Rakic. 2013. *Patterning and Cell Type Specification in the Developing CNS and PNS*. Amsterdam; Boston: Elsevier/AP. http://site.ebrary.com/id/10698618.

Sakers, Kristina, Allison M. Lake, Rohan Khazanchi, Rebecca Ouwenga, Michael J. Vasek, Adish Dani, and Joseph D. Dougherty. 2017. 'Astrocytes Locally Translate Transcripts in Their Peripheral Processes'. *Proceedings of the National Academy of Sciences*, April, 201617782. https://doi.org/10.1073/pnas.1617782114.

Salhab, Abdulrahman, Karl Nordström, Gilles Gasparoni, Kathrin Kattler, Peter Ebert, Fidel Ramirez, Laura Arrigoni, et al. 2018. 'A Comprehensive Analysis of 195 DNA Methylomes Reveals Shared and Cell-Specific Features of Partially Methylated Domains'. *Genome Biology* 19 (1): 150. https://doi.org/10.1186/s13059-018-1510-5.

Sander, Jeffry D., and J. Keith Joung. 2014. 'CRISPR-Cas Systems for Editing, Regulating and Targeting Genomes'. *Nature Biotechnology* 32 (4): 347–55. https://doi.org/10.1038/nbt.2842.

Sanosaka, Tsukasa, Takuya Imamura, Nobuhiko Hamazaki, MuhChyi Chai, Katsuhide Igarashi, Maky Ideta-Otsuka, Fumihito Miura, et al. 2017. 'DNA Methylome Analysis Identifies Transcription Factor-Based Epigenomic Signatures of Multilineage Competence in Neural Stem/Progenitor Cells'. *Cell Reports* 20 (12): 2992–3003. https://doi.org/10.1016/j.celrep.2017.08.086.

Saraiva, Margarida, Jillian R. Christensen, Alla V. Tsytsykova, Anne E. Goldfeld, Steven C. Ley, Dimitris Kioussis, and Anne O'Garra. 2005. 'Identification of a Macrophage-Specific Chromatin Signature in the IL-10 Locus'. *Journal of Immunology (Baltimore, Md.: 1950)* 175 (2): 1041–46.

Sato, Takahiko, Noriaki Sasai, and Yoshiki Sasai. 2005. 'Neural Crest Determination by Co-Activation of Pax3 and Zic1 Genes in Xenopus Ectoderm'. *Development* 132 (10): 2355–63. https://doi.org/10.1242/dev.01823.

Saxonov, Serge, Paul Berg, and Douglas L. Brutlag. 2006. 'A Genome-Wide Analysis of CpG Dinucleotides in the Human Genome Distinguishes Two Distinct Classes of Promoters'. *Proceedings of the National Academy of Sciences of the United States of America* 103 (5): 1412–17. https://doi.org/10.1073/pnas.0510310103.

Schleiden, M. J., Theodor Schwann, and Henry Smith. 1847. *Microscopical Researches into the Accordance in the Structure and Growth of Animals and Plants. Translated from the German of Theodor Schwann, by Henry Smith.* London,: Sydenham Society,. https://doi.org/10.5962/bhl.title.17276.

Schmidt, Florian, Nina Gasparoni, Gilles Gasparoni, Kathrin Gianmoena, Cristina Cadenas, Julia K. Polansky, Peter Ebert, et al. 2017. 'Combining Transcription Factor Binding Affinities with Open-Chromatin Data for Accurate Gene Expression Prediction'. *Nucleic Acids Research* 45 (1): 54–66. https://doi.org/10.1093/nar/gkw1061.

Schones, Dustin E., Kairong Cui, Suresh Cuddapah, Tae-Young Roh, Artem Barski, Zhibin Wang, Gang Wei, and Keji Zhao. 2008. 'Dynamic Regulation of Nucleosome Positioning in the Human Genome'. *Cell* 132 (5): 887–98. https://doi.org/10.1016/j.cell.2008.02.022.

Schübeler, Dirk. 2015. 'Function and Information Content of DNA Methylation'. *Nature* 517 (7534): 321–26. https://doi.org/10.1038/nature14192.

Schulz, W. A., C. Steinhoff, and A. R. Florl. 2006. 'Methylation of Endogenous Human Retroelements in Health and Disease'. *Current Topics in Microbiology and Immunology* 310: 211–50.

Schumacher, Anne, Stefanie Ehrentraut, Markus Scharm, Hongsheng Wang, Roland Hartig, Herbert C. Morse, and Ana Claudia Zenclussen. 2018. 'Plasma Cell Alloantigen 1 and IL-10 Secretion Define Two Distinct Peritoneal B1a B Cell Subsets With Opposite Functions, PC1high Cells Being Protective and PC1low

Cells Harmful for the Growing Fetus'. *Frontiers in Immunology* 9: 1045. https://doi.org/10.3389/fimmu.2018.01045.

Shah, Shivanee, and Liang Qiao. 2008. 'Resting B Cells Expand a CD4+CD25+Foxp3+ Treg Population via TGF-Beta3'. *European Journal of Immunology* 38 (9): 2488–98. https://doi.org/10.1002/eji.200838201.

Shannon, Paul, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. 2003. 'Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks'. *Genome Research* 13 (11): 2498–2504. https://doi.org/10.1101/gr.1239303.

Shayevitch, Ronna, Dan Askayo, Ifat Keydar, and Gil Ast. 2018. 'The Importance of DNA Methylation of Exons on Alternative Splicing'. *RNA* 24 (10): 1351–62. https://doi.org/10.1261/rna.064865.117.

Shen, Ping, and Simon Fillatreau. 2015. 'Suppressive Functions of B Cells in Infectious Diseases'. *International Immunology* 27 (10): 513–19. https://doi.org/10.1093/intimm/dxv037.

Silva, Jose, Winifred Mak, Ilona Zvetkova, Ruth Appanah, Tatyana B. Nesterova, Zoe Webster, Antoine H. F. M. Peters, Thomas Jenuwein, Arie P. Otte, and Neil Brockdorff. 2003. 'Establishment of Histone H3 Methylation on the Inactive X Chromosome Requires Transient Recruitment of Eed-Enx1 Polycomb Group Complexes'. *Developmental Cell* 4 (4): 481–95. https://doi.org/10.1016/S1534-5807(03)00068-6.

Sofueva, Sevil, Eitan Yaffe, Wen-Ching Chan, Dimitra Georgopoulou, Matteo Vietri Rudan, Hegias Mira-Bontenbal, Steven M Pollard, Gary P Schroth, Amos Tanay, and Suzana Hadjur. 2013. 'Cohesin-Mediated Interactions Organize Chromosomal Domain Architecture: Functional Role for Cohesin in Chromosome Domain Structure'. *The EMBO Journal* 32 (24): 3119–29. https://doi.org/10.1038/emboj.2013.237.

Spitz, François, and Eileen E. M. Furlong. 2012. 'Transcription Factors: From Enhancer Binding to Developmental Control'. *Nature Reviews Genetics* 13 (9): 613–26. https://doi.org/10.1038/nrg3207.

Srinivasan, Rahul, Tsai-Yi Lu, Hua Chai, Ji Xu, Ben S. Huang, Peyman Golshani, Giovanni Coppola, and Baljit S. Khakh. 2016. 'New Transgenic Mouse Lines for Selectively Targeting Astrocytes and for Studying Calcium Signals in Astrocyte Processes in Situ and in Vivo'. *Neuron* 92 (6): 1181–95. https://doi.org/10.1016/j.neuron.2016.11.030.

Stadler, Michael B., Rabih Murr, Lukas Burger, Robert Ivanek, Florian Lienert, Anne Schöler, Erik van Nimwegen, et al. 2011. 'DNA-Binding Factors Shape the Mouse Methylome at Distal Regulatory Regions'. *Nature* 480 (7378): 490–95. https://doi.org/10.1038/nature10716.

Stroud, Hume, Suhua Feng, Shannon Morey Kinney, Sriharsa Pradhan, and Steven E Jacobsen. 2011. '5-Hydroxymethylcytosine Is Associated with Enhancers and Gene Bodies in Human Embryonic Stem Cells'. *Genome Biology* 12 (6): R54. https://doi.org/10.1186/gb-2011-12-6-r54.

Suetake, Isao, Fuminori Shinozaki, Junichi Miyagawa, Hideyuki Takeshima, and Shoji Tajima. 2004. 'DNMT3L Stimulates the DNA Methylation Activity of Dnmt3a

and Dnmt3b through a Direct Interaction'. *The Journal of Biological Chemistry* 279 (26): 27816–23. https://doi.org/10.1074/jbc.M400181200.

Supek, Fran, Matko Bošnjak, Nives Škunca, and Tomislav Šmuc. 2011. 'REVIGO Summarizes and Visualizes Long Lists of Gene Ontology Terms'. Edited by Cynthia Gibas. *PLoS ONE* 6 (7): e21800. https://doi.org/10.1371/journal.pone.0021800.

Sur, Mriganka, and John L. R. Rubenstein. 2005. 'Patterning and Plasticity of the Cerebral Cortex'. *Science* 310 (5749): 805–10. https://doi.org/10.1126/science.1112070.

Szklarczyk, Damian, John H. Morris, Helen Cook, Michael Kuhn, Stefan Wyder, Milan Simonovic, Alberto Santos, et al. 2017. 'The STRING Database in 2017: Quality-Controlled Protein-Protein Association Networks, Made Broadly Accessible'. *Nucleic Acids Research* 45 (D1): D362–68. https://doi.org/10.1093/nar/gkw937.

Tahiliani, Mamta, Kian Peng Koh, Yinghua Shen, William A. Pastor, Hozefa Bandukwala, Yevgeny Brudno, Suneet Agarwal, et al. 2009. 'Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1'. *Science (New York, N.Y.)* 324 (5929): 930–35. https://doi.org/10.1126/science.1170116.

Takahashi, Kazutoshi, and Shinya Yamanaka. 2006. 'Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors'. *Cell* 126 (4): 663–76. https://doi.org/10.1016/j.cell.2006.07.024.

Takizawa, Takumi, Kinichi Nakashima, Masakazu Namihira, Wataru Ochiai, Atsumi Uemura, Makoto Yanagisawa, Naoyuki Fujita, Mitsuyoshi Nakao, and Tetsuya Taga. 2001. 'DNA Methylation Is a Critical Cell-Intrinsic Determinant of Astrocyte Differentiation in the Fetal Brain'. *Developmental Cell* 1 (6): 749–58. https://doi.org/10.1016/S1534-5807(01)00101-0.

Tasic, Bosiljka, Boaz P. Levi, and Vilas Menon. 2017. 'Single-Cell Transcriptomic Characterization of Vertebrate Brain Composition, Development, and Function'. In *Decoding Neural Circuit Structure and Function: Cellular Dissection Using Genetic Model Organisms*, edited by Arzu Çelik and Mathias F. Wernet, 437–68. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-57363-2_18.

The ENCODE Project Consortium. 2012. 'An Integrated Encyclopedia of DNA Elements in the Human Genome'. *Nature* 489 (7414): 57–74. https://doi.org/10.1038/nature11247.

Thomas-Chollier, Morgane, Andrew Hufton, Matthias Heinig, Sean O'Keeffe, Nassim El Masri, Helge G Roider, Thomas Manke, and Martin Vingron. 2011. 'Transcription Factor Binding Predictions Using TRAP for the Analysis of ChIP-Seq Data and Regulatory SNPs'. *Nature Protocols* 6 (12): 1860–69. https://doi.org/10.1038/nprot.2011.409.

Thorvaldsdottir, H., J. T. Robinson, and J. P. Mesirov. 2013. 'Integrative Genomics Viewer (IGV): High-Performance Genomics Data Visualization and Exploration'. *Briefings in Bioinformatics* 14 (2): 178–92. https://doi.org/10.1093/bib/bbs017.

Tiwari, Neha, Abhijeet Pataskar, Sophie Péron, Sudhir Thakurela, Sanjeeb Kumar Sahu, María Figueres-Oñate, Nicolás Marichal, Laura López-Mascaraque, Vijay K. Tiwari, and Benedikt Berninger. 2018. 'Stage-Specific Transcription Factors

Drive Astrogliogenesis by Remodeling Gene Regulatory Landscapes'. *Cell Stem Cell* 23 (4): 557-571.e8. https://doi.org/10.1016/j.stem.2018.09.008.

Townsend, S. E., and C. C. Goodnow. 1998. 'Abortive Proliferation of Rare T Cells Induced by Direct or Indirect Antigen Presentation by Rare B Cells in Vivo'. *The Journal of Experimental Medicine* 187 (10): 1611–21. https://doi.org/10.1084/jem.187.10.1611.

Trevant, Brune, Tripti Gaur, Sadiq Hussain, John Symons, Barry S. Komm, Peter V.N. Bodine, Gary S. Stein, and Jane B. Lian. 2008. 'Expression of Secreted Frizzled Related Protein 1, A Wnt Antagonist, in Brain, Kidney, and Skeleton Is Dispensable for Normal Embryonic Development'. *Journal of Cellular Physiology* 217 (1): 113–26. https://doi.org/10.1002/jcp.21482.

Tsai, Hui-Hsin, Huiliang Li, Luis C. Fuentealba, Anna V. Molofsky, Raquel Taveira-Marques, Helin Zhuang, April Tenney, et al. 2012. 'Regional Astrocyte Allocation Regulates CNS Synaptogenesis and Repair'. *Science (New York, N.Y.)* 337 (6092): 358–62. https://doi.org/10.1126/science.1222381.

Valinluck, Victoria, Hsin-Hao Tsai, Daniel K. Rogstad, Artur Burdzy, Adrian Bird, and Lawrence C. Sowers. 2004. 'Oxidative Damage to Methyl-CpG Sequences Inhibits the Binding of the Methyl-CpG Binding Domain (MBD) of Methyl-CpG Binding Protein 2 (MeCP2)'. *Nucleic Acids Research* 32 (14): 4100–4108. https://doi.org/10.1093/nar/gkh739.

Valouev, Anton, Steven M. Johnson, Scott D. Boyd, Cheryl L. Smith, Andrew Z. Fire, and Arend Sidow. 2011. 'Determinants of Nucleosome Organization in Primary Human Cells'. *Nature* 474 (7352): 516–20. https://doi.org/10.1038/nature10002.

Vaquerizas, Juan M., Sarah K. Kummerfeld, Sarah A. Teichmann, and Nicholas M. Luscombe. 2009. 'A Census of Human Transcription Factors: Function, Expression and Evolution'. *Nature Reviews. Genetics* 10 (4): 252–63. https://doi.org/10.1038/nrg2538.

Vickaryous, Matthew K., and Brian K. Hall. 2006. 'Human Cell Type Diversity, Evolution, Development, and Classification with Special Reference to Cells Derived from the Neural Crest'. *Biological Reviews* 81 (03): 425. https://doi.org/10.1017/S1464793106007068.

Visel, Axel, Matthew J. Blow, Zirong Li, Tao Zhang, Jennifer A. Akiyama, Amy Holt, Ingrid Plajzer-Frick, et al. 2009. 'ChIP-Seq Accurately Predicts Tissue-Specific Activity of Enhancers'. *Nature* 457 (7231): 854–58. https://doi.org/10.1038/nature07730.

Waddington, Conrad Hal. 1957. *The Strategy of the Genes : A Discussion of Some Aspects of Theoretical Biology*. London: Allen & Unwin.

Walz, Wolfgang. 2000. 'Role of Astrocytes in the Clearance of Excess Extracellular Potassium'. *Neurochemistry International* 36 (4): 291–300. https://doi.org/10.1016/S0197-0186(99)00137-0.

Wang, H., D.-M. Shin, S. Abbasi, S. Jain, A. L. Kovalchuk, N. Beaty, S. Chen, I. Gonzalez-Garcia, and H. C. Morse. 2012. 'Expression of Plasma Cell Alloantigen 1 Defines Layered Development of B-1a B-Cell Subsets with Distinct Innate-like Functions'. *Proceedings of the National Academy of Sciences* 109 (49): 20077–82. https://doi.org/10.1073/pnas.1212428109.

Wang, Hong-Qiang, Lindsey K. Tuominen, and Chung-Jui Tsai. 2011. 'SLIM: A Sliding Linear Model for Estimating the Proportion of True Null Hypotheses in Datasets

with Dependence Structures'. *Bioinformatics (Oxford, England)* 27 (2): 225–31. https://doi.org/10.1093/bioinformatics/btq650.

Wang, Zheng-Yu, Hiroshi Sato, Saritha Kusam, Sarita Sehra, Lisa M. Toney, and Alexander L. Dent. 2005. 'Regulation of IL-10 Gene Expression in Th2 Cells by Jun Proteins'. *Journal of Immunology (Baltimore, Md.: 1950)* 174 (4): 2098–2105.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Second edition. Use R! Cham: Springer.

Winchenbach, Jan, Tim Düking, Stefan A. Berghoff, Sina K. Stumpf, Swen Hülsmann, Klaus-Armin Nave, and Gesine Saher. 2016. 'Inducible Targeting of CNS Astrocytes in Aldh1l1-CreERT2 BAC Transgenic Mice'. *F1000Research* 5 (December). https://doi.org/10.12688/f1000research.10509.1.

Wolf, S. D. 1996. 'Experimental Autoimmune Encephalomyelitis Induction in Genetically B Cell-Deficient Mice'. *Journal of Experimental Medicine* 184 (6): 2271–78. https://doi.org/10.1084/jem.184.6.2271.

Xie, Wei, Matthew D. Schultz, Ryan Lister, Zhonggang Hou, Nisha Rajagopal, Pradipta Ray, John W. Whitaker, et al. 2013. 'Epigenomic Analysis of Multilineage Differentiation of Human Embryonic Stem Cells'. *Cell* 153 (5): 1134–48. https://doi.org/10.1016/j.cell.2013.04.022.

Yamada, Keiko, and Masahiko Watanabe. 2002. 'Cytodifferentiation of Bergmann Glia and Its Relationship with Purkinje Cells'. *Anatomical Science International* 77 (2): 94–108. https://doi.org/10.1046/j.0022-7722.2002.00021.x.

Yao, Lijing, Benjamin P. Berman, and Peggy J. Farnham. 2015. 'Demystifying the Secret Mission of Enhancers: Linking Distal Regulatory Elements to Target Genes'. *Critical Reviews in Biochemistry and Molecular Biology* 50 (6): 550–73. https://doi.org/10.3109/10409238.2015.1087961.

Yoshida, Ryoko, Mayu Suzuki, Ryota Sakaguchi, Eiichi Hasegawa, Akihiro Kimura, Takashi Shichita, Takashi Sekiya, Hiroshi Shiraishi, Kouji Shimoda, and Akihiko Yoshimura. 2012. 'Forced Expression of Stabilized C-Fos in Dendritic Cells Reduces Cytokine Production and Immune Responses in Vivo'. *Biochemical and Biophysical Research Communications* 423 (2): 247–52. https://doi.org/10.1016/j.bbrc.2012.05.097.

Yuan, Guo-Cheng, Yuen-Jong Liu, Michael F. Dion, Michael D. Slack, Lani F. Wu, Steven J. Altschuler, and Oliver J. Rando. 2005. 'Genome-Scale Identification of Nucleosome Positions in S. Cerevisiae'. *Science (New York, N.Y.)* 309 (5734): 626–30. https://doi.org/10.1126/science.1112178.

Zeisel, Amit, Hannah Hochgerner, Peter Lönnerberg, Anna Johnsson, Fatima Memic, Job van der Zwan, Martin Häring, et al. 2018. 'Molecular Architecture of the Mouse Nervous System'. *Cell* 174 (4): 999-1014.e22. https://doi.org/10.1016/j.cell.2018.06.021.

Zhang, Y., K. Chen, S. A. Sloan, M. L. Bennett, A. R. Scholze, S. O'Keeffe, H. P. Phatnani, et al. 2014. 'An RNA-Sequencing Transcriptome and Splicing Database of Glia, Neurons, and Vascular Cells of the Cerebral Cortex'. *Journal of Neuroscience* 34 (36): 11929–47. https://doi.org/10.1523/JNEUROSCI.1860-14.2014.

Zhang, Yingying, Renata Jurkowska, Szabolcs Soeroes, Arumugam Rajavelu, Arunkumar Dhayalan, Ina Bock, Philipp Rathert, et al. 2010. 'Chromatin Methylation Activity of Dnmt3a and Dnmt3a/3L Is Guided by Interaction of the

ADD Domain with the Histone H3 Tail'. *Nucleic Acids Research* 38 (13): 4246–53. https://doi.org/10.1093/nar/gkq147.

Zhang, Yong, Tao Liu, Clifford A. Meyer, Jérôme Eeckhoute, David S. Johnson, Bradley E. Bernstein, Chad Nusbaum, et al. 2008. 'Model-Based Analysis of ChIP-Seq (MACS)'. *Genome Biology* 9 (9): R137. https://doi.org/10.1186/gb-2008-9-9-r137.

Zheng, Yan, Yuanyuan Zha, Robbert M. Spaapen, Rebecca Mathew, Kenneth Barr, Albert Bendelac, and Thomas F. Gajewski. 2013. 'Egr2-Dependent Gene Expression Profiling and ChIP-Seq Reveal Novel Biologic Targets in T Cell Anergy'. *Molecular Immunology* 55 (3): 283–91. https://doi.org/10.1016/j.molimm.2013.03.006.

Zibetti, Cristina, Sheng Liu, Jun Wan, Jiang Qian, and Seth Blackshaw. 2017. 'Lhx2 Regulates Temporal Changes in Chromatin Accessibility and Transcription Factor Binding in Retinal Progenitor Cells.' *BioRxiv*, December. https://doi.org/10.1101/238279.

Ziller, Michael J., Reuven Edri, Yakey Yaffe, Julie Donaghey, Ramona Pop, William Mallard, Robbyn Issner, et al. 2015. 'Dissecting Neural Differentiation Regulatory Networks through Epigenetic Footprinting'. *Nature* 518 (7539): 355–59. https://doi.org/10.1038/nature13990.

Ziller, Michael J., Hongcang Gu, Fabian Müller, Julie Donaghey, Linus T.-Y. Tsai, Oliver Kohlbacher, Philip L. De Jager, et al. 2013. 'Charting a Dynamic DNA Methylation Landscape of the Human Genome'. *Nature* 500 (7463): 477–81. https://doi.org/10.1038/nature12433.

Zuin, J., J. R. Dixon, M. I. J. A. van der Reijden, Z. Ye, P. Kolovos, R. W. W. Brouwer, M. P. C. van de Corput, et al. 2014. 'Cohesin and CTCF Differentially Affect Chromatin Architecture and Gene Expression in Human Cells'. *Proceedings of the National Academy of Sciences* 111 (3): 996–1001. https://doi.org/10.1073/pnas.1317788111.

Zweifel, Stefan, Guillaume Marcy, Quentin Lo Guidice, Deqiang Li, Christophe Heinrich, Kasum Azim, and Olivier Raineteau. 2018. 'HOPX Defines Heterogeneity of Postnatal Subventricular Zone Neural Stem Cells'. *Stem Cell Reports* 11 (3): 770–83. https://doi.org/10.1016/j.stemcr.2018.08.006.

E. Appendix

	Description	Hat1	Prmt2	Ino80c	Hmgb1	Hmgn1	Anp32e	Npm1	Nap1l1	Asf1a	Smc3	Stag1	Rad21	Ctcf
ACTE1	Telencephalon astrocytes, fibrous	0.0151	0.123	0.0298	0.209	0.621	0.154	0.159	0.130	0.0887	0.101	0.0504	0.139	0.0663
ACTE2	Telencephalon astrocytes, protoplasmic	0.0115	0.0902	0.0204	0.118	0.434	0.0568	0.132	0.0739	0.0979	0.0866	0.0438	0.0608	0.0385
ACOB	Olfactory astrocytes	0.0168	0.138	0.0172	0.167	0.684	0.149	0.186	0.134	0.104	0.0969	0.0665	0.101	0.0586
ACNT1	Non-telencephalon astrocytes, protoplasmic	0.0122	0.0541	0.0116	0.0836	0.253	0.0384	0.0851	0.0369	0.0788	0.0585	0.0291	0.0449	0.0270
ACNT2	Non-telencephalon astrocytes, fibrous	0.0262	0.0902	0.0155	0.111	0.455	0.149	0.112	0.119	0.0541	0.0703	0.0605	0.129	0.0419
ACMB	Dorsal midbrain Myoc-expressing astrocyte-like	0.0557	0.232	0.0381	0.251	1.06	0.195	0.278	0.255	0.0968	0.129	0.0648	0.209	0.0685
ACBG	Bergmann glia	0.0326	0.253	0.0271	0.286	0.791	0.0730	0.232	0.146	0.151	0.170	0.0734	0.0912	0.0578

Figure 4-1 Single cell gene expression of genes involved in chromatin organization. *Related to* ***Table 17****. Displayed data were obtained from http://mousebrain.org/ provided by the Linnarsson Lab (Zeisel et al., 2018). Values represent the mean expression relative to the highest detected expression value. ACTE2 and ACBG cell types correspond to the CTX and CB astrocytes from this thesis, respectively.*

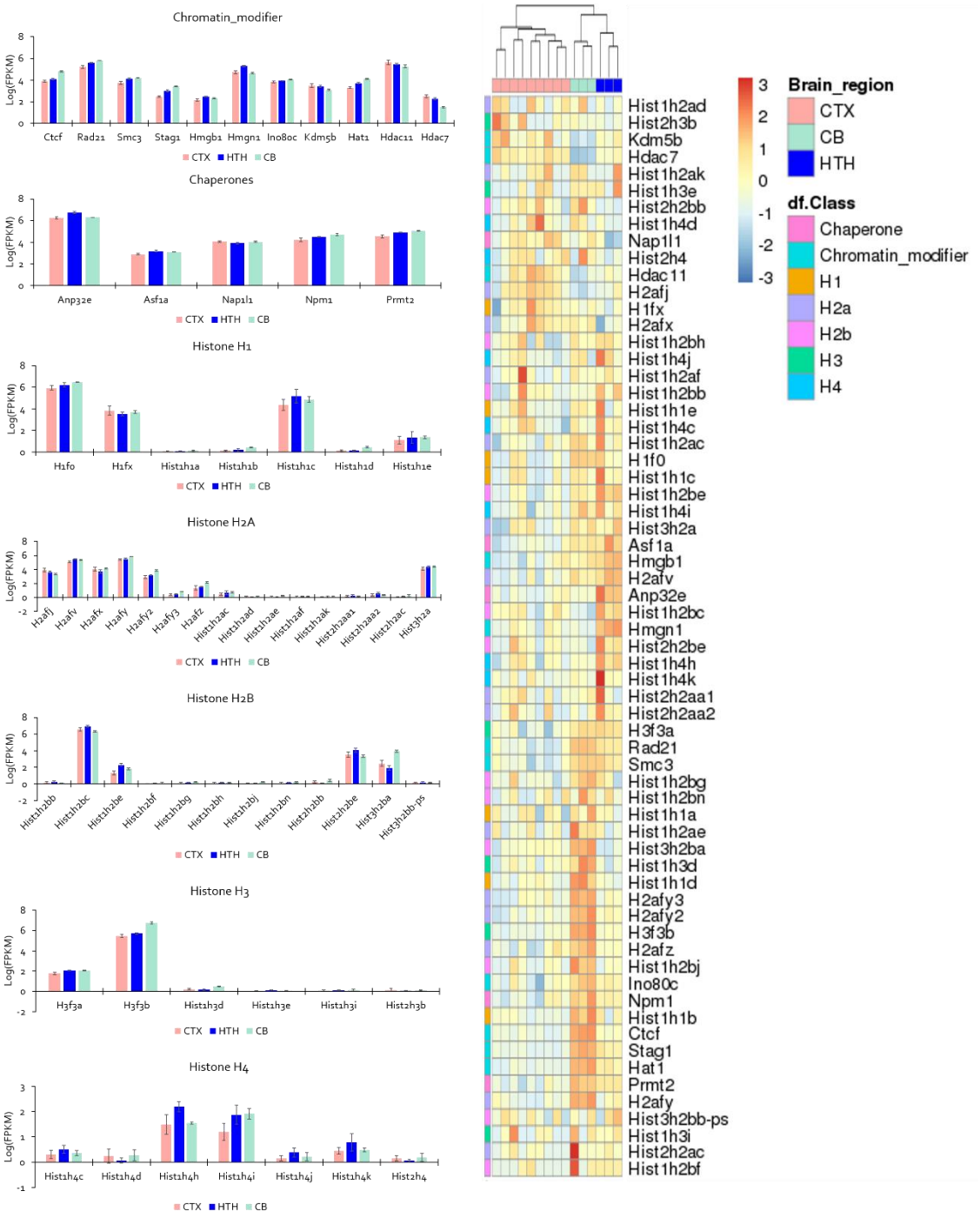


Figure 4-2 Gene expression of chromatin-organizer and -modifier, and histones from Boisvert et al. 2018*. On the left the mean ± STDEV of the FPKM are shown. On the right expression of the same genes are represented as row-Z-scores. CTX= Mean (visual + motor + somatosensory cortex); HTH= Hypothalamus; CB= Cerebellum.*

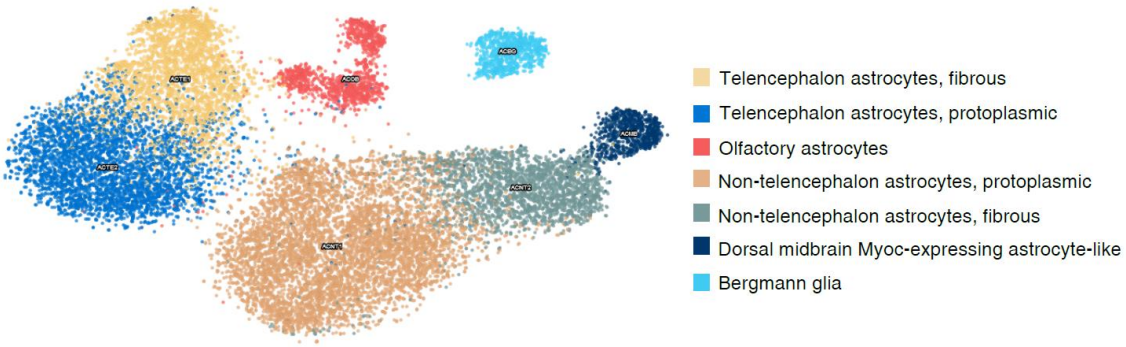


Figure 4-3 Major clusters of single astrocytes represented in a T-distributed Stochastic Neighbor Embedding (gt-SNE plot), colored by cluster identity (top). *Figure was obtained from the loom viewer provided by the Linnarsson Lab (http://loom.linnarssonlab.org/) (Zeisel et al., 2018).*

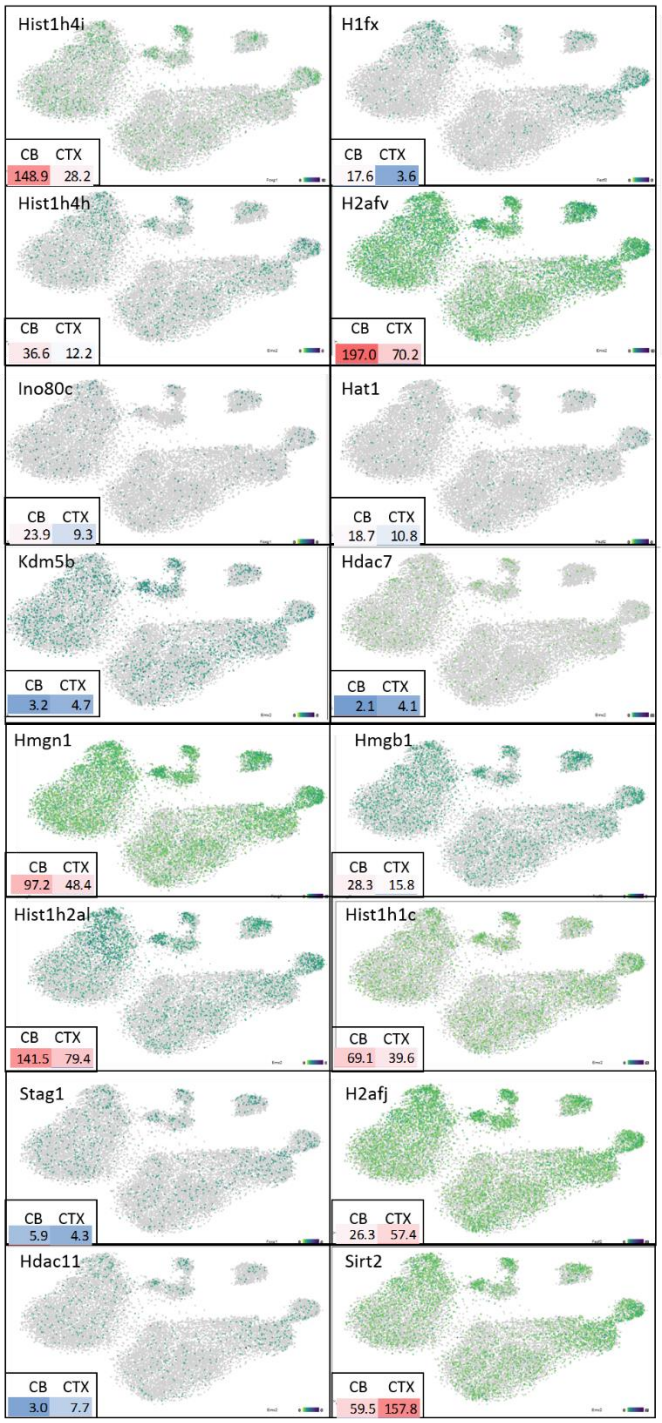


Figure 4-4 Single cell gene expression of selected genes related to chromatin organization found to be differentially expressed between CB and CTX astrocytes shown on the gt-SNE layout from Figure 4-3 Major clusters of single astrocytes represented in a T-distributed Stochastic Neighbor Embedding (gt-SNE plot), colored by cluster identity (top). *Figure was obtained from the loom viewer provided by the Linnarsson Lab (http://loom.linnarssonlab.org/) (Zeisel et al., 2018).* ***Figure 4-3****. Boxes in the lower left corner display expression values of respective gene obtained from this study. Figures were obtained from the loom viewer provided by the Linnarsson Lab (http://loom.linnarssonlab.org/) (Zeisel et al., 2018).*

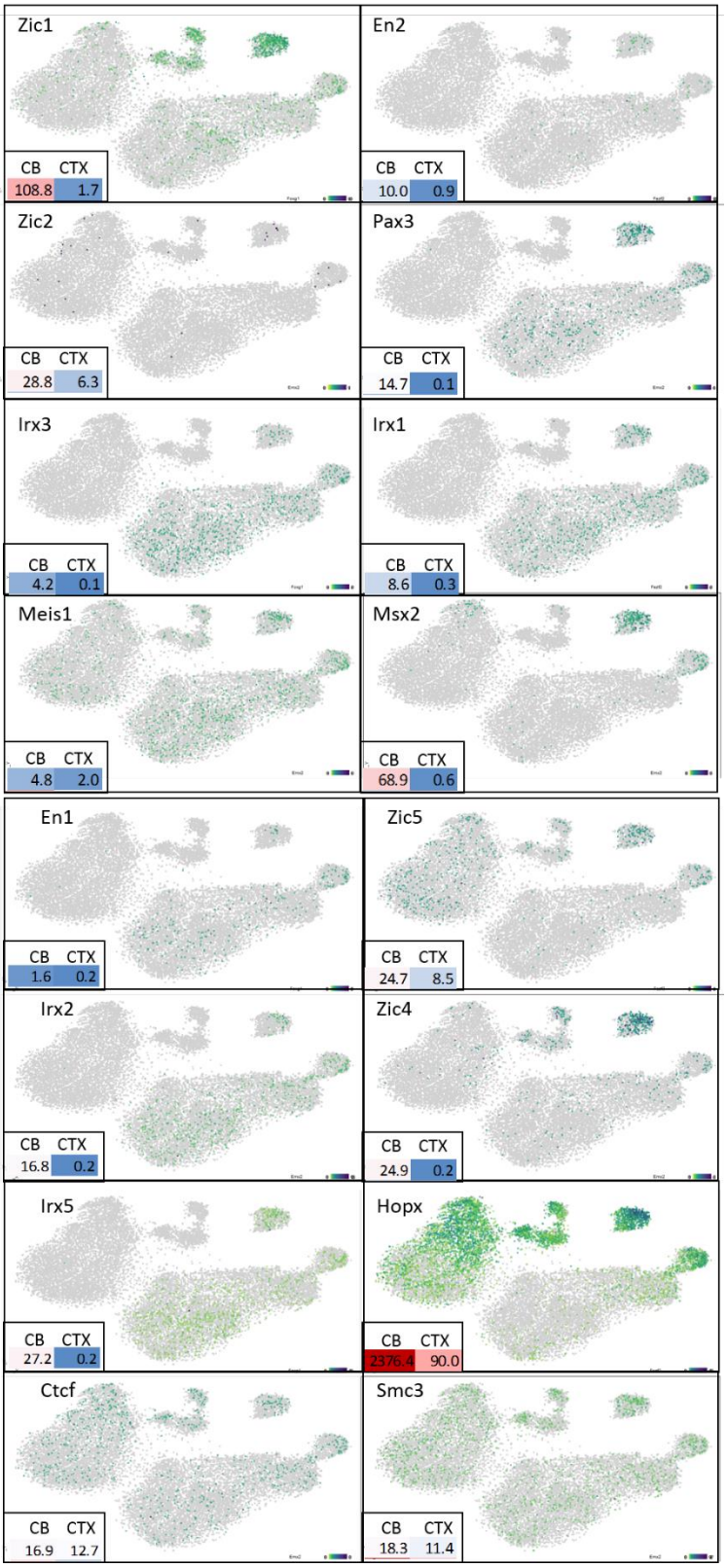


Figure 4-5 Single cell gene expression of selected genes found to be higher expressed in CB astrocytes shown on the gt-SNE layout from Figure 4-3. *Boxes in the lower left corner display the expression values of respective gene obtained from this study. Figures were obtained from the loom viewer provided by the Linnarsson Lab (http://loom.linnarssonlab.org/) (Zeisel et al., 2018).*

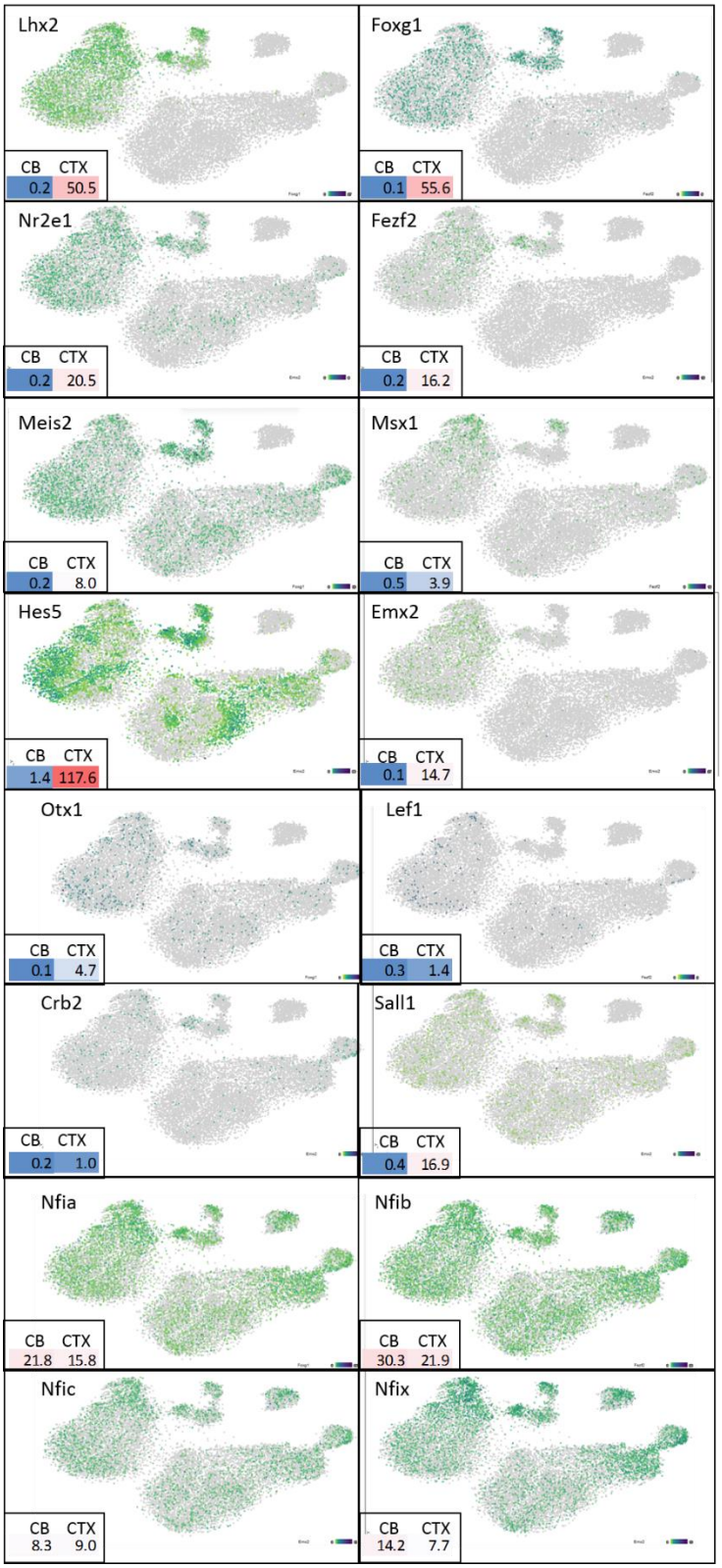


Figure 4-6 Single cell gene expression of selected genes found to be higher expressed in CTX astrocytes, and Nfi family members shown on the gt-SNE layout from Figure 4-3. *Boxes in the lower left corner display the expression values of respective gene obtained from this study. Figures were obtained from the loom viewer provided by the Linnarsson Lab (http://loom.linnarssonlab.org/) (Zeisel et al., 2018).*

Figure 4-7 In situ hybridization of early specification factors in mouse E11.5 embryos. *Colored based on the stronger expression in CB or CTX. Circles indicate cortical (reddish) and cerebellar (green) brain regions. Images were obtained from the Allen Developing Mouse Brain Atlas.*

Table 22 Mean gene expression of immune globulins in the analyzed plasma cell subtypes.

Symbol	Lag3+	Lag3+ day1	Lag3-	Lag3- day1
Ighe	2.61	4.55	78.45	140.41
Ighg2b	6070.52	1585.17	39132.76	35671.95
Ighg2c	3713.04	1444.17	30590.01	31379.09
Ighg3	792.35	430.55	5743.51	7381.58
Ighj1	15.82	38.05	144.14	78.26
Ighj3	8.70	38.94	118.62	161.27
Ighj4	12.35	34.87	151.31	171.55
Ighm	166413.97	257023.15	86381.97	80636.36
Ighv10-1	13.41	80.76	100.71	180.79
Ighv10-3	11.08	34.42	125.73	147.65
Ighv1-12	7.12	24.77	90.26	119.37
Ighv1-18	28.88	36.17	288.40	384.63
Ighv1-19	28.78	46.75	364.49	292.54
Ighv1-20	2.31	4.66	22.87	32.19
Ighv1-22	21.40	30.77	177.14	331.36
Ighv12-3	1.14	3.25	11.21	29.44
Ighv1-26	73.69	123.40	586.24	1051.85
Ighv1-31	7.04	14.62	78.76	110.79
Ighv13-2	1.11	8.56	16.66	32.62
Ighv1-34	11.80	20.88	138.94	129.01
Ighv1-36	3.82	5.16	11.25	23.83
Ighv1-37	0.75	2.73	3.81	62.24
Ighv1-39	41.34	64.63	293.09	560.16
Ighv1-4	2.40	9.13	39.15	45.14
Ighv14-1	3.28	9.88	36.07	62.95
Ighv14-2	7.23	21.07	67.45	117.24
Ighv1-42	4.58	10.39	66.81	115.02
Ighv1-77	4.84	8.05	47.17	52.65
Ighv1-78	12.36	46.61	156.85	290.81
Ighv1-80	10.36	14.18	94.87	174.44
Ighv1-81	17.38	18.30	115.44	205.50
Ighv1-82	18.65	28.26	164.05	338.82
Ighv1-84	3.47	13.35	66.01	85.76
Ighv1-85	4.00	11.37	53.06	70.94
Ighv1-9	22.94	42.25	281.77	397.77
Ighv2-2	8.41	41.71	143.25	173.01
Ighv2-3	7.36	27.41	92.52	183.91
Ighv2-4	4.13	6.98	45.03	26.04
Ighv2-5	2.61	13.27	29.99	42.79
Ighv2-6	2.37	25.23	18.37	127.31
Ighv2-6-8	0.46	4.22	4.67	13.80
Ighv2-9	2.23	16.26	16.56	36.07
Ighv2-9-1	4.32	15.68	58.36	97.51
Ighv3-1	10.27	27.63	77.38	95.97
Ighv3-3	0.62	0.92	3.44	2.32
Ighv3-5	1.96	8.28	17.61	24.84
Ighv3-6	58.43	117.98	548.66	867.30
Ighv3-8	9.48	18.03	108.40	121.97
Ighv4-1	20.83	90.34	108.30	332.50

Symbol	Lag3+	Lag3+ day1	Lag3-	Lag3- day1
Ighv14-3	10.42	24.83	115.32	154.19
Ighv1-43	0.51	4.21	8.58	9.51
Ighv14-4	6.07	13.60	46.18	98.09
Ighv1-47	1.51	6.49	47.86	64.41
Ighv1-49	1.15	0.00	3.87	6.61
Ighv1-5	7.26	7.90	55.68	162.33
Ighv1-50	3.88	7.06	47.90	146.93
Ighv15-2	0.43	0.75	37.10	6.04
Ighv1-52	4.43	5.53	44.23	69.22
Ighv1-53	22.61	91.44	237.90	427.80
Ighv1-54	5.19	7.89	56.84	95.28
Ighv1-55	37.44	78.63	377.04	545.65
Ighv1-56	2.04	1.20	10.15	22.84
Ighv1-58	8.47	17.57	81.98	208.48
Ighv1-59	17.93	30.62	175.28	241.99
Ighv1-61	12.29	21.56	104.72	180.98
Ighv1-62-1	1.44	0.75	2.81	9.84
Ighv1-62-2	15.94	42.75	119.07	256.05
Ighv1-62-3	11.44	21.18	78.95	129.51
Ighv1-63	4.34	8.80	25.77	216.75
Ighv1-64	19.44	57.09	236.12	388.71
Ighv1-66	9.16	11.68	78.56	66.89
Ighv1-67	1.58	0.61	2.53	5.02
Ighv1-69	6.53	15.77	51.29	144.32
Ighv1-7	2.61	10.41	47.60	70.56
Ighv1-71	0.32	0.92	2.95	9.65
Ighv1-72	15.22	32.10	127.73	207.24
Ighv1-74	8.60	29.26	99.76	234.55
Ighv1-75	7.38	30.89	207.76	263.25
Ighv1-76	12.66	35.28	187.63	309.80
Ighv5-9-1	9.13	34.13	73.63	169.56
Ighv6-3	16.59	142.52	120.21	256.55
Ighv6-6	8.55	109.43	79.99	172.59
Ighv7-1	5.54	28.40	68.70	100.24
Ighv7-3	11.76	72.01	83.97	144.95
Ighv7-4	0.43	3.59	2.78	11.74
Ighv8-12	7.05	30.62	88.07	115.84
Ighv9-1	1.77	6.65	38.25	51.86
Ighv9-2	5.39	6.92	60.94	75.83
Ighv9-3	23.35	60.08	309.78	264.01
Ighv9-4	10.07	16.59	68.88	89.76
Igkc	1068699.95	1263315.29	765997.23	761523.05
Igkj1	580.39	1270.37	1107.54	978.90
Igkj4	427.45	1022.67	711.59	734.77
Igkv10-94	2526.90	5617.02	4373.53	5296.12
Igkv10-95	167.30	499.46	317.46	262.88
Igkv10-96	10053.70	18128.74	18330.26	21777.55
Igkv1-110	8353.43	29829.16	11548.97	13173.79
Igkv11-125	101.03	331.95	245.19	158.86

Ighv5-12	5.77	12.55	62.67	65.21
Ighv5-15	2.30	3.13	27.21	23.42
Ighv5-16	12.77	23.98	99.92	129.47
Ighv5-17	16.23	39.00	142.66	225.96
Ighv5-2	1.50	2.66	7.85	7.65
Ighv5-4	6.19	33.22	47.76	109.33
Ighv5-6	8.72	16.51	103.01	146.24
Ighv5-9	2.31	4.86	23.40	50.67
Igkv12-98	677.36	2177.83	1068.30	1358.80
Igkv1-35	1.41	5.44	2.56	5.15
Igkv13-84	636.00	827.86	1820.34	1291.74
Igkv13-85	0.48	0.09	0.33	0.14
Igkv14-100	976.66	2984.54	1865.06	2037.67
Igkv14-111	3255.07	6163.33	9548.52	10290.04
Igkv14-126	10003.07	28975.87	3352.22	2861.16
Igkv14-130	130.39	292.23	385.65	472.53
Igkv15-103	4612.90	13342.38	4652.14	6593.78
Igkv16-104	4330.55	9163.22	5158.52	4496.23
Igkv17-121	1763.73	2227.06	3775.69	3915.09
Igkv17-127	3179.81	5943.96	6742.71	6436.71
Igkv18-36	15.65	33.05	15.69	28.59
Igkv1-88	694.87	1981.31	892.69	954.79
Igkv1-99	616.43	1723.40	1210.21	961.21
Igkv19-93	8.00	21.20	34.33	24.31
Igkv2-109	4002.55	12053.07	5918.78	5850.59
Igkv2-112	786.30	3462.97	936.66	680.47
Igkv2-137	2531.35	10382.50	3349.53	4349.87
Igkv3-1	302.11	471.07	798.97	1107.82
Igkv3-10	572.91	1336.61	2250.30	1839.83
Igkv3-12	733.76	1720.90	2365.36	3124.58
Igkv3-2	1637.28	3228.53	4890.97	4975.51
Igkv3-3	22.97	88.16	156.69	166.70
Igkv3-4	1588.60	3510.77	3510.98	3253.61
Igkv3-5	1400.66	3123.76	3476.10	3733.21
Igkv3-7	485.45	1073.17	1658.17	1526.20
Igkv3-9	53.31	45.72	126.10	107.43
Igkv4-50	740.62	1677.32	1389.75	1657.57
Igkv4-51	44.39	104.92	69.37	81.96
Igkv6-14	258.48	452.56	1054.37	930.07
Igkv6-15	7095.53	11597.76	13159.77	12306.12
Igkv6-17	2049.58	4136.83	5284.43	5521.01
Igkv6-20	754.64	1479.62	1929.86	1491.24

Igkv1-117	6485.52	25844.96	13080.94	12757.48
Igkv1-122	639.07	1781.72	567.83	559.28
Igkv1-131	201.31	558.03	236.36	409.46
Igkv1-132	47.13	260.39	108.13	255.88
Igkv1-133	531.00	1351.35	672.29	913.25
Igkv1-135	9567.37	30401.82	10170.06	18544.82
Igkv12-38	131.33	203.17	306.22	381.47
Igkv12-41	1611.40	2467.65	3308.26	2762.48
Igkv12-44	5713.66	14793.94	6445.17	8103.39
Igkv12-46	3865.19	8100.30	7805.52	7279.18
Igkv12-89	1828.96	4595.19	1548.27	1178.69
Igkv4-53	2239.77	5373.87	4338.69	3760.84
Igkv4-54	34.66	73.53	91.48	89.10
Igkv4-55	1518.37	2980.41	2411.41	2549.13
Igkv4-57	2200.04	4693.35	3820.32	3186.50
Igkv4-57-1	1126.28	2515.74	2703.71	2306.08
Igkv4-58	299.91	823.59	858.63	994.60
Igkv4-59	2471.25	7069.66	3712.58	5077.59
Igkv4-61	1076.05	1673.87	3096.42	2251.04
Igkv4-62	189.40	286.09	183.69	367.36
Igkv4-63	1486.09	3283.86	1519.82	1807.64
Igkv4-68	1531.09	3028.20	2086.41	2562.63
Igkv4-69	422.65	385.31	1645.33	347.29
Igkv4-70	583.98	1538.16	945.02	1021.05
Igkv4-71	352.86	643.31	524.46	471.54
Igkv4-72	1020.34	1494.90	3028.11	3287.32
Igkv4-74	549.42	965.15	1006.13	947.72
Igkv4-78	36.23	166.63	95.72	61.17
Igkv4-79	295.41	884.34	540.52	662.96
Igkv4-80	671.61	1604.58	1395.59	1373.42
Igkv4-81	32.94	100.59	27.13	17.74
Igkv4-86	3906.97	11924.99	2538.74	2224.75
Igkv4-90	416.49	1031.61	897.70	1627.25
Igkv4-91	1356.30	4251.66	1854.49	2141.28
Igkv4-92	31.20	96.64	75.18	130.28
Igkv5-37	60.61	124.15	185.39	174.89
Igkv5-39	29323.56	43384.78	42104.23	32318.07
Igkv5-43	3882.43	9271.75	6302.65	7584.13
Igkv5-45	1047.89	3010.38	1820.05	2028.85
Igkv5-48	4633.18	6646.39	9848.33	10156.36
Igkv6-13	840.65	1722.31	1756.55	1559.74

Igkv6-23	5669.43	6440.10	15337.17	19201.06
Igkv6-25	734.40	2543.31	2446.86	2961.34
Igkv6-29	149.11	430.39	554.35	426.00
Igkv6-32	1278.38	1880.32	3545.44	3647.37
Igkv7-33	394.04	1409.69	146.54	155.03
Igkv8-16	68.68	449.86	84.36	69.44
Igkv8-18	81.34	269.95	92.39	120.93
Igkv8-19	3415.72	8426.03	3804.56	3456.21
Igkv8-21	1039.27	1928.18	2616.72	4531.62
Igkv8-24	6811.07	16632.54	6430.19	7029.77
Igkv8-26	14.01	44.95	4.30	14.75
Igkv8-27	9998.45	24398.94	13801.88	14205.36
Igkv8-28	1786.08	6503.42	2907.34	3290.64
Igkv8-30	6987.18	18135.22	6615.20	8495.07
Igkv9-120	6948.22	13389.52	6573.22	7506.63
Igkv9-123	182.27	574.85	150.46	206.26
Igkv9-124	4275.15	9764.94	2850.73	5528.53
Igkv9-129	180.56	393.01	265.39	181.35

Table 23 Top 50 genes that are higher expressed in naïve Lag3+ compared to naïve Lag3- (left) or higher expressed in naïve Lag3- compared to naïve Lag3+ (right).

Gene symbol	logFC	PValue	FDR
Lag3	4.91	1.33E-07	0.001528
Cpq	2.35	8.83E-07	0.002004
Vav3	1.96	1.16E-06	0.002004
Bcl2	1.67	1.4E-06	0.002004
Zfp36l1	2.05	1.71E-06	0.002178
Egr1	1.95	2.04E-06	0.002343
Inpp4a	1.89	3.49E-06	0.002981
Kctd12	1.45	3.66E-06	0.002981
Prg2	1.84	4.22E-06	0.002981
Il4i1	2.70	4.59E-06	0.002981
Sema6d	1.91	4.65E-06	0.002981
Otud1	1.90	5.44E-06	0.003125
Bcl2l15	1.96	6.53E-06	0.003273
Ppfibp2	1.45	6.68E-06	0.003273
Sirpa	1.32	7.47E-06	0.003273
Sell	2.59	7.49E-06	0.003273
Jade2	1.31	8.61E-06	0.003273
Krt222	1.45	9.23E-06	0.003273
AW112010	1.97	9.69E-06	0.003273
Klf4	1.94	9.72E-06	0.003273
Rgcc	1.52	1.04E-05	0.00327
Vegfa	2.01	1.06E-05	0.00327
Bhlhe40	2.58	1.08E-05	0.00327
Ggh	2.00	1.18E-05	0.00346
Oosp2	1.62	1.44E-05	0.00395
Ly6k	1.67	1.48E-05	0.00395
Cd200	1.65	1.77E-05	0.00461
Serpina3f	1.45	2.06E-05	0.00506
Plekha1	1.24	2.13E-05	0.00506
Nfkbiz	1.12	2.22E-05	0.00506
Ackr3	1.92	2.25E-05	0.00506
Rnf130	1.47	2.42E-05	0.00535
Snn	1.14	2.87E-05	0.00615
Egr2	2.55	3.00E-05	0.00627
Phlda1	1.47	3.24E-05	0.00665
Il10	1.57	3.48E-05	0.00680
Ephx1	1.93	3.76E-05	0.00719
AB124611	1.49	5.07E-05	0.00840
Btla	1.34	5.60E-05	0.00894
Irf2bp2	1.15	5.77E-05	0.00907
Clec2i	1.80	5.85E-05	0.00908
Pglyrp1	1.31	6.71E-05	0.00942
Abhd6	1.49	6.95E-05	0.00962
Srsf7	0.89	7.30E-05	0.00985
Fgf1	3.38	7.36E-05	0.00985
Ctss	1.32	7.70E-05	0.01002
Twsg1	1.26	7.88E-05	0.01002
Zswim6	1.46	7.94E-05	0.01002
Gpr18	1.49	8.27E-05	0.01010
Rtp4	1.62	8.39E-05	0.01010

Gene symbol	logFC	PValue	FDR
Ighe	-4.86	5.86E-07	0.002004
Ighg3	-2.79	7.65E-07	0.002004
Cdk1	-1.76	9.26E-07	0.002004
Mcm5	-1.81	1.25E-06	0.002004
Cdc45	-2.12	3.06E-06	0.002981
Kpna2	-1.41	0.000004	0.002981
Ada	-1.14	4.83E-06	0.002981
Ighg2c	-3.01	4.93E-06	0.002981
Eno1	-1.34	5.79E-06	0.003164
Ncapg	-1.84	8.08E-06	0.003273
Rrm2	-1.63	8.94E-06	0.003273
Kif11	-1.64	9.79E-06	0.003273
Slpi	-2.10	9.81E-06	0.003273
Ighg2b	-2.66	1.03E-05	0.003273
Ccnb2	-1.69	1.04E-05	0.003273
Nuf2	-2.30	1.26E-05	0.003614
Cep55	-1.82	0.000014	0.003929
Spc24	-2.04	1.88E-05	0.004793
Rfc4	-1.21	0.000021	0.005062
Nmral1	-1.27	2.18E-05	0.005062
Ssr2	-1.57	2.89E-05	0.00615
Prelid1	-1.00	3.35E-05	0.00674
Cdkn3	-1.82	3.49E-05	0.00680
Cdc20	-1.83	4.37E-05	0.00822
Tg	-1.33	4.74E-05	0.00840
Selm	-1.17	4.78E-05	0.00840
Ighv1-63	-2.70	4.86E-05	0.00840
Ddost	-1.15	4.97E-05	0.00840
Ahcy	-1.19	5.06E-05	0.00840
Igkv4-69	-1.91	5.10E-05	0.00840
Ighv1-75	-4.83	5.12E-05	0.00840
NA	-3.48	5.12E-05	0.00840
Top2a	-1.38	5.61E-05	0.00894
Ighv15-2	-6.47	6.09E-05	0.00924
Tigit	-1.83	6.12E-05	0.00924
H2-Eb1	-0.87	6.34E-05	0.00936
Spc25	-1.26	6.39E-05	0.00936
Kif22	-1.89	6.44E-05	0.00936
Tcf19	-1.50	6.58E-05	0.00942
Mad2l1	-1.14	6.72E-05	0.00942
Cxcr3	-2.33	7.37E-05	0.00985
Igkv3-10	-1.93	7.74E-05	0.01002
Pbk	-2.26	7.84E-05	0.01002
Ighv1-47	-5.14	8.04E-05	0.01004
Dnpep	-1.03	8.16E-05	0.01008
Srm	-0.92	8.80E-05	0.01021
Gstp2	-1.30	9.19E-05	0.01034
Tmem97	-1.47	9.61E-05	0.01058
Shcbp1	-1.71	9.65E-05	0.01058
Cenpe	-1.81	9.77E-05	0.01058

Table 24 Top 50 genes that are higher expressed in Lag3⁺ day1 compared to naïve Lag3⁻ day1 (left) or higher expressed in Lag3⁻ day1 compared to Lag3⁺ day1 (right).

Gene symbol	logFC	PValue	FDR
Lag3	5.93	3.40E-09	2.84E-05
Vav3	2.96	7.17E-09	2.84E-05
Cpq	3.10	1.98E-08	2.84E-05
Sema6d	2.95	2.88E-08	3.20E-05
Rab30	1.99	7.30E-08	5.65E-05
Btla	2.48	8.11E-08	5.65E-05
Bcl2	1.97	9.73E-08	6.21E-05
Prg2	2.32	1.54E-07	7.37E-05
Snn	1.82	1.77E-07	7.46E-05
Oosp2	2.28	2.06E-07	7.46E-05
P2rx4	1.72	2.27E-07	7.68E-05
Cd200	2.29	3.33E-07	9.80E-05
Slamf9	2.10	3.51E-07	9.80E-05
Ephx1	2.88	3.84E-07	9.80E-05
Rgcc	1.92	4.09E-07	9.80E-05
Otud1	2.20	4.24E-07	9.93E-05
Pglyrp1	2.09	4.65E-07	0.000104
Egr1	1.99	4.82E-07	0.000104
Gpr18	2.32	5.91E-07	0.000118
Dsel	2.84	6.05E-07	0.000118
Sesn1	2.25	6.69E-07	0.000124
2610035D17Rik	2.17	6.79E-07	0.000124
Hsd17b11	1.96	8.53E-07	0.000138
Ypel3	1.39	9.13E-07	0.000144
Il13ra1	1.78	9.47E-07	0.000145
Ggh	2.30	1.02E-06	0.000154
Vegfa	2.22	1.08E-06	0.000159
Fam214a	1.44	1.13E-06	0.000165
Ly6k	1.94	1.19E-06	0.000166
Iglc2	2.54	1.21E-06	0.000166
Egr2	3.22	1.25E-06	0.000168
Fez2	1.65	1.31E-06	0.000172
Thbd	2.23	1.41E-06	0.000180
Kcnmb4os2	1.56	1.51E-06	0.000188
Hdac9	1.68	1.58E-06	0.000193
Ptpn6	2.22	1.60E-06	0.000193
Ipcef1	1.61	1.63E-06	0.000195
Serpina3f	1.68	1.68E-06	0.000198
Klf4	2.06	1.74E-06	0.000202
Sepp1	2.25	1.82E-06	0.000205
Ppt1	1.50	1.91E-06	0.000210
Itgav	1.52	1.92E-06	0.000210
Rasgef1a	2.38	2.07E-06	0.000221
Epb41l2	1.49	2.17E-06	0.000223
Cnst	1.52	2.20E-06	0.000223
Tspan32	1.46	2.26E-06	0.000227
Sell	2.51	2.52E-06	0.000248
Clec2i	2.29	2.55E-06	0.000248
Samsn1	1.72	2.61E-06	0.000249
Ctss	1.71	2.86E-06	0.000265

Gene symbol	logFC	PValue	FDR
Cdk1	-2.83	1.10E-08	2.84E-05
Rrm2	-3.38	1.47E-08	2.84E-05
Kif11	-3.56	1.59E-08	2.84E-05
Mcm5	-2.92	1.63E-08	2.84E-05
Cep55	-4.38	1.78E-08	2.84E-05
Ighg3	-4.07	2.81E-08	3.20E-05
Top2a	-3.28	3.32E-08	3.20E-05
Cdca5	-4.17	3.34E-08	3.20E-05
Cdc45	-3.28	6.85E-08	5.65E-05
Spc24	-4.21	8.30E-08	5.65E-05
Ccnb2	-3.01	8.36E-08	5.65E-05
Ttk	-4.21	1.08E-07	6.39E-05
Ccr9	-2.45	1.11E-07	6.39E-05
Prc1	-2.85	1.26E-07	6.91E-05
Cdc20	-3.82	1.45E-07	7.30E-05
Ighg2b	-4.46	1.46E-07	7.30E-05
Brca1	-3.32	1.87E-07	7.46E-05
Ptgr1	-3.37	1.93E-07	7.46E-05
Hmmr	-4.38	2.01E-07	7.46E-05
Ighg2c	-4.41	2.03E-07	7.46E-05
Foxm1	-6.02	2.09E-07	7.46E-05
Kpna2	-1.88	2.13E-07	7.46E-05
Ube2c	-3.72	2.14E-07	7.46E-05
Ighv1-63	-4.59	3.04E-07	9.80E-05
Rad51ap1	-3.41	3.28E-07	9.80E-05
Eno1	-1.76	3.34E-07	9.80E-05
Bub1	-3.71	3.43E-07	9.80E-05
Nuf2	-3.61	3.81E-07	9.80E-05
Mcm10	-4.08	3.87E-07	9.80E-05
Ighe	-4.92	4.03E-07	9.80E-05
Tacc3	-2.44	4.05E-07	9.80E-05
Clspn	-4.24	4.10E-07	9.80E-05
H2-Eb1	-1.47	4.10E-07	9.80E-05
Banf2os	-3.38	4.71E-07	0.000104
Cdca8	-3.69	4.72E-07	0.000104
Birc5	-3.83	5.00E-07	0.000106
Ncapg	-2.45	5.44E-07	0.000114
2810417H13Rik	-3.57	5.56E-07	0.000114
Shcbp1	-3.24	6.14E-07	0.000118
Uhrf1	-4.37	6.19E-07	0.000118
Ccna2	-3.10	6.67E-07	0.000124
Cenpe	-3.44	7.04E-07	0.000126
Neil3	-3.78	7.51E-07	0.000133
Pik3r5	-2.47	7.64E-07	0.000133
Gmnn	-1.71	7.97E-07	0.000137
Cdkn3	-2.92	8.38E-07	0.000138
Cdca3	-3.39	8.40E-07	0.000138
Spc25	-2.02	8.52E-07	0.000138
Tyms	-1.94	8.87E-07	0.000141
Cdc6	-2.76	9.48E-07	0.000145

Table 25 Top 50 genes that are higher expressed in naïve LAG-3⁺ compared to LAG-3⁺ day1 (right) or higher expressed in LAG-3⁺ day1 compared to naïve LAG-3⁺ (left).

Symbol	logFC	PValue	FDR
Tg	-2.624	1.4E-07	5.6E-04
Oasl2	-3.418	1.4E-07	5.6E-04
Slamf9	-2.716	1.5E-07	5.6E-04
Lgals3bp	-2.091	2.8E-07	6.2E-04
Tgtp2	-3.175	3.2E-07	6.2E-04
Ly6c2	-2.237	5.1E-07	8.1E-04
Rtp4	-2.669	5.9E-07	8.1E-04
Pik3cg	-1.601	6.4E-07	8.1E-04
Ifitm3	-2.014	8.2E-07	9.4E-04
Isg15	-2.211	9.0E-07	9.4E-04
Ly6c1	-3.026	1.3E-06	1.2E-03
Gbp10	-1.968	1.3E-06	1.2E-03
Zbp1	-1.624	1.5E-06	1.2E-03
Pydc4	-1.866	2.6E-06	2.0E-03
Tgtp1	-3.331	4.1E-06	2.6E-03
Xaf1	-1.715	7.6E-06	3.5E-03
Ada	-1.073	8.9E-06	3.9E-03
Usp18	-1.825	1.0E-05	4.2E-03
Gm43302	-1.762	1.1E-05	4.2E-03
Ifi47	-1.601	1.1E-05	4.2E-03
Stat1	-1.371	1.2E-05	4.2E-03
Gm5431	-2.353	1.3E-05	4.4E-03
Phf11b	-1.683	1.4E-05	4.4E-03
AW011738	-1.400	1.5E-05	4.4E-03
Npc2	-1.386	1.5E-05	4.4E-03
Herc6	-1.822	1.5E-05	4.4E-03
Serpina3f	-1.414	2.4E-05	6.1E-03
Lamp2	-1.260	2.6E-05	6.4E-03
Vmp1	-1.176	3.0E-05	6.9E-03
Phf11d	-2.006	3.1E-05	6.9E-03
Psmd14	-1.191	3.5E-05	7.4E-03
Ggta1	-0.985	3.7E-05	7.7E-03
Rsph1	-1.502	4.1E-05	8.3E-03
Ube2l6	-1.711	4.2E-05	8.3E-03
Ifi204	-1.986	5.3E-05	9.9E-03
Gm11827	-1.632	5.9E-05	1.1E-02
Fcgr2b	-0.980	6.1E-05	1.1E-02
Gm4955	-1.607	6.1E-05	1.1E-02
Ifi27l2a	-2.071	6.1E-05	1.1E-02
Lgals9	-1.147	6.3E-05	1.1E-02
Atp6v0a2	-1.046	6.4E-05	1.1E-02
Ifi27	-1.137	6.9E-05	1.1E-02
H2-T22	-0.920	6.9E-05	1.1E-02
Ly6a	-1.288	7.8E-05	1.2E-02
Cldn7	-2.970	7.9E-05	1.2E-02
Nt5c3	-0.952	8.0E-05	1.2E-02
Lag3	-1.923	8.4E-05	1.2E-02
Irf7	-1.593	8.9E-05	1.2E-02
Cox7a2l	-1.146	9.7E-05	1.3E-02
A430005L14Rik	-1.079	9.9E-05	1.3E-02

Symbol	logFC	PValue	FDR
Ccr9	2.432	1.96E-07	0.00056
Trim7	2.270	3.22E-06	0.00231
Ccr10	2.660	3.68E-06	0.00249
AI504432	1.645	5.34E-06	0.00323
NA	1.498	6.00E-06	0.00330
Foxm1	4.491	6.03E-06	0.00330
Itgb7	2.386	6.51E-06	0.00340
Cbfa2t3	2.332	7.29E-06	0.00355
Fhit	3.201	7.72E-06	0.00355
Pycard	1.263	9.56E-06	0.00407
Rapgef1	2.235	1.18E-05	0.00424
Pik3r5	1.884	1.38E-05	0.00443
Osbpl3	2.451	1.42E-05	0.00443
H2-Eb1	1.044	1.48E-05	0.00443
Rasgrp2	1.338	2.29E-05	0.00607
Ccnd2	1.300	2.33E-05	0.00607
Sspn	2.071	2.35E-05	0.00607
Ltb	2.373	2.38E-05	0.00607
Chka	2.208	2.62E-05	0.00640
Cbll1	1.297	2.89E-05	0.00692
Runx2	2.149	3.07E-05	0.00694
Kif11	1.596	3.25E-05	0.00712
Ctse	1.048	3.29E-05	0.00712
Il2rg	1.038	3.74E-05	0.00767
Usp2	2.966	4.71E-05	0.00917
Cep55	1.945	4.95E-05	0.00948
Cdca5	1.982	5.17E-05	0.00974
Bhlhe40	2.099	6.38E-05	0.01062
Smim3	1.985	7.14E-05	0.01131
Nfkbid	1.765	7.19E-05	0.01131
Ryr1	1.320	7.65E-05	0.01183
Cntnap1	2.631	8.03E-05	0.01183
Ccdc17	1.134	8.54E-05	0.01200
Ptgr1	1.779	8.54E-05	0.01200
Sgpp2	3.052	8.57E-05	0.01200
Slc22a15	4.466	9.20E-05	0.01253
Sema4d	1.183	9.42E-05	0.01253
Cxcr3	2.838	9.44E-05	0.01253
Nebl	2.330	9.54E-05	0.01253
Ighg2b	1.988	9.84E-05	0.01253
Lpp	2.111	0.000109	0.01287
H2-Ab1	1.188	0.000112	0.01299
Gm10720	6.993	0.000123	0.01350
Tmem138	2.019	0.000135	0.01424
Ttk	1.983	0.00016	0.01584
Spc24	1.860	0.000168	0.01624
Top2a	1.255	0.000179	0.01667
Neil3	2.119	0.000182	0.01671
Nek6	1.437	0.000185	0.01671
Prc1	1.297	0.000185	0.01671

Table 26 Top 50 genes that are higher expressed in naïve Lag3- compared to Lag3- day1 (right) or higher expressed in Lag3- day1 compared to naïve Lag3- (left).

Symbol	logFC	PValue	FDR
Rtp4	-3.795	2.6E-08	1.5E-04
Gbp10	-2.718	4.9E-08	1.5E-04
Oasl2	-3.516	5.6E-08	1.5E-04
Lgals3bp	-2.300	7.2E-08	1.5E-04
Isg15	-2.770	7.9E-08	1.5E-04
Usp18	-2.992	8.4E-08	1.5E-04
Ifitm3	-2.454	9.3E-08	1.5E-04
Gm43302	-2.640	2.1E-07	3.0E-04
Tgtp2	-2.823	4.7E-07	6.0E-04
Ifit1	-3.670	5.9E-07	6.8E-04
Pydc4	-2.010	7.6E-07	8.0E-04
Xaf1	-2.049	9.3E-07	8.9E-04
Ly6c2	-1.894	1.3E-06	1.2E-03
Zbp1	-1.382	3.3E-06	2.3E-03
H2-T22	-1.220	3.4E-06	2.3E-03
Ighv1-63	-3.018	4.3E-06	2.7E-03
Bst2	-1.740	4.6E-06	2.8E-03
Ube2l6	-2.046	5.0E-06	2.8E-03
Rnf213	-1.320	5.2E-06	2.8E-03
Ifi47	-1.603	5.8E-06	3.0E-03
Slamf9	-1.660	6.3E-06	3.1E-03
Tgtp1	-2.818	8.1E-06	3.9E-03
Herc6	-1.797	9.2E-06	4.2E-03
Ada	-0.969	1.0E-05	4.4E-03
Rsph1	-1.603	1.1E-05	4.5E-03
Ly6a	-1.527	1.2E-05	4.6E-03
Stat1	-1.249	1.4E-05	5.3E-03
Igtp	-1.258	1.4E-05	5.3E-03
AW011738	-1.258	1.8E-05	6.1E-03
Ly6c1	-2.058	1.9E-05	6.1E-03
Tg	-1.318	2.3E-05	6.8E-03
Ifi27l2a	-2.219	2.4E-05	6.8E-03
Gm4955	-1.621	2.9E-05	7.8E-03
Lgals9	-1.155	3.0E-05	7.8E-03
H2-Q7	-1.168	4.4E-05	1.1E-02
Ifi204	-1.816	4.4E-05	1.1E-02
Il22	-5.505	4.5E-05	1.1E-02
Ifi27	-1.079	6.0E-05	1.3E-02
Serpina3f	-1.179	6.1E-05	1.4E-02
Parp9	-0.960	7.4E-05	1.5E-02
Rsad2	-1.682	7.7E-05	1.6E-02
Ifi35	-0.993	8.0E-05	1.6E-02
Irgm2	-1.146	8.5E-05	1.7E-02
Cldn7	-2.727	8.8E-05	1.7E-02
Mx1	-1.981	8.8E-05	1.7E-02
Tlr7	-1.504	9.4E-05	1.7E-02
H2-T23	-1.078	9.8E-05	1.8E-02
NA	-1.232	1.0E-04	1.9E-02
Ccl5	-2.812	1.5E-04	2.5E-02
Ighv2-6	-2.741	1.7E-04	2.7E-02

Symbol	logFC	PValue	FDR
NA	1.433	2.7E-06	2.2E-03
Igkv4-69	2.289	3.4E-06	2.3E-03
Atp1b1	1.424	1.1E-05	4.5E-03
Ctse	1.003	1.6E-05	5.7E-03
Ltb	2.021	1.9E-05	6.1E-03
Cd28	1.034	1.9E-05	6.2E-03
Ccdc17	1.168	2.3E-05	6.8E-03
Pycard	1.001	2.4E-05	6.8E-03
Trim7	1.523	2.4E-05	6.8E-03
Sspn	1.669	2.9E-05	7.8E-03
Reln	1.110	3.0E-05	7.8E-03
Lpp	2.090	3.8E-05	9.4E-03
Il2rg	0.896	4.7E-05	1.1E-02
Snn	0.934	6.5E-05	1.4E-02
Gm15513	1.052	7.1E-05	1.5E-02
Fhit	1.913	8.0E-05	1.6E-02
Ctso	0.783	1.1E-04	2.0E-02
Anxa2	2.267	1.3E-04	2.2E-02
Itgam	1.680	1.3E-04	2.2E-02
Il18	1.300	1.4E-04	2.3E-02
Rapgef1	1.406	1.6E-04	2.6E-02
Tspan32	0.875	1.8E-04	2.8E-02
Evi2a	0.838	2.2E-04	3.1E-02
Rasgef1a	1.311	2.9E-04	3.9E-02
Dnm3	1.193	3.6E-04	4.5E-02
St6gal1	0.713	3.7E-04	4.5E-02
Tmem173	0.820	4.2E-04	5.0E-02

Table 27 Genes with unique downregulation in Lag3⁺ day1 cells

Symbol	Lag3- day1 R1	Lag3- day1 R2	Lag3- day1 R3	Lag3- R2	Lag3- R1	Lag3+ day1 R1	Lag3+ day1 R2	Lag3+ R2	Lag3+ R1
2610524H06Rik	32.12	34.46	36.43	26.47	34.54	7.83	11.47	17.14	18.45
2810468N07Rik	2.15	1.59	2.11	2.02	2.41	1.07	0.75	2.83	2.39
Akap1	1.34	1.22	0.89	0.94	0.96	0.18	0.38	0.80	1.02
Anln	0.86	0.75	0.88	0.53	0.62	0.08	0.02	0.21	0.29
Apaf1	1.21	1.20	1.24	1.32	1.20	0.60	0.42	0.88	0.93
Arhgap11a	1.92	1.01	1.88	1.64	1.48	0.16	0.29	0.63	0.95
Atp2a3	12.39	6.29	11.85	10.93	10.67	3.00	2.08	6.90	5.59
Brca1	1.19	1.26	1.05	0.74	0.61	0.12	0.11	0.22	0.31
Btk	7.32	6.29	6.45	6.77	7.42	3.15	3.56	5.36	5.11
Car2	0.55	0.69	0.79	0.74	0.55	0.09	0.05	0.55	0.76
Casc5	0.43	0.37	0.44	0.26	0.34	0.05	0.08	0.16	0.30
Cbfa2t3	7.10	3.93	6.23	8.33	6.48	1.16	1.43	6.48	6.57
Cbll1	2.61	2.24	2.54	2.57	2.70	1.28	1.21	2.75	3.37
Ccnb2	11.95	11.04	13.69	12.60	9.35	1.65	1.40	3.03	3.74
Ccnd2	14.84	11.05	15.03	14.24	15.23	5.43	4.83	12.84	12.41
Ccr9	2.66	2.74	2.59	3.78	3.02	0.53	0.45	2.80	2.46
Cdca8	2.59	2.50	2.44	1.69	1.61	0.18	0.20	0.33	0.72
Celf1	4.27	3.59	4.36	4.17	3.82	2.45	2.61	4.02	4.06
Chst12	5.65	4.25	6.47	5.65	6.24	2.93	2.79	5.59	4.94
Cntnap1	0.40	0.42	0.43	0.57	0.46	0.14	0.06	0.57	0.63
Cpeb2	11.06	8.86	9.50	11.95	10.30	5.36	4.29	8.09	9.07
D3Ertd254e	0.31	0.39	0.40	0.35	0.48	0.16	0.19	0.57	0.46
Dnmt1	2.28	1.91	2.47	2.15	1.85	0.84	0.48	1.23	1.76
Elac2	1.82	1.77	1.73	2.04	1.95	0.91	0.83	1.74	1.57
Eme2	2.10	1.27	2.01	2.11	2.33	0.97	0.63	1.91	1.75
Erdr1	21.21	16.67	18.68	25.11	15.76	9.75	8.84	17.42	17.88
Ermp1	0.95	1.05	0.86	0.92	1.00	0.41	0.58	0.97	0.90
Fas	6.29	5.82	7.31	7.91	7.68	3.20	2.32	5.15	5.40
Fmn2	0.15	0.25	0.24	0.19	0.19	0.02	0.03	0.17	0.21
Gbf1	2.74	1.33	2.56	2.80	2.51	0.87	0.75	2.01	2.78
Gga3	2.42	1.65	2.34	2.25	2.43	1.61	1.33	2.71	2.56
Gm22579	26.72	14.89	19.40	24.71	23.53	13.73	8.58	25.95	43.90
Gmnn	16.67	19.14	20.73	13.88	13.54	5.86	5.68	11.01	9.26
H2-Eb2	4.88	5.33	5.53	5.03	5.78	2.51	2.60	4.43	4.41
Hn1	37.21	36.34	41.58	47.10	41.75	18.16	19.77	28.01	32.79
Hnrnpul2	4.33	3.39	4.50	4.67	4.59	2.40	2.84	5.42	4.49
Ighg2b	27362.19	24059.34	23646.32	24957.14	32071.51	1329.10	951.04	5481.98	3561.50
Ighg2c	25677.73	22195.67	18208.49	17542.64	27114.34	1173.53	901.59	3119.29	2435.24
Iqgap2	0.35	0.20	0.48	0.48	0.38	0.03	0.01	0.23	0.33
Irf5	4.83	2.84	4.92	3.86	5.26	2.27	2.10	4.61	4.75
Kmt2b	1.52	0.93	1.36	1.28	1.42	0.55	0.65	1.20	1.19
Larp1	4.74	3.93	4.72	5.30	4.81	2.99	3.16	4.53	5.06
Larp4b	4.11	3.46	4.81	3.80	4.38	2.66	2.30	5.10	3.89
Lat	2.97	3.29	2.83	2.61	2.06	1.03	1.30	2.26	5.80
Mapk1ip1l	2.62	2.23	2.68	2.28	2.22	1.60	1.55	2.77	3.22
Med24	1.70	1.31	1.46	1.43	1.82	0.68	0.78	1.35	1.40
Mum1	0.76	0.53	0.78	0.79	0.72	0.40	0.35	0.89	0.83
Myc	6.02	5.12	6.31	5.58	4.94	4.01	4.53	10.19	6.78
Nln	1.48	2.12	2.05	2.58	2.02	0.86	0.92	1.83	2.22
Pabpc4	5.99	6.25	6.29	5.64	5.22	3.38	4.67	6.84	7.76
Palm3	0.73	0.26	0.52	0.57	0.53	0.02	0.01	0.34	0.12
Pfn2	0.47	0.40	0.87	0.59	0.75	0.07	0.22	0.94	0.75
Prc1	2.12	2.19	2.75	2.11	1.74	0.33	0.32	0.68	0.91
Prrc2b	1.50	0.82	1.45	1.61	1.57	0.79	0.65	1.64	1.54
Pycr2	32.90	25.92	29.87	28.23	29.80	12.71	12.88	19.73	25.42
Racgap1	2.75	2.29	2.65	2.71	1.89	0.53	0.41	0.96	1.34

Symbol	Lag3- day1 R1	Lag3- day1 R2	Lag3- day1 R3	Lag3- R2	Lag3- R1	Lag3+ day1 R1	Lag3+ day1 R2	Lag3+ R2	Lag3+ R1
Rad51ap1	1.97	1.63	2.19	1.22	1.02	0.18	0.18	0.38	0.48
Rhbdf2	0.66	0.24	0.46	0.50	0.55	0.09	0.14	0.40	0.70
Rrm2	30.94	31.72	33.68	20.40	19.78	3.60	2.59	5.40	7.53
Runx2	1.74	1.02	1.79	2.07	1.51	0.33	0.30	1.27	1.48
Safb	2.76	2.17	2.47	2.51	2.40	1.17	1.40	2.25	2.73
Sdad1	2.66	2.35	2.20	2.20	1.99	1.41	1.09	2.67	2.21
Sema4d	3.22	2.52	2.98	3.06	2.93	0.99	1.11	2.08	2.67
Serinc5	8.88	9.20	9.63	12.46	11.87	5.61	6.86	10.75	9.44
Slc39a14	0.89	0.86	0.93	0.87	0.99	0.29	0.18	0.48	0.56
Slc43a3	5.37	4.39	5.46	3.53	3.87	1.53	1.34	2.23	3.24
Slc4a7	1.83	1.48	2.09	1.76	1.66	0.88	0.80	1.42	1.60
Slx4	1.15	0.67	0.88	1.13	1.12	0.18	0.29	0.70	0.60
Smim3	1.94	2.42	1.85	2.23	1.91	0.45	0.60	1.86	2.29
Top2a	5.47	5.77	5.70	3.88	3.34	0.50	0.66	1.11	1.66
U2af2	3.35	3.11	3.61	3.02	4.42	1.91	2.57	4.22	5.33
Vps52	3.95	2.45	3.57	3.88	3.68	1.59	1.31	3.09	3.61
Zfp599	0.30	0.37	0.15	0.35	0.26	0.00	0.04	0.35	0.26

Table 28 Genes with unique upregulation in Lag3⁺ day1 cells

Symbol	Lag3- day1 R1	Lag3- day1 R2	Lag3- day1 R3	Lag3- R2	Lag3- R1	Lag3+ day1 R1	Lag3+ day1 R2	Lag3+ R2	Lag3+ R1
Igkv1-110	8124.02	11937.76	7689.99	8807.46	7965.95	23928.29	18914.91	9219.71	3059.28
Igkv1-117	7509.98	13270.31	6123.28	8170.94	10898.86	21332.98	15823.05	6910.30	2647.44
Igkv2-137	3264.76	3579.54	2321.49	2464.30	2404.04	8005.36	6887.89	2683.81	1047.97
Igkv8-28	2021.45	3066.93	1845.26	2007.89	2222.94	5296.19	4049.25	1888.69	744.89
Igkv2-112	432.96	621.47	379.35	718.36	641.85	2625.75	2339.13	798.48	364.18
Igkv1-122	358.17	484.78	335.01	477.42	345.53	1542.97	1022.77	650.52	294.28
Xbp1	489.68	595.28	429.17	394.16	458.01	649.91	711.02	395.83	356.41
Igkv8-16	27.77	72.05	46.41	74.71	47.41	338.39	306.41	65.83	36.11
Slamf7	153.07	179.19	145.58	136.15	150.62	235.48	201.44	124.30	86.94
Igkv4-78	29.38	53.71	45.60	74.39	64.59	163.42	77.66	34.32	19.50
Rabac1	77.28	119.35	62.80	59.85	94.43	156.26	198.51	78.56	74.93
Myl12a	59.71	63.42	60.22	51.04	45.19	113.04	124.36	71.84	72.46
Gabarap	64.48	85.59	60.71	49.71	67.98	107.47	135.53	63.10	58.37
Cox7a2l	35.59	48.21	34.18	39.29	37.45	100.09	100.43	48.58	42.05
Cr1l	60.66	68.19	57.99	64.47	60.85	87.18	86.58	62.61	47.99
Rpl18	51.29	59.17	43.10	43.09	52.33	63.16	74.12	29.88	43.92
Ypel3	20.86	22.88	18.83	28.43	27.86	56.19	53.53	29.20	33.17
Rab30	13.14	14.20	12.89	21.59	16.27	53.86	52.86	29.38	31.29
Tnfrsf13c	27.24	30.05	30.43	23.42	26.99	48.14	51.16	31.74	30.63
Ppt1	14.15	17.00	14.64	18.42	17.86	46.40	39.81	29.48	22.76
Rgs10	22.99	24.23	24.70	20.75	22.14	44.97	50.62	29.90	34.12
Gpx4	16.67	21.96	11.12	19.39	21.06	44.44	40.41	19.58	15.91
Btla	6.50	8.19	8.47	10.04	8.62	42.35	43.78	27.92	19.52
Ddrgk1	27.46	28.20	24.48	23.94	24.30	40.03	36.17	22.33	22.06
Eif4ebp1	17.59	23.27	16.85	16.04	16.79	38.65	45.76	21.63	22.58
Use1	18.50	18.30	17.98	20.01	18.33	36.89	39.20	27.20	25.44
Atraid	23.05	23.05	21.80	22.88	22.17	35.09	37.46	23.90	23.98
Pcbp2	19.18	21.14	15.92	16.50	16.12	31.61	27.71	18.44	13.59
Mif4gd	9.45	11.43	9.20	8.56	8.81	29.71	23.13	17.42	13.93
Ppp1r11	18.51	19.68	17.75	14.70	17.36	20.99	25.00	14.21	14.06
Katna1	7.32	7.78	7.58	6.63	5.75	19.87	15.95	9.50	8.59
Surf1	11.65	12.73	10.34	11.18	10.90	19.10	17.02	10.79	9.56
Rab9	9.37	10.44	9.63	9.75	8.46	18.96	18.44	12.00	9.61
Eif4b	12.21	13.75	11.39	12.15	11.22	18.66	16.08	10.83	10.02
Epb41l2	5.48	6.37	6.07	5.38	4.53	18.35	15.33	11.25	9.06
4921524J17Rik	8.61	10.19	8.47	7.01	7.22	16.94	18.71	12.34	9.65
Far1	9.45	9.64	10.79	7.45	7.83	16.71	17.28	11.17	9.84
Ctsl	5.34	6.16	4.78	5.13	5.47	14.55	13.73	8.52	7.60
Il13ra1	4.35	5.12	4.05	3.98	4.06	14.12	16.92	10.59	7.88
Leprotl1	9.25	10.08	8.40	7.28	8.15	13.36	13.77	8.70	8.44
Gm11827	3.40	3.09	3.53	2.20	3.53	10.64	9.55	3.48	3.05
Mfap1b	5.48	5.39	5.50	5.38	5.38	10.60	8.70	6.51	5.78
Sypl	3.78	3.98	3.74	3.37	2.95	10.25	9.06	5.56	3.96
Vrk1	5.98	6.98	5.74	5.44	4.82	9.48	8.75	5.34	4.58
Osbpl9	5.10	6.40	5.70	6.99	5.94	9.34	9.90	6.50	5.71
Snap23	4.51	5.12	5.05	4.49	4.41	8.61	7.38	5.47	4.35
Srfbp1	4.18	4.20	3.18	3.28	2.99	8.21	8.09	4.84	4.35
Cfl2	5.08	5.47	5.13	4.65	3.84	8.04	7.23	5.14	3.97
Mtpap	5.14	6.06	5.97	5.90	4.63	7.68	6.86	4.49	4.20
Enpp1	6.24	6.90	5.25	5.06	5.35	7.17	7.88	3.91	4.34
Bbs9	2.04	2.73	1.92	2.42	2.07	6.30	6.30	3.59	3.26
Nck1	3.39	3.82	4.32	3.65	3.14	6.12	5.62	3.71	3.57
Thbd	1.27	1.27	1.57	1.09	1.59	5.23	7.58	2.51	2.12
Ap5m1	2.77	2.89	3.08	2.89	2.50	4.74	3.73	2.39	2.49