2021

Developmental Dynamics of 5-Hydroxymethylcytosine and its Role in the Terminal Differentiation of Neurons

Elitsa Stoyanova

DEVELOPMENTAL DYNAMICS OF 5-HYDROXYMETHYLCYTOSINE AND ITS ROLE IN THE TERMINAL DIFFERENTIATION OF NEURONS

A Thesis Presented to the Faculty of
The Rockefeller University
in Partial Fulfillment of the Requirements for
the degree of Doctor of Philosophy

by
Elitsa Stoyanova
June 2021

DEVELOPMENTAL DYNAMICS OF 5-HYDROXYMETHYLCYTOSINE AND ITS ROLE IN THE TERMINAL DIFFERENTIATION OF NEURONS

Elitsa Stoyanova, Ph.D.

The Rockefeller University 2021

While epigenetic dynamics in mitotic differentiating cells have been characterized in depth, little is known about how postmitotic neurons regulate their chromatin state. Modulation of transcription through the regulation of accessibility of transcription factor binding sites is essential for the regulation of migration, synapse formation and terminal differentiation. Although 5-hydroxymethylcytosine (5hmC), the oxidized form of 5-methylcytosine, accumulates to high levels in neuronal lineages and refines the genomic binding of MeCP2, the functional consequences of 5hmC deposition in differentiating neurons have not been determined. We report high resolution characterization of the genomic landscape of developing postmitotic Purkinje neurons, confirm the relationship between 5hmC and gene expression and identify a novel class of genes that are demethylated in the absence of cell division. Deletion of the 5hmC writers Tet1, Tet2, and Tet3 from postmitotic Purkinje cells alters cytosine modification in regulatory domains, impairs gene expression, hinders developmental transitions, and causes hyper-excitability and increased susceptibility to excitotoxic drugs. These data demonstrate that 5hmC and the Tet proteins are required for terminal differentiation of Purkinje cells through demethylation of important regulatory regions. Our data, along with recent reports of the role of Tet proteins in neurodegeneration suggest an essential role in development and function of all neurons.

For my family and my friends,
without whom none of this would be possible

ACKNOWLEDGMENTS

I had absolutely no idea what I was in for when I started graduate school. I had no plans to be a neuroscientist – I hadn't even taken a single course in neuroscience in college! But once I rotated in the Heintz lab and met Nat, I knew that is the place I would spend the next 6 years at.

I would like to start by thanking Nat for having an infectious excitement for science. It is inspiring and motivating, sometimes even crazy in the best way possible. Looking back at what I did, I can't believe you convinced me it was a smart idea to generate a quadruple transgenic mouse in my third year. Thank you for not only shaping me as a scientist but also being a continuous support in my future endeavors.

I feel incredibly lucky to have been mentored by two of the most skilled postdocs in the beginning of my graduate work – Xiao Xu and Francesco Piccolo. You not only thought me how to be a good experimentalist, but also how to think about my project – big and small. I sincerely believe that without your training and help my PhD would have been much less enjoyable and successful.

I would like to thank all the past and present members of the Heintz lab. They made the last 6 years incredibly fun and rewarding, and the challenging moments - easier. In particular, I'd like to thank Michael Riad who has been an excellent collaborator and a friend, absolutely crucial to wrapping up the story of 5hmC and the TET proteins in Purkinje cells.

I've received a great deal of advice and encouragement from my committee members – Dr. Titia de Lange and Dr. Sasha Tarakhovsky. Thank you for thoughtful and critical contributions to my project. I would also like to thank my external examiner – Dr. Anne Schaefer – for agreeing to be part of final examination.

I've been incredibly lucky to be at Rockefeller. I would like to thank every member of the Dean's office – Sid, Emily, Martha, Kristen, Andrea, Cris and Stephanie – for their continued support during my studies and their unbelievable flexibility when issues arise. I always felt that you are on my side and that feeling is invaluable.

I would like to my undergraduate research mentors Kathryn Cole, Ian Woods and Matt Deardorff. You took a chance on me and sparked a fire that has been burning incessantly.

I received a tremendous amount of support and mentoring from the scientists at the Rockefeller cores. I am indebted to Thomas Carroll for his patience and endless sources of knowledge and troubleshooting expertise. Without you my journey into Bioinformatics would have been impossible. I also would like to thank the rest of the members of the Rockefeller Bioinformatics Resource Core - Ji-Dung Luo, Matthew Paul, Doug Barrows and Wei Wang. I am so grateful for Svetlana Mazel and her incredible team - Selamawit Tadesse, Stanka Semova, Songyan Han and Samer Shalaby - at The Rockefeller University Flow Cytometry Resource Center. Their professionalism and expertise are unmatched and made for a great time sorting nuclei for hours. I would like to also thank Connie Zhao and Christine Lai from The Rockefeller University Genomics Resource Center and Rada Norinsky from the Rockefeller University Transgenic and Reproductive Technology Center.

Last but not least, I would like to thank all my family and friends for putting up with me throughout this process. I know it wasn't easy, but I love you for it. I couldn't have done this without your support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

5caC = 5-carboxycytosine
5fC = 5-formylcytosine
5hmC = 5-hydroxymethylcytosine
5mC = 5-methylcytosine
ATAC-Seq = assay for transposase accessible chromatin using sequencing
BER = base-excision repair
BS = sodium bisulphite
OxBS = oxidative bisulphite
ChIP-Seq = chromatin immunoprecipitation using sequencing
CNS = central nervous system
CpG = cytosine-guanine dinucleotide
DAPI = 4',6-diamidino-2-phenylindole
DMR = differentially methylated regions
DNA = deoxyribonucleic acid
DNMT = DNA methyltransferase
MECP2 = methyl CpG binding protein 2
MBD = methyl-binding domain
mRNA = messenger RNA
RNA = ribonucleic acid
TDG = thymine DNA glycosylase
TET = ten-eleven translocation
TRAP-Seq = Translation ribosome affinity purification and sequencing
TSS = transcription start site
FANS-Seq = Fluorescently activated nuclear sorting and sequencing

CHAPTER 1. Introduction

1.1 Neuronal Diversity: From Purkinje Neurons to Pyrimidines

The brain is a complex and diverse structure that contains multiple cell types with unique molecular, morphologic and functional properties. These differences have fascinated scientists since Ramon y Cajal completed his beautiful illustrations of neuronal populations (Ramón y Cajal 1992) and to this day are the main inspiration behind research programs around the world. Neurons range from becoming postmitotic as late as three weeks post birth and remaining relatively small with few projections, to exiting the division cycle as early as two weeks post conception and developing long axonal projections and elaborate dendritic arbors. The regulatory processes leading to this tremendous amount of diversity has sparked numerous studies from differential histone tail modifications to transcription factor binding and RNA interference. All of these studies point to one fact – each specific cell type has a unique interplay of regulatory genetic, epigenetic and molecular mechanisms leading to the establishment of its identity.

A study of the composition of DNA methylation of the cerebellar Purkinje and granule cells (Kriaucionis and Heintz 2009) led to discovery of a new form of cytosine modifications – DNA hydroxymethylation (5hmC). It was further shown that while DNA methylation is present at stable levels in multiple tissues and organs, DNA hydroxymethylation is most abundant in the central nervous system and it is strongly correlated with active gene expression, and its abundance and genomic localization is highly cell-type specific. At this point in time, I started working on my thesis, trying to decipher why 5hmC is so important for the nervous system. Many questions entertained our imagination – does 5hmC promote transcription? Is its role to reverse repression? Why are there three enzymes writing this mark? Why the nervous system? Taken together, all these questions were the motivating force behind my thesis project over the last 6 years.

1.2 Profiling the brain at a cell-type specific level

Scientists have attempted to characterize the neuronal diversity in great detail through molecular and electrophysiological tools, but one big hurdle remains – how do we isolate homogenous populations of neurons? Some approaches, such as the Translating Ribosome Affinity Purification and Sequencing (TRAP-Seq) method have allowed us to profile specific rare or abundant neuronal translatomes, but it requires difficult to scale transgenic mouse lines and is not applicable to humans (Heiman et al. 2008; Doyle et al. 2008). On the other hand, single cell sequencing has alleviated some of these issues. The technique can be applied to any species, regardless of availability of transgenic animals or suitable probes, but it is often quite expensive and profiling rare populations of neurons is a formidable endeavor(Hwang, Lee, and Bang 2018). Fluorescently activated nuclear sorting and sequencing (FANS-Seq) is somewhat in the middle, where with the aid of cell-type specific antibodies or probes, nuclei from any cell type can be isolated in bulk, which allows for not only transcriptome profiling, but also epigenetic mapping of the forms of accessible chromatin, DNA methylation and histone marks (X. Xu et al. 2018).

Combining these techniques, scientists have been able to decipher the transcriptomes of many neurons and compile massive atlases of neuronal population classifications. The big question of how these transcriptomes are regulated still remains. We have very little information on epigenetic regulation, such as the complex interplay of DNA and histone tail modifications,

coupled with transcription factor dynamics and RNA interference, in the CNS or how proper developmental transitions and terminal differentiation are accomplished.

1.3 Control of gene expression

Chromatin organization

The organization and structure of DNA inside the nucleus is of crucial importance to the regulation of gene expression. DNA polymers, wrapping around histone octamers, form nucleosomes, which are coiled and condensed into chromatin fibers. Chromatin can be subdivided into two main categories – heterochromatin and euchromatin, referring to either transcriptionally silent or active regions of the genome. Heterochromatic regions are physically compressed, through DNA methylation or repressive complexes such as PRC2 and histone marks like H3K27me3 (Simon and Kingston 2009; Boyer et al. 2006), thus restricting access to transcription factor binding sites and promoters. Euchromatin, on the other hand, is marked by high accessibility and active histone marks such as H3K4me3.

DNA methylation

DNA methylation is an ancient form of gene expression regulation present in many species (Zemach et al. 2010). However, cytosine methylation has been lost in certain eukaryotes like *Caenorhabditis elegans, Drosophila melanogaster* and yeast, partly due to its mutagenic property to spontaneously deaminate causing a cytosine to thymine transition (Holliday and Grigg 1993). Roughly 60-80% of all mammalian CpG sites are methylated and regulate multiple pathways in development and disease such as repression of transposable elements, X-chromosome inactivation, genomic imprinting, heritability and most cancers (Smith and Meissner 2013; Baylin and Jones 2016). There are three stages of cytosine methylation - *de novo* establishment, maintenance and demethylation. De novo DNA methylation is established by the enzymes DNMT3A and DNMT3B and can occur in any context such as CpG or CpH (H = A, T or C) (Okano et al. 1999, 3). Conversely, methylation maintenance can occur only in CpG context as DNMT1, the maintenance DNA methyltransferase, recognizes only hemi-methylated cytosines (Hermann, Goyal, and Jeltsch 2004). This ensures re-establishment of methylation on the newly synthesized strands and preserving epigenetic information during the mitotic phase of development. The genomic distribution of 5mC is diverse. Typically, promoter methylation is correlated with expression suppression as it prevents transcription factors and machinery from binding, while gene body methylation is more diverse and cell-type specific (Stadler et al. 2011). Methylation of cytosines contributes to silencing in two ways. First, out of 542 transcription factors surveyed, 22% showed a lowered biding affinity for their motifs when methylated, suggesting it impedes transcription activation at promoters (Yin et al. 2017). Second, 5mC itself is a binding site for multiple proteins that are part of heterochromatin formation, such as the methyl-CpG-binding domain (MBD) proteins (Meehan et al. 1989). There are 5 different methyl binding proteins MBD1-4 and MeCP2. MeCP2 is a canonical repressor of gene expression, binding methylated cytosines and recruiting other repressive complexes with roles in heterochromatin condensation (Della Ragione et al. 2016, 2). However, MeCP2 has been reported to have roles in other functions such as binding DNA hydroxymethylation in non-CpG context in postmitotic adult neurons, previously unreported in other tissues and cells (Mellén, Ayata, and Heintz 2017). Furthermore, adding more complexity to an already complicated

protein, MeCP2 has been reported to be a transcriptional activator when coupled to different complexes. MeCP2 is also the gene disrupted in Rett syndrome, an epigenetic neurodevelopmental disorder (Katz et al. 2016)

1.4 DNA hydroxymethylation and the TET proteins

Discovery of 5hmC and the TET proteins

In 2009, while investigating the peculiar organization of the highly euchromatic nucleus of Purkinje cells, the Heintz lab discovered a new stable and abundant DNA modification – 5-hydroxymethylcytosine (Kriaucionis and Heintz 2009). Not only did this study show that this mark is abundant in neurons, but also that granule and Purkinje neurons accumulate significantly different amounts of it suggesting cell type specific difference in the function and dynamics of 5hmC. At the same time, the Rao and Aravind labs were studying base J-binding proteins (JBPs) and discovered that Tet1,2 and 3 are their homologs and are predicted to be 2-oxoglutarate and Fe(II) dependent dioxygenases, which can oxidize 5-methylcytosine (**Figure 1.1A-B**) (Iyer et al. 2009). This hypothesis was verified by using a recombinant Tet protein and it further displayed that Tet can oxidize 5hmC further to 5-formyl and 5-carboxylcytosine (Tahiliani et al. 2009; Ito et al. 2011). It is yet unknown whether 5hmC and its oxidation derivatives are only intermediates in the DNA demethylation cycle or if they have readers with particular downstream functions. A quantitative mass spectroscopy proteomics study in mouse embryonic stem cells, neuronal progenitor cells and adult mouse brain tissue attempted to shed some light on this issue. They demonstrated multiple binding partners for each modification such as Uhrf2, Zhx2, Thap11, the Rfx proteins, Klf4, including multiple proteins involved in DNA repair (Spruijt et al. 2013). While a caveat to the study is that it does not distinguish whether a protein binds directly or indirectly to the modified cytosine, it does strengthen the involvement in DNA demethylation hypothesis as multiple DNA repair proteins are found to bind the modified cytosines.

Since its discovery, 5hmC and the Tet proteins have become central to many cancer and developmental studies, yet their role in the central nervous system still remains elusive.

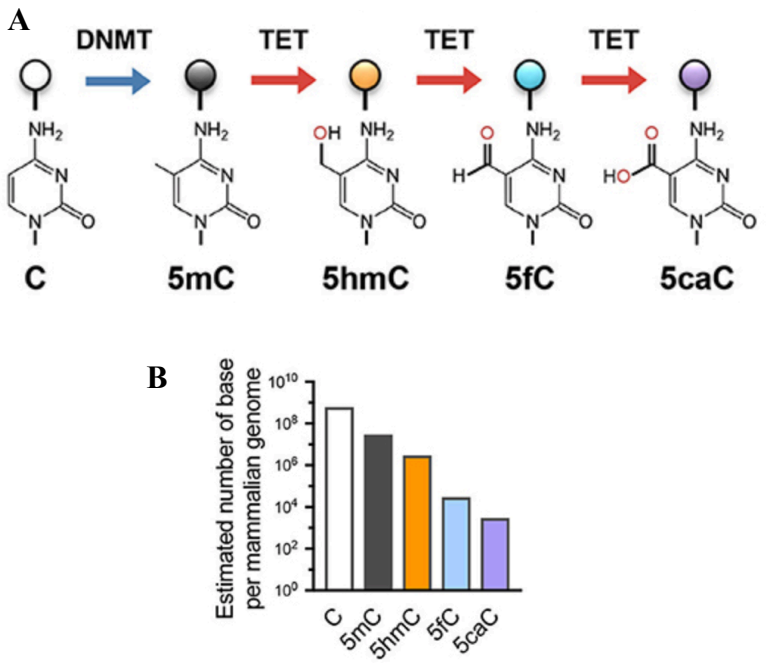


Figure 1.1. The Tet proteins oxidize 5-methylcysotine to its further derivatives 5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxycytosine.
A. The DNA methyltransferase (DNMTs) protein family coverts unmodified cytosine to 5-methylcytosine. The TET protein family oxidizes 5-methylcytosine to 5-hydroxymethylcytosine, 5-formylcytosine and 5-cabroxycytosine. **B**. Approximate abundance of modified and unmodified cytosines in the human/mouse genome. 5-methylcytosine is roughly 5%, mostly present in CpG context. 5-hydroxymethylcytosine is roughly 1-10% of the methylated cytosines depending on the cell type. 5-formylcytosine and 5-carboxycytosine are in much lower amounts – about an order of magnitude lower than the previous oxi-cytosine form, either due to lower processivity of the TET proteins or removal by TDG/BER pathway. Adapted from (C.-W. J. Lio and Rao 2019).

Major role of 5hmC in DNA demethylation

Removal of 5mC can happen in various ways in different species and cell-types. For example, in plants 5mC can be excised directly by the ROS1/DME family of DNA glycosylases and then replaced by the BER pathway (J.-K. Zhu 2009). There are no known homologues to the ROS1/DME proteins in mammals, yet three distinct mechanism of DNA demethylation have been described – active, passive and functional.

Active demethylation is a replication independent process, where oxidized cytosines like 5-fC and 5-caC are excised by thymine DNA glycosylase (TDG), a DNA repair protein, followed by the base excision repair pathway to replace the abasic sites with unmodified cytosines (He et al. 2011; Maiti and Drohat 2011). It has been confirmed by multiple studies that *in vitro* TDG can excise 5fC and 5caC but not 5hmC (Maiti and Drohat 2011; L. Zhang et al. 2012; Hashimoto, Hong, et al. 2012). These reactions have also been replicated using purified proteins from the BER pathway and TDG, demonstrating an excision of 5fC and 5caC, generation of an abasic site, followed by conversion to a single stranded break and repair (Weber et al. 2016).

Passive demethylation, on the other hand, is a replication dependent process and can only occur in actively dividing cells. When 5mC is oxidized to 5hmC and a cell divides, DNMT1, the maintenance DNA methyltransferase, does not recognize the hemi-hydroxymethylated cytosine so a remethylation event does not occur. This has also been shown for the other two oxi-mC forms 5fC and 5caC (Hermann, Goyal, and Jeltsch 2004, 1; Hashimoto, Liu, et al. 2012). After a couple of cycles of replication, all previous oxidized sites are replaced by unmodified cytosines.

Functional demethylation refers to the oxidation of 5mC to 5hmC, without subsequent replication or excision. MeCP2, a methyl-binding repressive protein, is shown to strongly bind 5mCpG, but not 5hmCpG. Thus, an oxidation event of 5mCpG removes a binding site and leads to reduced presence of repressive proteins and increased transcription (Mellén et al. 2012; Mellén, Ayata, and Heintz 2017).

While both active and passive DNA demethylation have been reported and characterized in zygotes, primordial germ cells and embryonic stem cells, the only established mechanism in neurons is functional DNA methylation. Passive DNA demethylation is impossible in neurons as they are postmitotic, whereas active DNA demethylation is strongly believed to occur, but is has not yet been proven.

5hmC as a regulator of embryonic development

Multiple studies have shown that the Tet proteins have important roles very early in development, in particular at stages that require massive demethylation events and *de novo* methylation. Right after fertilization, while the DNA from the egg and sperm have not yet fused, the male pronucleus becomes demethylated through a Tet3-dependent pathway (Wossidlo et al. 2011; Gu et al. 2011). Tet3-deficient morula have lower levels of Oct4, yet they form blastocysts normally, but fail to develop completely due to massive abnormalities by E11.5 (Gu et al. 2011).

Tet1-deficient ES cells and Tet1- and Tet2-deficient ES cells have been reported to stably form large hemorrhagic teratomas that have an enrichment of trophoblast cells, showing that Tet1 has a regulatory role in lineage commitment. This is likely due to a disruption of the transcriptional profile of ES cells as the Tet1 depletion leads to altered expression patterns of transcription factors such as Cdx2, Eomes, Elf5, Pax6 and NeuroD2 (Dawlaty et al. 2011; 2013).

Ultimately, both the Tet1-deficient and Tet1- and Tet2-deficient mice develop normally, suggesting a minimal role for the enzymes or that Tet3 can efficiently compensate for their loss.

During a global event resulting in DNA modification alterations between E9.5 and E10.5, 5mC is converted to 5hmC by Tet1 and Tet2, without any detectable oxidation to 5fC or 5caC by immunohistochemistry, followed by passive demethylation by replication (Hackett et al. 2013; Vincent et al. 2013). Further studies in primordial germ cells (PGCs) show that a loss of Tet1 leads to expression changes in over 1000 genes, including a significant number of meiosis genes (Yamaguchi et al. 2012). This resulted in female gametes having a defect in the meiotic synapsis, becoming developmentally arrested, and ultimately leading to smaller litter sizes.

The exploration of demethylation events in embryonic development suggests that they occur in a similar way. Tet3 in zygotes and Tet1 and Tet2 in PGCs oxidize 5mC to 5hmC, which is then followed by dilution through replication. This showcases the role of 5hmC in antagonizing maintenance DNA methyltransferases and promoting passive DNA demethylation.

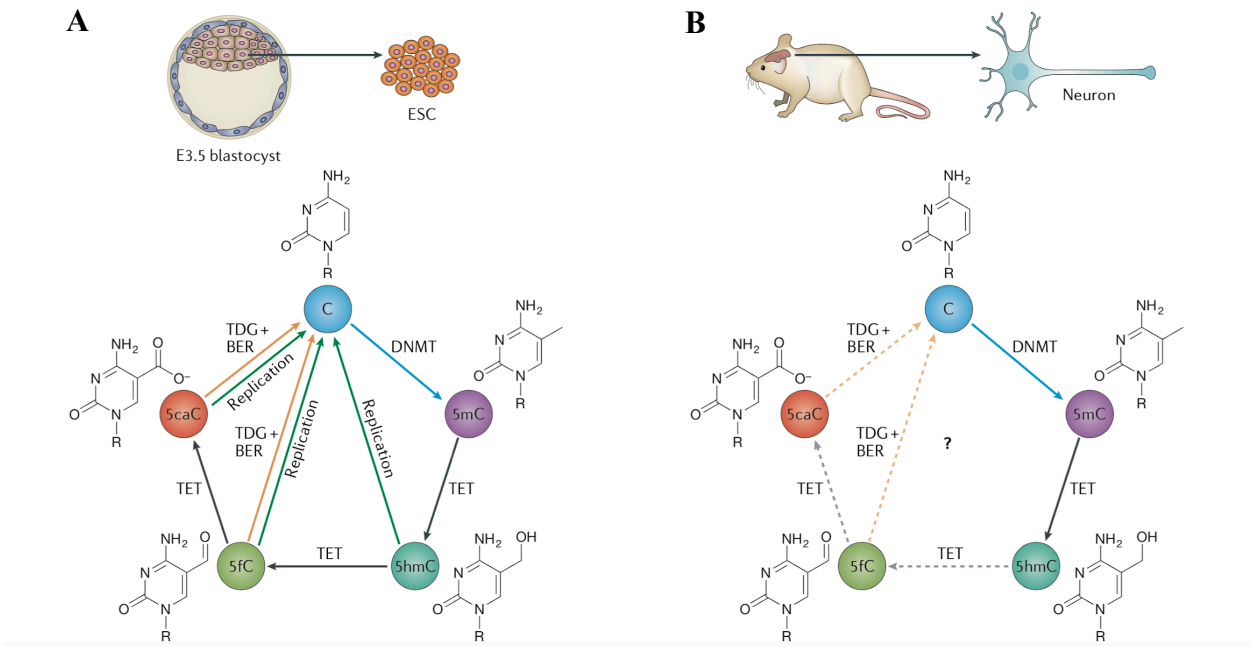


Figure 1.2. Active and passive DNA demethylation pathways in distinct cell types. **A**. DNA demethylation pathway in embryonic stem cells (ESCs), where all the necessary proteins for active demethylation, TDG and BER, are expressed, and the cells are actively dividing, which allows for passive demethylation to occur as well. **B**. DNA demethylation pathway in postmitotic neurons. Some neurons express all the necessary proteins for active demethylation (TDG and BER), however there is very little evidence that active demethylation occurs. 5-formylcytosine and 5-carboxycytosine are present at very low levels, while 5-hydroxymethylcytosine is much more abundant. Passive demethylation does not occur as neurons are postmitotic. Adapted from: (Wu and Zhang 2017)

5hmC and the TET proteins are implicated in multiple blood disorders

The Tet proteins were named after the ten-eleven translocation that occurs in certain cases of acute myeloid and lymphocytic leukemia, where the *Mll1* gene from chromosome 10 is fused to *Tet1* on chromosome 11 (Pastor, Aravind, and Rao 2013). Loss-of-function mutations of the Tet proteins are commonly found in hematopoietic cancers, causing the increased methylation of genomic regulatory regions, often promoters and enhancers of tumor suppressor or DNA damage genes, which results in impaired transcription leading to oncogenesis. This notion is confirmed by multiple studies from different tissues and cell types that show that the role of the Tet enzymes is to be recruited to regulatory regions, oxidize 5mC to 5hmC and thus cause DNA demethylation, opening the binding site to activators or repressors of gene expression. In a recent study of mature B cells, researchers found that TET enzyme activity regulates class switch recombination by regulating the expression of *Aicda,* through demethylation of enhancer elements and facilitating chromatin accessibility (C.-W. J. Lio et al. 2019). In regulatory T cells, the Tet proteins are required for demethylation of CpG islands in the Foxp3 conserved noncoding sequence 2 intronic element, which is responsible for the stable expression of the gene (Sasidharan Nair, Song, and Oh 2016). A double knock out of Tet2 and Tet3 leads to a disruption in the Treg cell signature genes expression leading to an unstable Treg cell identity (Yue et al. 2019). Another double depletion of Tet2 and Tet3 in T cells results in disrupted development and expansion of self-reactive NKT cells and lymphoma, with genes showing increased DNA methylation and changes of expression of developmentally important genes (Tsagaratou et al. 2017). A Tet2 and Tet3 double knockout in B cells disrupts the pro- to pre-B cell transition in bone marrow and impacts the Igκ locus transcription and rearrangement, through impaired demethylation and accessibility of the enhancer locus (C.-W. Lio et al. 2016).

5hmC is a regulator of gene expression in the brain

While an impressive number of comprehensive studies have dissected the role of 5hmC and the Tet proteins in hematopoietic cells, very little is known about their function in the central nervous system (CNS). CNS cell types have shown consistently the highest amount of 5hmC among other tissues, which is partly a result of the long lifespan of non-dividing neurons, precluding the mark's dilution through division. Studies have shown that a strong correlation between 5hmC and active gene expression, with a high accumulation of the DNA modification over the gene bodies rather than in the promoters (Mellén et al. 2012). The conversion of 5mC to 5hmC in active genes leads to a functional demethylation of their gene bodies, as most repressors with methyl-binding domains do not recognize 5hmC and thus can no longer silence expression (Mellén, Ayata, and Heintz 2017). In a hippocampal Tet1 knockout mouse model, researchers have shown hypermethylation of multiple promoter regions, which caused aberrant expression levels of many neuronal activity regulated genes. This disruption of proper gene expression resulted in abnormal hippocampal long-term depression and impaired memory function (Rudenko et al. 2013). A follow-up study, in which Tet1 was overexpressed, showed global methylation changes again with impaired memory formation, suggesting that a fine balance of DNA modifications is crucial for the proper function of neurons (Kaas et al. 2013). In a study of the role of epigenetics regulation during cerebellar development and circuit formation, it was shown that axon guidance and ion channel genes accumulated significant amounts of 5hmC. A double knockdown of Tet1 and Tet3 in cerebellar granule cells revealed that hydroxymethylation is required for the dendritic arborization of neurons and strengthen the evidence that the Tet

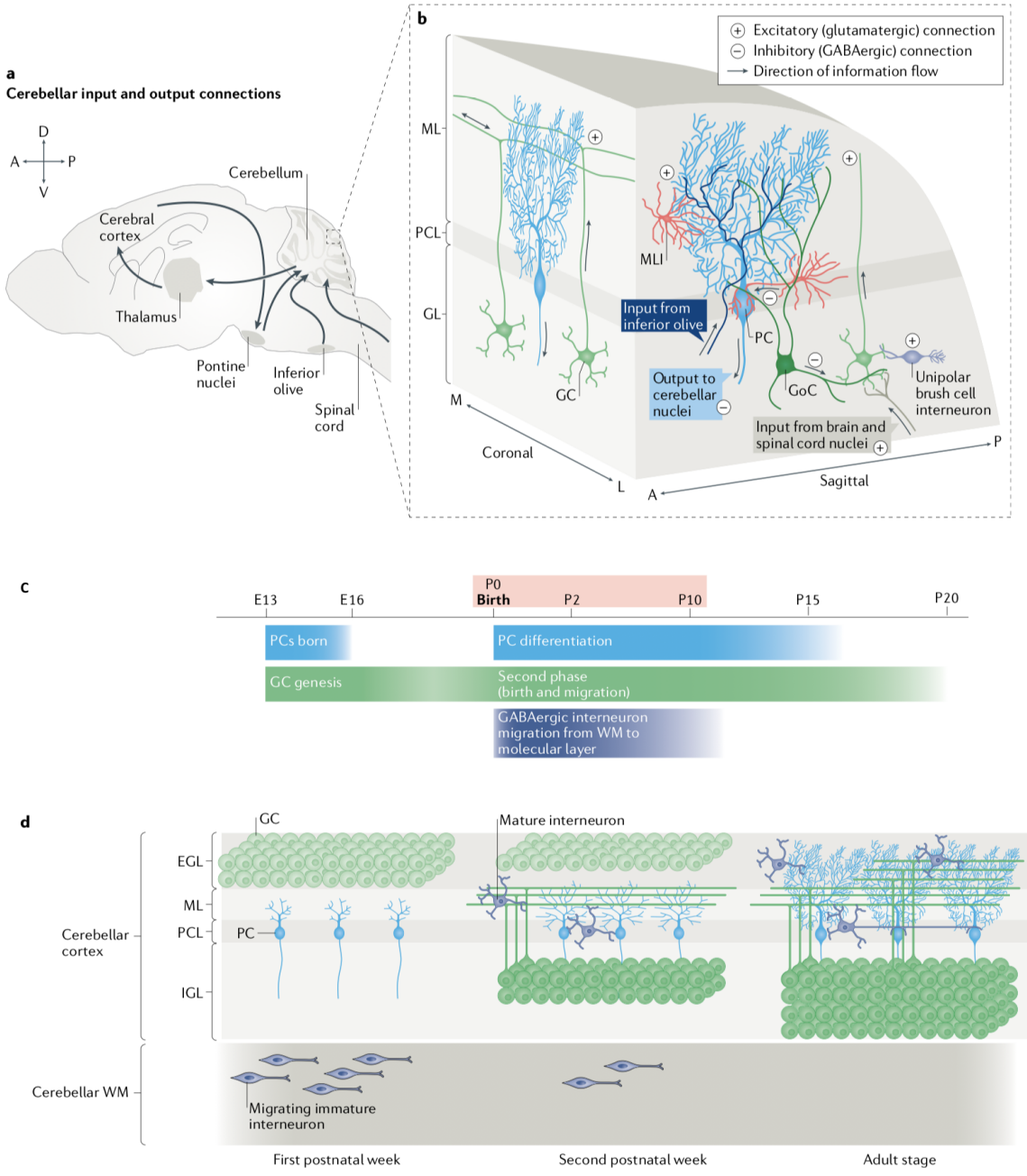
genes play a critical role in neuronal differentiation (X. Zhu et al. 2016). Much earlier in development, a Tet3 depletion leads to impaired transition from embryonic stem cells to neuronal progenitors, and ultimately results in a disrupted terminal neuronal differentiation (T. Li et al. 2015). Besides a clear role in the proper differentiation of neurons, it has been shown that Tet3 and 5hmC were upregulated following a sciatic nerve lesion, suggesting a potential role of DNA demethylation in axon regeneration (Weng et al. 2017). While most of these studies were conducted in mouse, a *Xenopus laevis* investigation which knocked down the expression of Tet3 showed a profound developmental defect resulting in small heads and no eyes (Y. Xu et al. 2012).

Recently, a human genetics study showed that non-coding and loss-of-function coding variants in Tet2 are associated with early-onset Alzheimer's disease and frontotemporal dementia (Cochran et al. 2020). In a Parkinson's disease (PD) study, Tet2 has been shown to be upregulated in humans, leading to widespread increase of enhancer DNA modifications (Marshall et al. 2020). Further investigation in a PD mouse model shows that a Tet2 loss prevents nigral dopaminergic neuronal loss. This is in line with multiple aging studies that show age-related changes in DNA methylation and positions the Tet proteins as not only key players in development but also at all stages of neuronal functions and potentially as therapeutic targets. However, while studies have been exploring individual Tet proteins' role in different neuronal cell types, no comprehensive triple Tet protein knockout has been reported to demonstrate their function and significance conclusively.

1.5 The Cerebellum

The main role of the cerebellum, which contains over half of the neurons in the adult brain, is to coordinate the sensory-motor processing (Sathyanesan et al. 2019). The granule cells receive input from the inferior olive, then synapse onto the Purkinje cells, which in turn send their output through the cerebellar deep nuclei to the thalamus and cerebral cortex (Figure 1.4A). The morphological structure of the cerebellar cortex consists of a Purkinje neuron monolayer, which on one side is surrounded by a thick granule cell layer, and on the other the Purkinje dendrites project into the molecular layer that also harbors multiple interneurons such as the basket, Lugaro, Golgi and unipolar brush cells (Figure 1.4B) (Sathyanesan et al. 2019). The developmental trajectories of the two principal neurons of the cerebellum are distinct. Purkinje cells are born around E13.5 in the rhombic lip and become postmitotic immediately. On the other hand, granule cells are dividing up until P21, with their peak proliferation period at P7 (Figure 1.4C). During the first postnatal week, the granule cells are located subpially in the external granule layer (EGL). During the second postnatal week they enter a phase of exponential division phase and migration towards the inner granule layer (IGL). This is accompanied by the growth of dendritic arbors of the Purkinje cells and migration of mature interneurons. By the adult phase, all of the granule cells have migrated from the EGL, Purkinje cells have developed their full dendritic arbors and developed synapses (Figure 1.4D) (Sathyanesan et al. 2019). Overall, the differences between the two principal neurons make the cerebellum an excellent structure for molecular studies of the brain.

Figure 1.3. Cerebellar development, cell types, features and circuitry.
A. Schematic of the main cerebellar input – coming from the spinal cord, pontine nuclei and inferior olive – and outputs – to cerebral cortex, through the thalamus. **B**. Cerebellar cell type composition – Purkinje cells (light blue), granule cells (light green), molecular layer interneurons (pink), Golgi cells (dark green), unipolar brush cells (purple), inferior olive input (dark blue), brain and spinal cord input (grey), arrows show the direction of information flow, + refers to excitatory glutamatergic connection, - refers to inhibitory GABAergic connection. **C**. Timeline of cellular birth periods, differentiation and migration. **D**. Expanded cellular view of the events in (**C**). EGL – external granule layer, PCL – Purkinje cell layer, IGL – internal granule layer, ML – molecular layer. Depicted are the granule cell expansion in the EGL, dendritic arborization of the Purkinje cells and migration of the interneurons during the first postnatal week. Migration of the granule cells towards the IGL and synapse formation onto Purkinje cells, during the second postnatal week. At the adult stage, granule cells, interneurons and Purkinje cells are fully integrated into the cerebellar circuit. The EGL has completely disappeared. P0 – postnatal day 0, A – anterior, P – posterior, D – dorsal, V – ventral. Adapted from (Sathyanesan et al. 2019)


a
Cerebellar input and output connections
D
A
P
V
Cerebellum
Cerebral cortex
Thalamus
Pontine nuclei
Inferior olive
Spinal cord
b
Excitatory (glutamatergic) connection
Inhibitory (GABAergic) connection
Direction of information flow
ML
PCL
GL
MLI
PC
GC
GoC
Input from inferior olive
Output to cerebellar nuclei
Unipolar brush cell interneuron
Input from brain and spinal cord nuclei
M
L
Coronal
A
P
Sagittal
c
E13
E16
P0
Birth
P2
P10
P15
P20
PCs born
PC differentiation
GC genesis
Second phase (birth and migration)
GABAergic interneuron migration from WM to molecular layer
d
GC
Mature interneuron
EGL
ML
PCL
IGL
PC
Cerebellar cortex
Cerebellar WM
Migrating immature interneuron
First postnatal week
Second postnatal week
Adult stage

1.6 Methodological overview

FANS-Seq

Fluorescently activate nuclear sorting and sequencing (FANS-Seq) is a methodology where cell type specific nuclei of wild type animals can be purified through antibody labeling and sorting (X. Xu et al. 2018). It can be applied to any species, where suitable antibodies exist and has been successful in mouse, rat, monkey and human. The technique, unlike traditional RNA sequencing, requires the extracted nuclei to be preserved with a fixative, but the quality of the RNA molecule is good enough for a deep molecular profiling. FANS-Seq is further compatible with downstream applications such as ATAC-Seq, OxBS-Seq and ChIP-Seq.

OxBS-Seq

Multiple methods for detection of 5hmC have been developed, from immunostaining in sections, to mass spectrometry and sequencing. The genomic localizations methods stem from two main approaches - DNA precipitation or a single nucleotide resolution sequencing. The DNA precipitation methods are the following:

- 5hmC, 5fC and 5caC immunoprecipitation – This method is akin to ChIP-Seq, where a specific antibody that recognizes the modified cytosines is used to pull down regions of enrichment. However, the pulldown efficiency is strongly depended on the density of modified bases and could be subject to antibody batch effect (Pastor, Aravind, and Rao 2013).
- 5-hmC selective chemical labelling (hME-Seal) – In this method, a glucose molecule is attached to 5hmC using the beta-glucosyl transferase enzyme. The glucose is then coupled to biotin using click chemistry and precipitated with streptavidin beads (C.-X. Song et al. 2011).
- CMS immunoprecipitation – In this method, a sodium bisulphite treatment of 5hmC results into cytosine methylene sulphonate (CMS), which is later used as an antigen for the pull down (Pastor, Aravind, and Rao 2013).

The single nucleotide resolution methods are:

- Bisulfite sequencing (BS-Seq) – In this method, treatment with sodium bisulphite reduces all cytosine species to two – 5mC and 5hmC are read as cytosine, and 5C, 5fC and 5caC are deaminated and read as thymine (Pastor, Aravind, and Rao 2013).
- Oxidative bisulfite sequencing (OxBS-Seq) – In this method, 5hmC is oxidized by potassium perruthenate (**Figure 1.5**). The same sample is split into two – one undergoes OxBS-Seq and the other BS-Seq, and then subtraction of the samples computationally can delineate 5hmC and 5mC (Pastor, Aravind, and Rao 2013).
- TAB-Seq – In this method, similarly to hME-Seal, 5hmC is protected by glucose and then all 5mC oxidized to 5caC by a recombinant Tet1 enzyme. The sample then undergoes bisulfite treatment and it creates a direct readout of 5hmC, without the necessity for computational separation of the 5mC and 5hmC signal (Pastor, Aravind, and Rao 2013).

- SMRT sequencing – This method, similarly to hME-Seal, a conjugate is added to all 5hmC bases. Since this method relies on polymerase stalling for base readout, a bulkier 5hmC allows for its identification (Pastor, Aravind, and Rao 2013).

ATAC-Seq

Assay for Transposase-Accessible Chromatin using sequencing (ATAC-Seq) is a technique that allows for the genomic profiling of chromatin accessibility (Chen et al. 2016; Buenrostro et al. 2013). It employs an enzyme, Tn5, that cuts the open space between two nucleosomes and can produce fragments that are equal to the nucleosome free region, or to a mono-nucleosome, di-, tri, etc. The technique is often used to identify accessible regulatory regions and transcription factor binding sites. Computationally, they can be used to identify differentially accessible sites that are changing in development and disease.

ChIP-Seq

Chromatin Immuno-Precipitation and Sequencing (ChIP-Seq) is a technique where antibodies against different histone tail modifications are used to bind and precipitate regions of enrichment (Barski et al. 2007). Those regions are sequenced and computational they can be used to bind differential enrichment sites, gene ontology of enriched gene sets and can address many other biological questions.

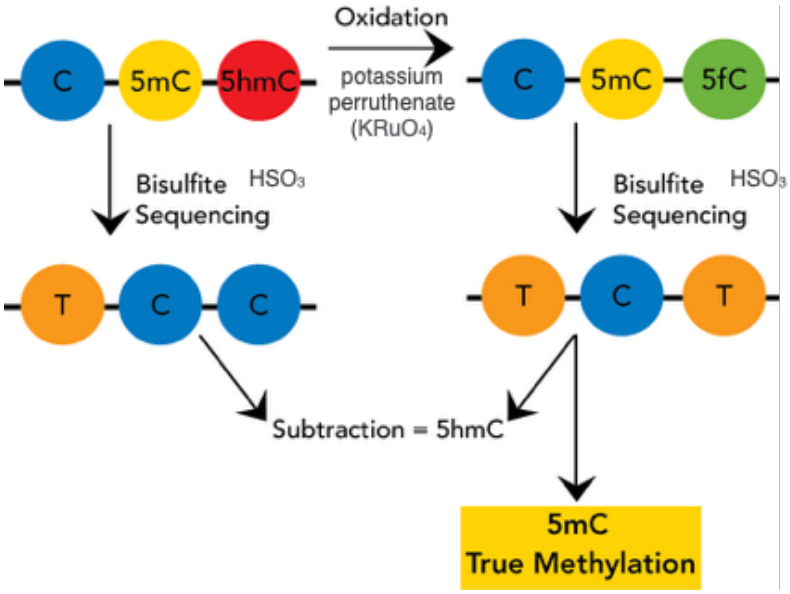


Figure 1.4. Oxidative bisulfite sequencing for simultaneous the detection of 5-hydroxymethylcytosine and 5-methylcytosine.
Schematic of OxBS-Seq and BS-Seq. OxBS-Seq is an augmented version of BS-Seq, which includes an oxidation by potassium perruthenate ($KRuO_4$) before the bisulfite (HSO_3) treatment. The oxidation step converts 5-hydroxymethylcytosine to 5-formylcytosines, while the bisulfite treatment protects 5-methylcytosine and deaminates unmodified cytosine and 5-formylcytosine to thymine. Once the two separate samples are sequenced, the true values for 5-methylcytosine and 5-hydroxymethylcytosine are derived computationally. Adapted from: (https://www.nugen.com/products/ultralow-methyl-seq-truemethyl-oxbs)

CHAPTER 2. Developmental dynamics of 5hmC in Purkinje neurons

2.1 Introduction

While much has been revealed about 5hmC and the Tet proteins and their function in embryonic development and certain blood cancers, our knowledge about their role in the central nervous system remains sparse, especially during the post-natal development of neurons. Differentiation of most neurons includes an extended postmitotic program in which dramatic increases in size and complexity are accompanied by accumulation of high levels of 5hmC (Mellén et al. 2012). For example, cerebellar Purkinje cells (PCs) complete their last division in mid-gestation, and they remain largely undifferentiated throughout embryogenesis. At birth, they enter a phase of postmitotic differentiation that continues until early adult life and includes elaboration of a large euchromatic nucleus, emergence of an exceedingly complex dendritic arbor, and formation of hundreds of thousands of synapses (Butts, Green, and Wingate 2014). There is evidence that the formation of 5hmC at exon start sites during circuit formation in the cerebellum is required for the proper expression of axon guidance and ion channel genes (X. Zhu et al. 2016). A knockdown of Tet1 and Tet3 in cultured cerebellar granule cultures showed an impaired dendritic arborization, a crucial part of circuit formation. While this study has shown how important 5hmC is in the development of granule neurons, they are a population that is actively dividing for a very long period, resulting of much of the 5hmC produced to be diluted through passive demethylation. The high levels of 5hmC characteristic of adult neurons vary according to cell type, and their genomic distributions are strongly correlated with gene expression, but we don't know how specifically these patterns are established during post-natal differentiation. Furthermore, genomic occupancy of the essential neuronal protein MeCP2 reflects its differential affinities for methylated and hydroxymethylated CG and CH dinucleotides (Mellén, Ayata, and Heintz 2017). Despite these findings, an essential function for the accumulation of 5hmC in postmitotic neurons has not been established. To investigate in detail the developmental patterns, dynamics and function of 5hmCG and 5hmCG accumulated, we decided to focus on the Purkinje neuron at different timepoints and explore the chromatin landscape through transcriptome profiling, methylation and hydroxymethylation, chromatin accessibility and two histone marks – H3K4me3 as activating and H3K27me3 as silencing.

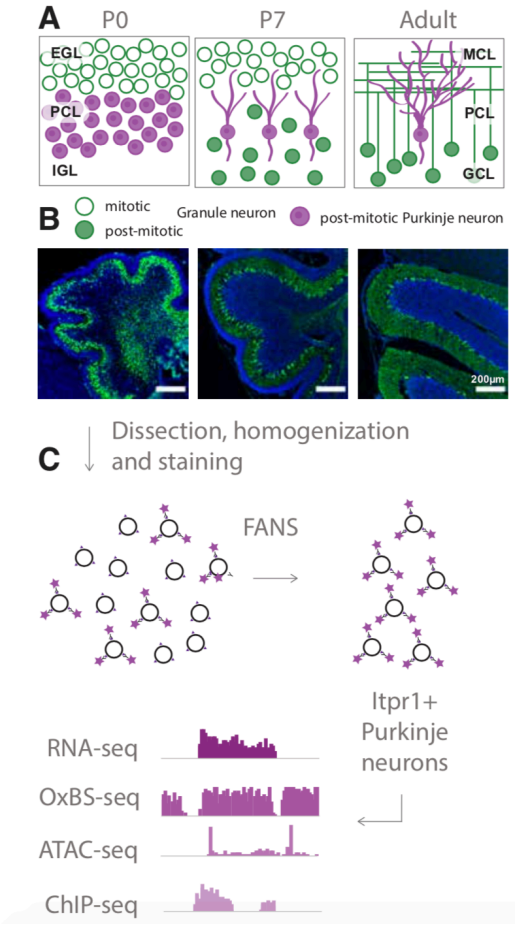


Figure 2.1. Schematic illustration of Purkinje nuclei isolation.

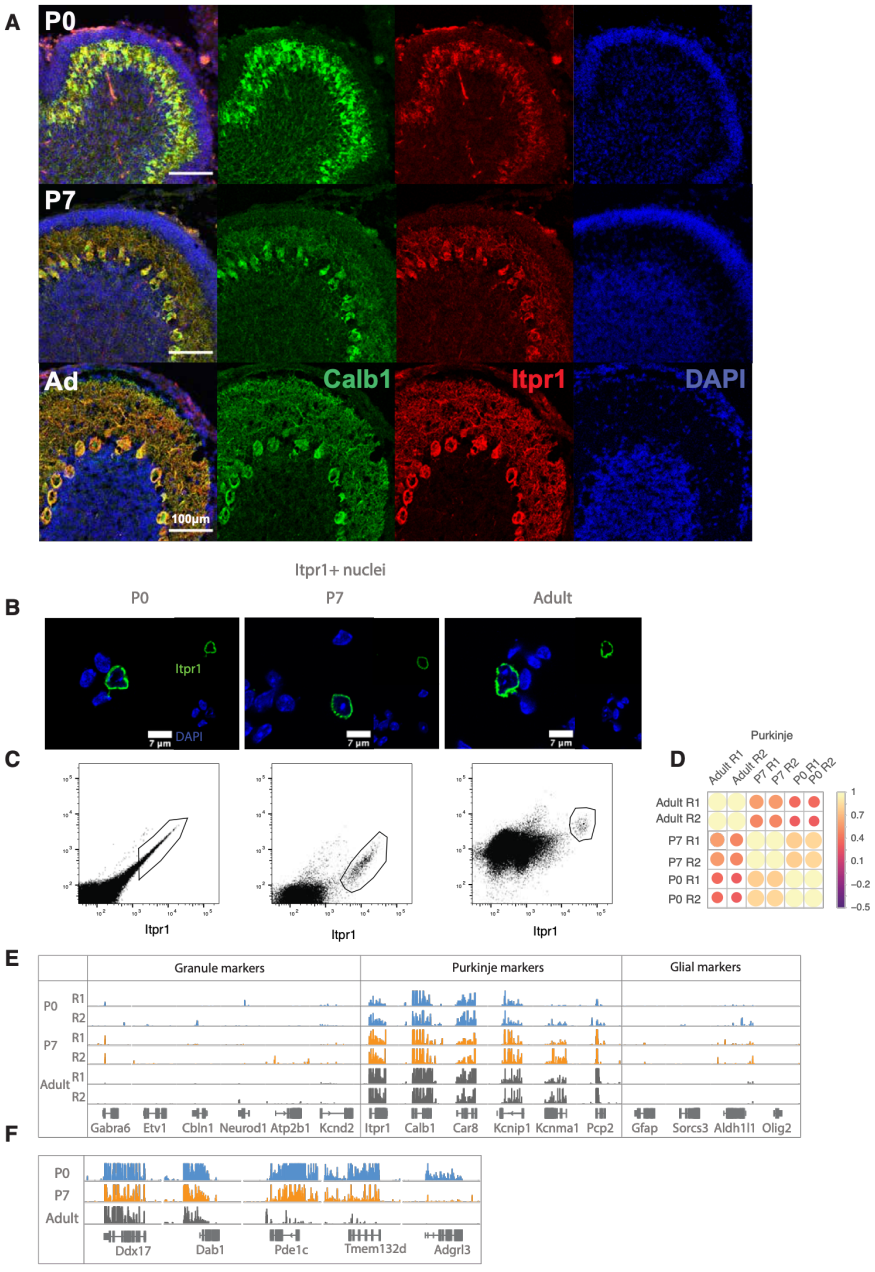
A. Schematic of PC differentiation and growth at P0, P7 and adult (approximately 8 weeks old) timepoints. EGL-external granule layer, PCL-Purkinje cell layer, IGL-internal granule layer, MCL-molecular cell layer, GCL- granule cell layer. **B**. Immunofluorescence staining with Calb1 (green) of PCs in murine cerebella at P0, P7 and adult timepoints. **C**. Workflow schematic of nuclei isolation, antibody staining with anti-Itpr1, fluorescence activated sorting, and downstream sequencing applications.

2.2 Chromatin landscape of differentiating Purkinje neurons

We chose to analyze Purkinje neurons at three important developmental stages: at the start of their accelerated differentiation (P0), during the rapid phase of their morphological development (P7) and as fully mature, differentiated neurons (adult) (**Figure 2.1A-B**) and characterize their patterns of 5hmC accumulation, transcription and chromatin on the path to terminal differentiation. To do so, we employed antibody assisted fluorescence activated nuclear sorting followed by sequencing (FANS-Seq) and performed RNA-Seq, OxBS-Seq, ChIP-Seq and ATAC-Seq (**Figure 2.1C**). FANS-Seq has already been utilized to purify Purkinje cells successfully using anti-Itpr1 antibodies (X. Xu et al. 2018). First, we confirmed that Itpr1 is a Purkinje specific marker by assessing its overlap with Calb1, which is a canonical Purkinje protein, and that it is highly expressed at all three timepoints through immunofluorescence (**Figure 2.2A**). Then, we isolated nuclei and stained them with Itpr1 and confirmed that it labels nuclei properly and can be used to purify nuclei using flow cytometry (**Figure 2.2B**). Finally, we employed FANS to purify Itpr1+ PC nuclei in two biological replicates at the three developmental stages (**Figure 2.2C**). Since PCs are extremely rare and difficult to isolate, we had to optimize low input protocols for genomic profiling. We used ~20,000 nuclei for transcriptional profiling, ~200,000 nuclei for bisulfite sequencing (BS-Seq) and oxidative bisulfite sequencing (OxBS-Seq), 25,000 nuclei for the assay for transposase-accessible chromatin sequencing (ATAC-Seq), and 25,000 nuclei for H3K27me3 and H3K4me3 chromatin immunoprecipitation sequencing (ChIP-Seq). We started with the transcriptome profiling in order to establish further that Itrp1 allows for the purification of Purkinje cells, without major contamination from other cell types such as glia and granule cells. First, we assessed the quality of the datasets by performing Pearson correlation (**Figure 2.2D**) and showing scores of over 0.9 for the pairs of replicates. Then, we checked for expression of markers from contaminating cell types (**Figure 2.2E**). The Purkinje datasets did not show any signs of contamination from glia or granule cells. Furthermore, the datasets appeared to be of high enough quality where at each developmental stage we could detect differential expression of genes and isoforms (**Figure 2.2F**). This gave us the confidence that the methodology is robust, so we continued with the generation of OxBS-Seq, ATAC-Seq and ChIP-Seq datasets. We performed Pearson correlation between the values for methylation and hydroxymethylation over gene bodies in both CpG and CpH context, showing high scores for each time point (**Figure 2.3A-B**). The quality of the OxBS-Seq data was further validated by examining the oxidation rates for the C to T conversion, which were all within the expected range (**Figure 2.3D**). Finally, the ATAC-Seq and ChIP-Seq datasets were validated using Pearson correlation of enrichment around the gene promoters, which showed scores of over 0.9 in each condition (**Figure 2.3C**).

Figure 2.2. Confirmation of Itpr1 as suitable marker for Purkinje nuclei isolation.

A. Immunofluorescence staining of Purkinje cells at P0, P7 and adult timepoints with Calb1 (green) and Itpr1 (red). **B.** Example of Purkinje cell nuclei stained with Itpr1 (green) post dissociation and pre-sorting, counterstained with DAPI, a heterochromatin marker. **C**. Example plots of fluorescence activated nuclear sorting of Purkinje cells at P0, P7 and adult timepoints. **D**. Pearson correlation of RNA-Seq datasets. **E**. IGV representation of Purkinje specific markers (Itpr1, Calb1, Car8, Kcnip1, Kcnma1, Pcp2) enrichment and depletion of granule (Gabra6, Etv1, Cbln1, Neurod1, Atp2b1, Kcnd2) and glial (Gfap, Sorcs3, Aldh1l1, Olig2) markers. **F** IGV representation of Purkinje genes dynamics during differentiation.


A
P0
P7
Ad
Calb1
Itpr1
DAPI
100μm
B
Itpr1+ nuclei
P0
P7
Adult
Itpr1
DAPI
7 μm
C
Itpr1
D
Purkinje
Adult R1
Adult R2
P7 R1
P7 R2
P0 R1
P0 R2
1
0.7
0.4
0.1
−0.2
−0.5
E
Granule markers
Purkinje markers
Glial markers
P0
P7
Adult
R1
R2
Gabra6
Etv1
Cbln1
Neurod1
Atp2b1
Kcnd2
Itpr1
Calb1
Car8
Kcnip1
Kcnma1
Pcp2
Gfap
Sorcs3
Aldh1l1
Olig2
F
Ddx17
Dab1
Pde1c
Tmem132d
Adgrl3

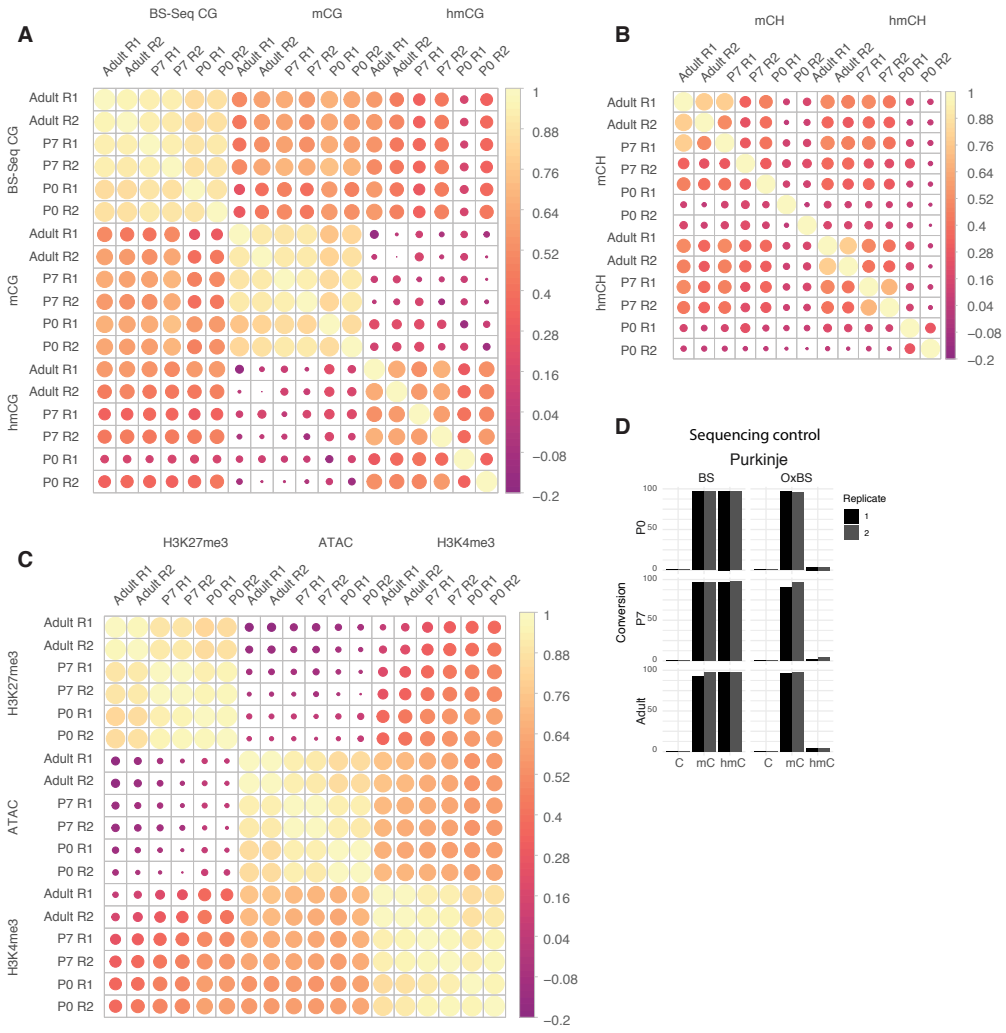


Figure 2.3. Quality control of sequencing datasets.
A. Pearson correlation of gene body accumulation of 5hmCG+5mCG (BS-Seq), 5hmCG and 5mCG (derived from maximum likelihood estimator model of BS-Seq and OxBS-Seq). **B**. Pearson correlation of gene body accumulation of 5hmCH and 5mCH (derived from maximum likelihood estimator model of BS-Seq and OxBS-Seq). **C**. Pearson correlation of ATAC-Seq, H3K4me3 and H3K27me3 ChIP-Seq dataset enrichment +/- 5kb around the gene promoters. **D**. Bisulfite conversion and oxidation efficiency in BS-Seq and OxBS-Seq datasets.

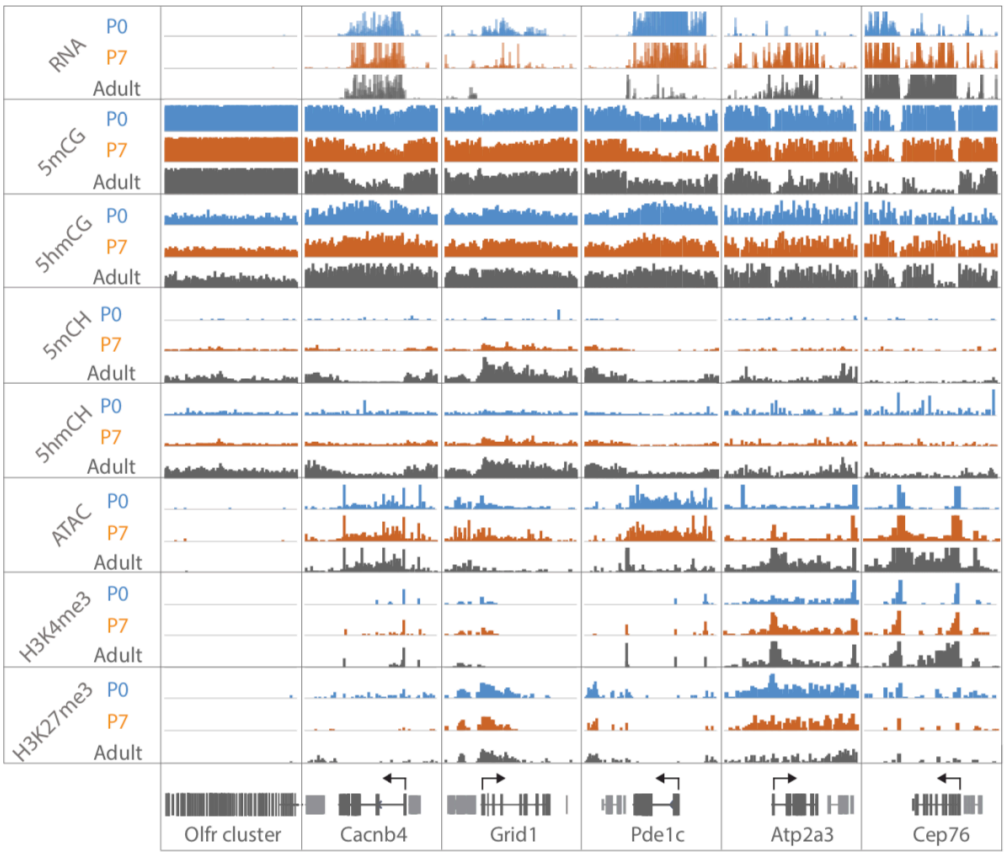


Figure 2.4. Genome browser representation of the Purkinje neuron chromatin landscape. Integrated genome viewer (IGV) representation of example regions of differentially regulated genes (Olfr cluster – always silent, Cacnb4 – always expressed, Grid1 and Pde1c – developmentally downregulated, Atp2a3 and Cep76 – developmentally upregulated). Top tracks show RNA expression in RPKM (reads per kilobase per million mapped reads), mCG tracks show methylation level in CG context from 0 to 0.8, hmCG tracks show hydroxymethylation level in CG context from 0 to 0.45, mCH and hmCH show methylation and hydroxymethylation level in CH context (H = A, C or T) from 0 to 0.04. ATAC tracks show ATAC-seq read density in RPKM from 0 to 10. H3K4me3 tracks show input normalized enrichment in RPKM from 0 to 5. H3K27me3 tracks show input normalized enrichment in RPKM from 0 to 3.

2.3 DNA hydroxymethylation is strongly correlated with gene expression in differentiating neurons

Once we confirmed the quality of all datasets, we embarked on exploring the gene expression regulation rules in the terminal differentiation of Purkinje cells. We identified multiple classes of constitutive expressed, silent or developmentally regulated genes, which show a wide range of changes in their DNA and histone modifications and chromatin accessibility during Purkinje neuron differentiation (**Figure 2.4**). Genes that are never expressed in the Purkinje lineage (e.g. Olfr clusters, which are expressed in olfactory neurons) have high levels of 5mCG, do not accumulate 5hmCG, and are not accessible. Genes that are actively transcribed throughout PC differentiation, defined as being consistently in the top quartile as ranked by expression at every single timepoint, (e.g. Cacnb4) have reduced levels of 5mCG, high levels of 5hmCG, they are accessible, and they are marked by the activating histone modification H3K4me3. While the epigenetic states of these classes of genes are stable, developmentally regulated genes are very dynamic.

We identified developmentally regulated genes by conducting a differential expression analysis between P0 and Adult Purkinje neurons, and filtering for significance of $p<0.01$ and log2 fold change of > 2 in either direction (**Figure 2.5A, Figure 2.6D-E**). We identified 432 developmentally upregulated genes (or genes up in Adult), which are enriched for genes involved in calcium ion transport, the inositol pathway and RNA splicing. The developmentally downregulated genes (or genes that were up in P0) were 922 and were enriched for genes in the cell signaling and axon guidance pathways (**Figure 2.5B**). Similar analysis comparing P7 and adult showed 299 upregulated genes and 478 downregulated ones. The analysis also displayed multiple categories enriched, with synaptic signaling and ion transport in P7, and RNA metabolism in adult (**Figure 2.6A-B**). A differential expression analysis between P0 and P7 timepoints yielded multiple changing genes (upregulated: 145, downregulated: 243), but there were no significant gene ontology categories (**Figure 2.6C**).

To explore the chromatin changes in developmentally downregulated genes, we compared them to two classes of repressed genes (always off): Hox transcription factors, which were expressed earlier in cell lineage and olfactory and vomeronasal receptors, which are never expressed in the Purkinje lineage. Hox transcription factors lack cytosine methylation or hydroxymethylation, are not accessible and are silenced through H3K27me3 histone modification enrichment. On the other hand, the olfactory and vomeronasal receptors are highly methylated, lack hydroxymethylation and are completely inaccessible as well (**Figure 2.6D**). Promoters of developmentally repressed genes in PCs become inaccessible, but this is not correlated with changes in DNA methylation or H3K27me3 occupancy **(Figure 2.7A)**. 5hmCG accumulated over the genes bodies early in development is not removed, suggesting that the presence 5hmCG is not sufficient to maintain transcription or initiate it, and the 5mCG levels remain stable (**Figure 2.5C-D)**.

Since we could not explain silencing through either loss of 5hmCG or significant accumulation of H3K27me3 occupancy, we decided to look into cytosine modifications in the CpG context. We did a k-means clustering analysis of the 5hmCH and 5mCG levels of developmentally downregulated genes and revealed at least two distinctly recognizable epigenetic patterns (**Figure 2.7B-D**). Those repressed genes that accumulate 5mCH and 5hmCH (modCHrich) are associated with lower promoter accessibility, accumulation of H3K27me3 and expression relative to those with low levels of CH modification (modCHpoor) (**Figure 2.7E-G**).

Although the data for modCHrich genes is consistent with repression through both polycomb repressive complexes and MeCP2, our data do not identify the mechanisms responsible for repression of the modCHpoor genes.

The chromatin landscape of developmentally upregulated genes is similar to those that are constitutively transcribed (**Figure 2.6E**). Their promoters become highly accessible, they accumulate high levels of H3K4me3, their gene bodies progressively lose 5mCG and gain 5hmCG, and they do not accumulate 5mCH or 5hmCH. That pattern of gene expression regulation in in accordance with previously reported studies (Mellén et al. 2012).

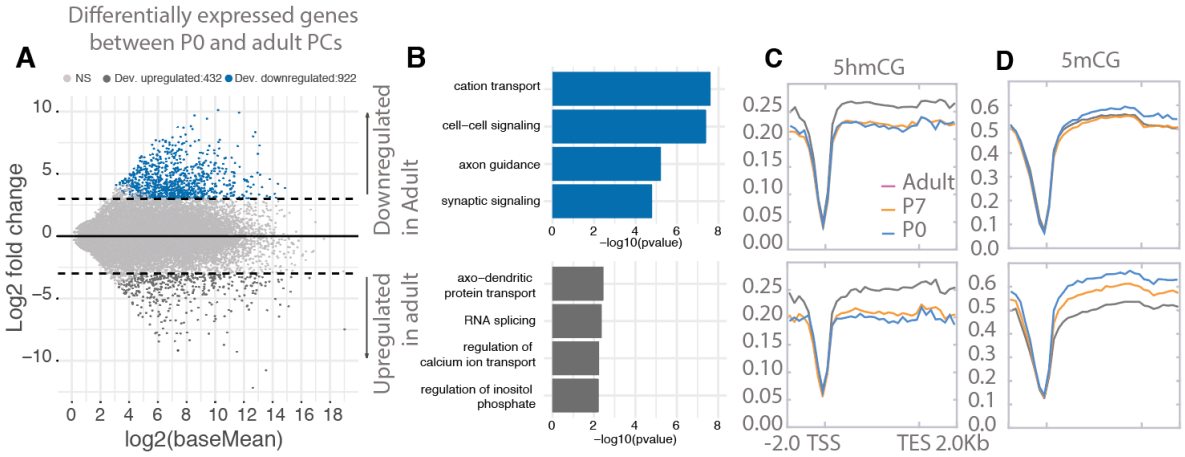
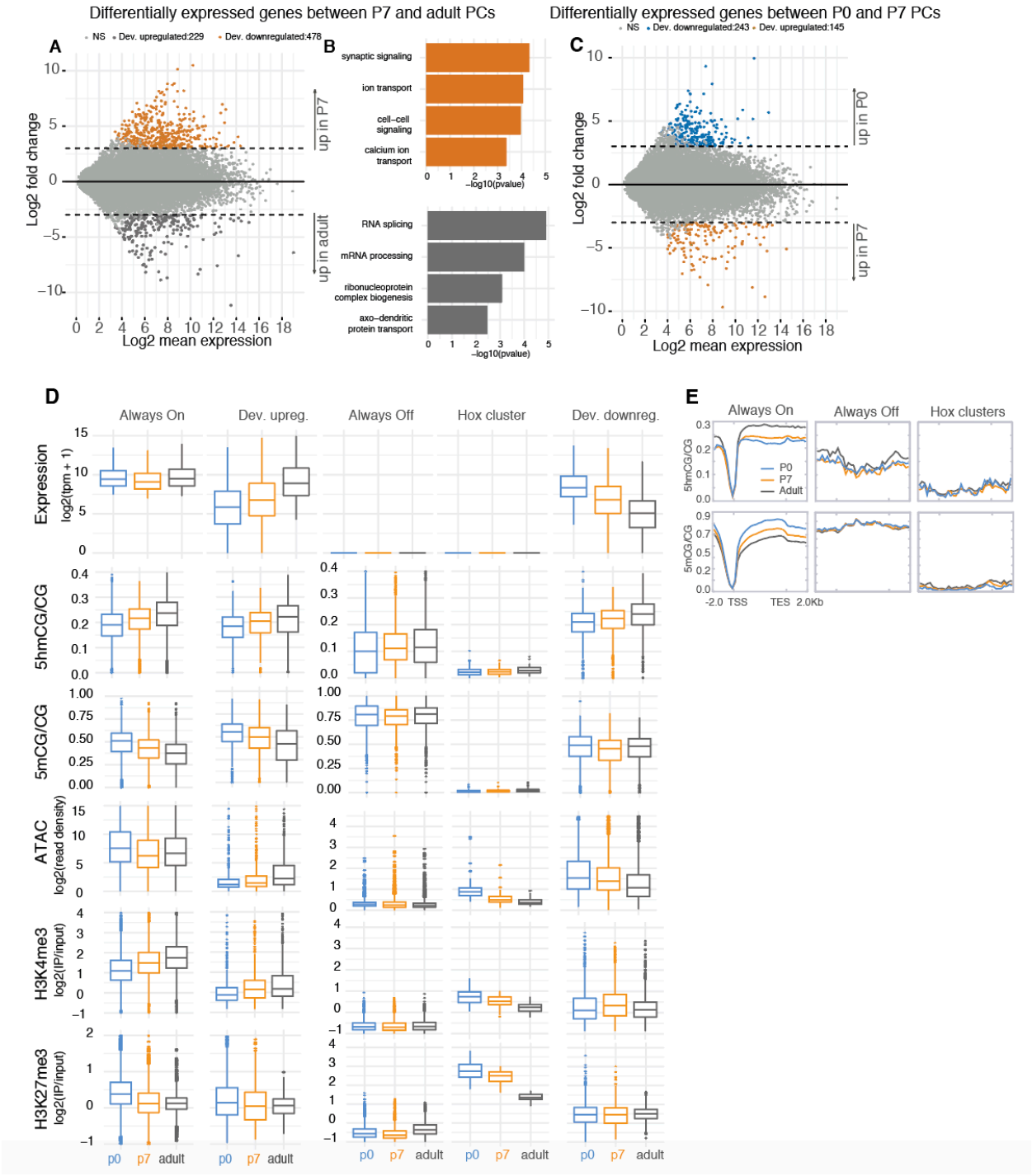


Figure 2.5. Differential expression between P0 and Adult Purkinje cells.
A. MA plot representing statistically significant ($p < 0.01$, log2(fold change) > 2) differential gene expression between P0 (blue) and adult (dark gray). Light grey dots represent genes that are not statistically significant. **B**. Gene ontology (GO) analysis of differentially expressed genes between P0 and Adult. **C**-**D**. Metagene plots representing the mean value of 5hmCG/CG (**C**) and 5mCG/CG (**D**) over the gene bodies of developmentally up- and down-regulated genes over the three timepoints (P0 – blue, P7 – orange and adult – dark grey).

Figure 2.6. Gene expression dynamics during Purkinje cell development.
A. MA plot representing differential gene expression between P7 and adult. Orange shaded dots represent genes enriched in P7 and gray dots represent genes enriched in adult, with absolute log2 fold change of 2 and p-adj value of 0.05. **B**. GO analysis of differentially expressed genes between P7 and adult. **C**. MA plot representing differential gene expression between P0 and P7 timepoints. Blue shaded dots represent genes enriched in P0 and orange dots represent genes enriched in P7, with absolute log2 fold change of 2 and p-adj value of 0.05. **D**. Quantification of expression, chromatin accessibility, 5mCG, 5hmCG, H3K4me3 and H3K27me3. **E**. Metagene representation of mean 5hmCG and 5mCG in control gene lists.


Differentially expressed genes between P7 and adult PCs
A
NS
Dev. upregulated:229
Dev. downregulated:478
Log2 fold change
Log2 mean expression
up in P7
up in adult
B
synaptic signaling
ion transport
cell–cell signaling
calcium ion transport
RNA splicing
mRNA processing
ribonucleoprotein complex biogenesis
axo–dendritic protein transport
−log10(pvalue)
Differentially expressed genes between P0 and P7 PCs
C
NS
Dev. downregulated:243
Dev. upregulated:145
Log2 fold change
Log2 mean expression
up in P0
up in P7
D
Always On
Dev. upreg.
Always Off
Hox cluster
Dev. downreg.
Expression
log2(tpm + 1)
5hmCG/CG
5mCG/CG
ATAC
log2(read density)
H3K4me3
log2(IP/input)
H3K27me3
log2(IP/input)
p0
p7
adult
E
Always On
Always Off
Hox clusters
5hmCG/CG
P0
P7
Adult
-2.0 TSS
TES 2.0Kb

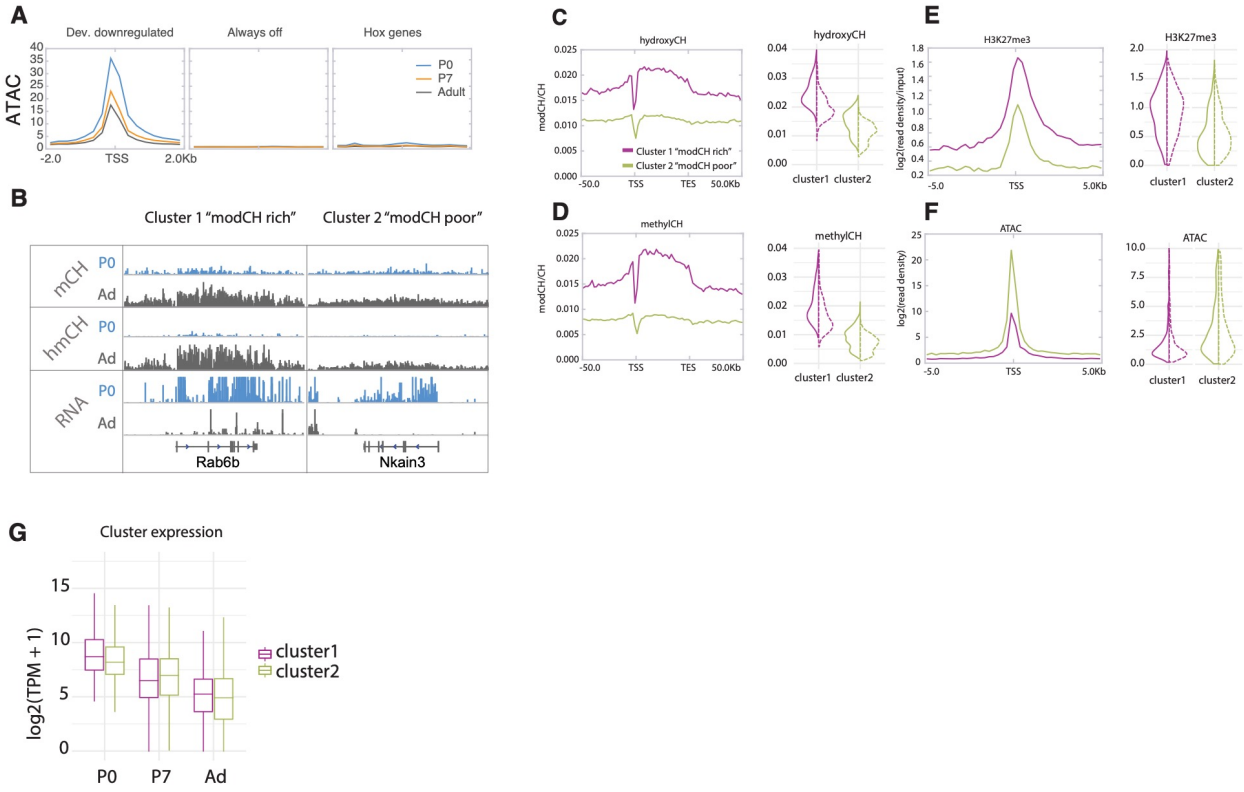


Figure 2.7. Two patterns of modified cytosines in repressed genes
A. Metagene plot representing ATAC-Seq read signal over the promoter regions of developmentally downregulated, Hox genes and always off genes (olfactory and vomeronasal receptors). **B**. IGV representation of repressed genes that accumulate (Rab6b) or not (Nkain3) modified cytosines in CH context. **C**-**D**. K-means clustering analysis of 5mCH and 5hmCH modifications of developmentally downregulated genes. Metagenes represent the average signal over the gene bodies. Violin plots summarize the modification value per gene. **E**-**F**. H3K27me3 (**E**) and chromatin accessibility (**F**) in the promoters of the genes in the two clusters. Metagenes represent the average signal over the gene transcription start sites. Violin plots summarize the modification value per gene. **G**. Expression levels log2(TPM+1) of the genes in each cluster.

2.4 Putative dynamic regulatory regions in Purkinje cells

The accessible chromatin landscape of Purkinje neuron is as dynamic during differentiation as their transcriptome. We used a computational method to identify all the regions in the genome at each time point that have a statistically significant enrichment of ATAC-seq signal. Then, we performed a differential accessibility analysis, which revealed multiple regions that are changing between P0 and adult (**Figure 2.8A**). We divided those regions into two groups based on how significant their change is ($p<0.01$) and its magnitude (log2 fold change > 4) – regions that either "gained" (more accessible in adult compared to P0) or "lost" accessibility (less accessible in adult compared to P0) and explored the cytosine modification dynamics in them. Interestingly, when a region becomes inaccessible, this is not accompanied by a gain of methylation or hydroxymethylation (**Figure 2.8B-C, Loss section**). Therefore, the loss of accessibility is not mediated through repressive DNA methylation. In regions that gain accessibility, both 5hmCG and 5mCG are lost (**Figure 2.8B-C, Gain section**). This is quite exciting, as it is suggesting that active demethylation is occurring in postmitotic cells in order to remodel regulatory regions. This idea is further explored in the next chapter of the thesis.

Since we had two groups of regions that are developmentally important, we decided to explore that transcription factor motifs we can identify. We saw multiple motifs that are differentially detected, which also display differential expression levels (**Figure 2.8D-E**). These findings suggest that there is a complex interplay of transcription factors during differentiation and their activity is in part regulated by allowing or restricting access to their binding sites.

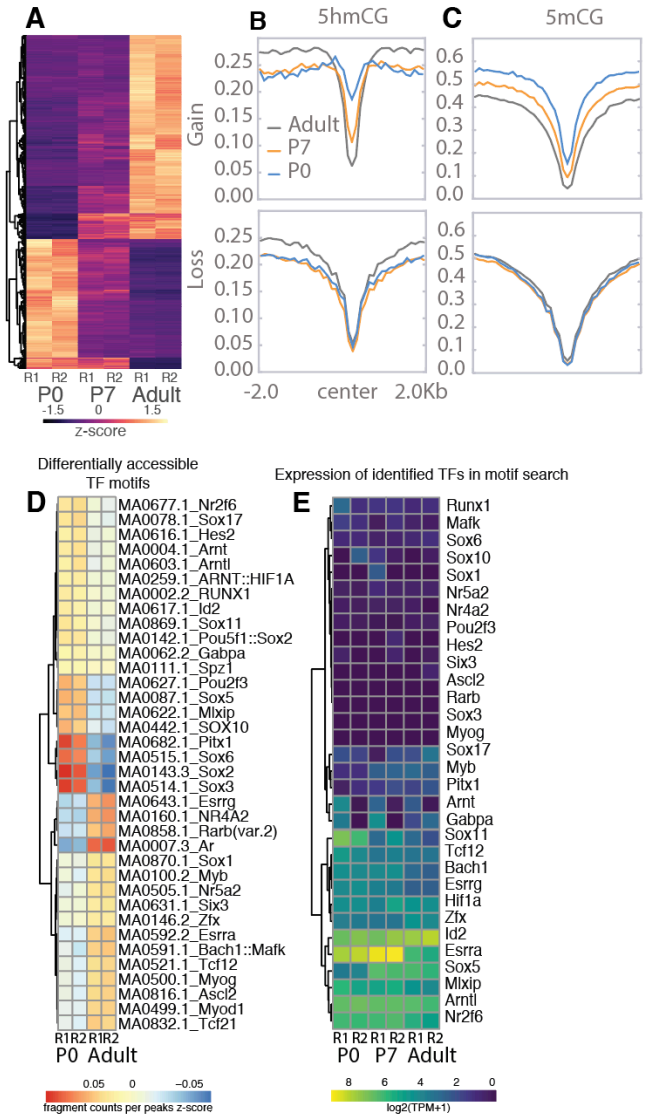


Figure 2.8. Open chromatin dynamics during Purkinje neuron differentiation.
A. Heatmap representing the differentially accessible regions between P0 and adult PCs ($p < 0.01$, log2(fold change) > 4). **B-C**. Metagene plots representing the mean values of 5hmCG/CG (**B**) and 5mCG/CG (**C**) over the centers and flanking regions of peaks that gained or lost accessibility relative to adult PCs. **D**. Differentially present motifs found in P0 and adult Purkinje cells from ATAC-Seq data, identified using ChromVar. **E**. Expression levels of identified motifs in log2(TPM+1).

2.5 A novel class of epigenetically regulated highly expressed Purkinje specific genes is actively demethylated

Although the majority of upregulated genes we explored follow the same pattern of increased 5hmCG accumulation with increased expression, our data reveal an additional, surprising feature of gene activation that has not been documented in postmitotic cells. Thus, in a small subset of highly expressed genes (Cep76, Itpr1, Mtss1) we noticed a profound loss of both 5hmCG and 5mCG (**Figure 2.9A**). To investigate the apparent demethylation over this class of genes, we computationally identified DNA methylation valleys (DMVs) in the following manner (**Figure 2.10A**). We used a published tool to identify undermethylated regions, which are defined as any region with under 15% methylation (Burger et al. 2013). Then, we selected every one that is over 5kb, and merged regions within 1kb distance, as we assumed, they would be part of the same regulatory region (**Figure 2.10B**). Those comprised our list of DNA methylation valleys (DMVs) (**Table 2.1**). We followed the analysis by dividing the DMVs into accessible or inaccessible based on their level of ATAC-Seq read enrichment (**Figure 2.10C**). Once we plotted the length of those regions, we noticed a great number is longer than 15kb, which was very interesting as the average length of active DMVs is about the size of a promoter (~5kb) (**Figure 2.10D**) (Jeong et al. 2014). Broad DMVs (>15kb) were present over both active and inactive genes, including the previously defined class of early developmental transcription factors that are inactive, hypo-methylated, inaccessible, and have elevated H3K27me3 levels (**Figure 2.9A**, Foxd1). Interestingly, these regions overlapped with broad peaks of chromatin accessibility and histone marks, close to 100kb, which is novel in itself (**Figure 2.10E).**

The evolution of DMVs in differentiating PCs is complex, and it reflects the transcriptional status of individual genes. Itpr1 and Mtss1 transcription steadily increases from P7 to adult, and it is accompanied by decreases in the level of 5mGC levels over their gene bodies. Both genes have high levels of 5hmCG at P0 which steadily decrease during differentiation. These findings were corroborated by differential methylation region (DMR) analysis as well. We performed DMR comparisons between P0 and Adult, and P7 and Adult. In both groups we identified multiple hypo-DMRs (hypo-methylated in Adult), overlapping with the genes that have DMVs. As methylation and hydroxymethylation are lost and DMVs broaden (**Figure 2.10I-K),** chromatin accessibility and H3K4me3 accumulation increase correspondingly (**Figure 2.9B-E**). These genes overlapping with broad DMVs also do not accumulate any 5mCH or 5hmCH (**Figure 2.10G)**.

Interestingly, many of the genes associated with active broad DMVs are Purkinje neuron specific and highly expressed (**Figure 2.10F)**. Gene ontology analysis indicates that they are involved in the inositol triphosphate/calcium signaling pathways, and a subset is associated with ataxia and autism (**Figure 2.10H).** These data provide strong evidence that DNA demethylation can occur in postmitotic neurons, and that it is enhanced in a specific class of genes that are extremely highly expressed and functionally important. Although mass spectroscopic analysis of genomic DNA from differentiating PCs (**Figure 2.10L**) failed to detect 5-formylcytosine or 5-carboxycytosine, their involvement as transient intermediates in the loss of 5mC and 5hmC cannot be ruled out because the small fraction of the genome covered by this gene class may preclude their detection.

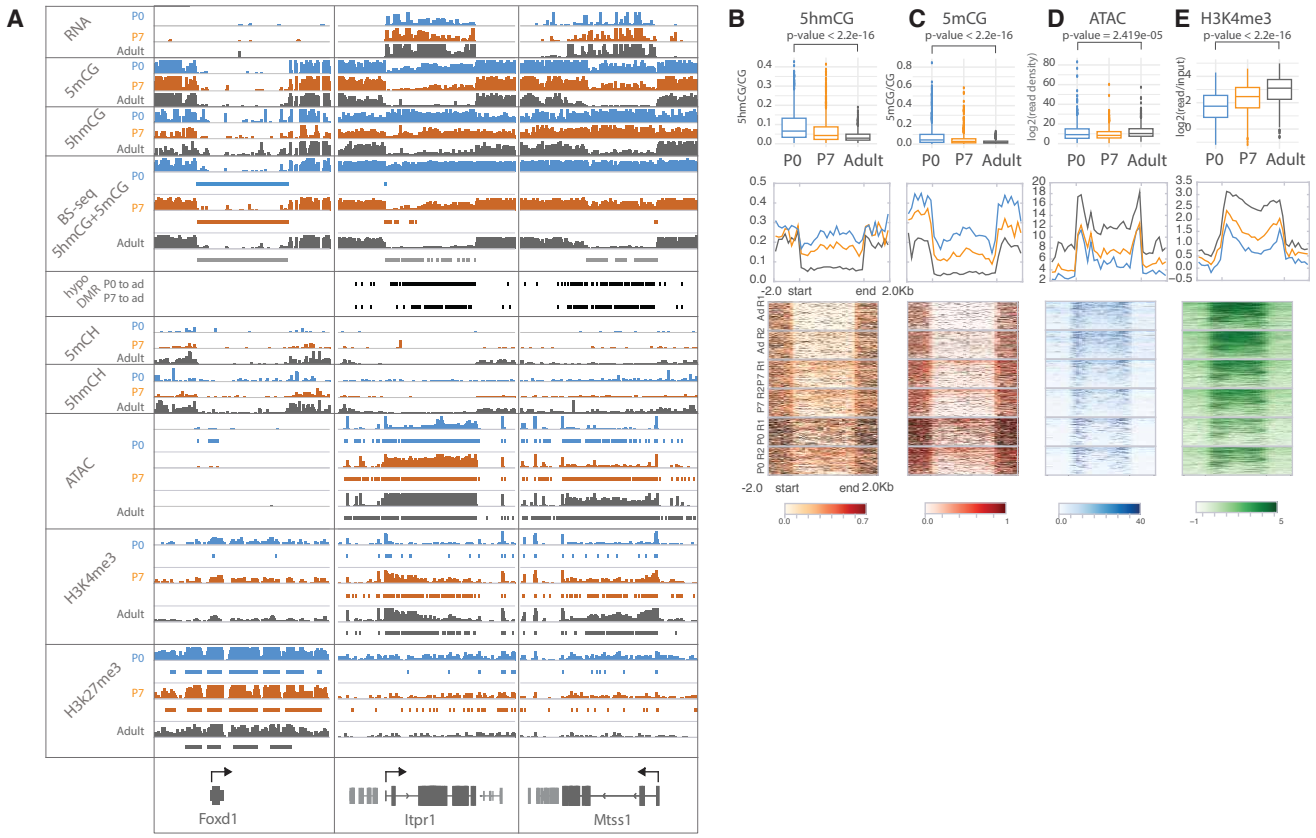


Figure 2.9. Novel class of epigenetically regulated highly expressed Purkinje specific genes. **A**. IGV representation of example regions of inactive DMV (Foxd1) and active PC specific DMVs (Itpr1, Mtss1). BS-Seq tracks shows combined levels of mCG and hmCG ranging from 0 to 1. Bars under BS-Seq tracks are computationally identified DMVs, black bars denote hypo-DMRs in adult, compared to P0 or P7. Bars under ATAC-Seq, H3K4me3 and H3K27me3 tracks denote broad peaks of signal enrichment. **B**-**E**. 5hmCG/CG (**B**), 5mCG/CG (**C**), ATAC-Seq log2(read density) RPKM enrichment (**D**) and H3K4me3 log2(input normalized) RPKM enrichment (**E**) quantification over DMVs at each timepoint. Boxplots show mean value per DMVs, the test for significance is Wilcoxon. Metagene plots show mean value ± 2kb around the DMV region, regardless of gene directionality. Heatmaps show data summarized in the metagene plots.

Figure 2.10. Characterization of Purkinje-specific DNA methylation valleys.
A. Schematic of DMV identification and filtering. **B**. Distribution of distances between DMVs. DMVs within 1kb of each other are assumed to be part of the same regulatory region and therefore merged. **C**. ATAC-Seq read density over DMVs shows a bimodal distribution, segregating DMVs into accessible and inaccessible (left panel). Inaccessible DMVs show enriched for H3K27me3 reads, while accessible DMVs show enrichment for H3K4me3 reads (right panel, dotted and straight lines designate replicates). **D**. Distribution of lengths of accessible and inaccessible DMVs. **E**. Length of broad peaks of ATAC-Seq and H3K4me3 ChIP-Seq overlapping large accessible DMVs. **F**. Expression of the top 500 most expressed genes in Purkinje cells and the genes associated with DMVs. **G**. DMVs do not accumulate any modifications in CH context. **H**. Genes associated with DMVs, red star denotes role in autism, gray in ataxia **I**. Cumulative length of DMVs per gene. **J-K**. Change of DMV length between P0 and adult (**J**) and P7 and adult (**K**). **L**. Mass spectrometry analysis of genomic DNA from adult PCs.

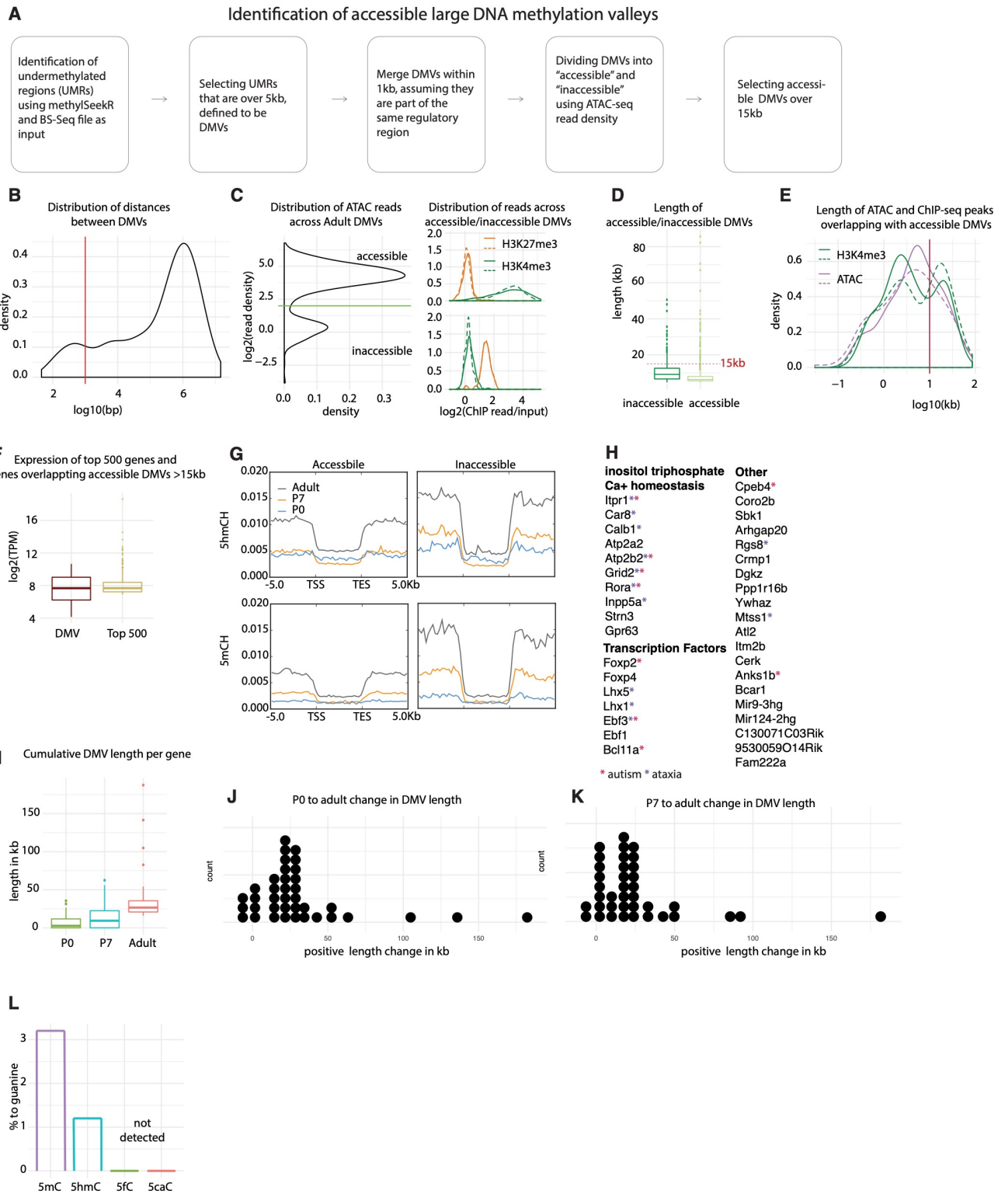

A
Identification of accessible large DNA methylation valleys
Identification of undermethylated regions (UMRs) using methylSeekR and BS-Seq file as input
Selecting UMRs that are over 5kb, defined to be DMVs
Merge DMVs within 1kb, assuming they are part of the same regulatory region
Dividing DMVs into "accessible" and "inaccessible" using ATAC-seq read density
Selecting accessible DMVs over 15kb
B
Distribution of distances between DMVs
density
log10(bp)
C
Distribution of ATAC reads across Adult DMVs
accessible
inaccessible
log2(read density)
density
Distribution of reads across accessible/inaccessible DMVs
H3K27me3
H3K4me3
log2(ChIP read/input)
D
Length of accessible/inaccessible DMVs
length (kb)
15kb
inaccessible accessible
E
Length of ATAC and ChIP-seq peaks overlapping with accessible DMVs
H3K4me3
ATAC
density
log10(kb)
F
Expression of top 500 genes and genes overlappting accessible DMVs >15kb
log2(TPM)
DMV
Top 500
G
Accessbile
Inaccessible
Adult
P7
P0
5hmCH
5mCH
-5.0 TSS TES 5.0Kb
H
inositol triphosphate Ca+ homeostasis
Itpr1**
Car8*
Calb1*
Atp2a2
Atp2b2**
Grid2**
Rora**
Inpp5a*
Strn3
Gpr63
Transcription Factors
Foxp2*
Foxp4
Lhx5*
Lhx1*
Ebf3**
Ebf1
Bcl11a*
Other
Cpeb4*
Coro2b
Sbk1
Arhgap20
Rgs8*
Crmp1
Dgkz
Ppp1r16b
Ywhaz
Mtss1*
Atl2
Itm2b
Cerk
Anks1b*
Bcar1
Mir9-3hg
Mir124-2hg
C130071C03Rik
9530059O14Rik
Fam222a
* autism * ataxia
I
Cumulative DMV length per gene
length in kb
P0 P7 Adult
J
P0 to adult change in DMV length
count
positive length change in kb
K
P7 to adult change in DMV length
positive length change in kb
L
% to guanine
not detected
5mC 5hmC 5fC 5caC

Table 2.1. Accessible DMVs with length over 15kb.

Chr	Start	End	Gene Name	Gene Description
chr6	108272646	108358260	Itpr1	inositol 1,4,5-trisphosphate receptor 1
chr6	108209931	108265472	Itpr1	inositol 1,4,5-trisphosphate receptor 1
chr4	8169631	8240202	Car8	carbonic anhydrase 8
chr11	100669657	100695452	Dhx58	DEXH (Asp-Glu-X-His) box polypeptide 58
chr5	120423843	120441118	Lhx5	LIM homeobox protein 5
chr17	79872208	79896806	Atl2	atlastin GTPase 2
chr10	90090835	90109116	Anks1b	ankyrin repeat and sterile alpha motif domain containing 1B
chr1	153652715	153672426	Rgs8	regulator of G-protein signaling 8
chr2	91942312	91967185	Dgkz	diacylglycerol kinase zeta
chr6	63300485	63382702	Grid2	glutamate receptor, ionotropic, delta 2
chr17	47896993	47925907	Foxp4	forkhead box P4
chr9	69288847	69325026	Rora	RAR-related orphan receptor alpha
chr5	122466487	122504054	Atp2a2	ATPase, Ca++ transporting, cardiac muscle, slow twitch 2
chr5	114566729	114599725	Fam222a	family with sequence similarity 222, member A
chr6	15182839	15201300	Foxp2	forkhead box P2
chr7	139427072	139494058	Inpp5a	inositol polyphosphate-5-phosphatase A
chr3	17783955	17802624	Mir124a-2	microRNA 124a-2
chr15	59046367	59083522	Mtss1	metastasis suppressor 1
chr5	37240217	37265126	Crmp1	collapsin response mediator protein 1
chr14	73368702	73385694	Itm2b	integral membrane protein 2B
chr6	63468817	63494400	Grid2	glutamate receptor, ionotropic, delta 2
chr6	113870078	113892648	Atp2b2	ATPase, Ca++ transporting, plasma membrane 2
chr15	36775123	36795083	Ywhaz	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein, zeta polypeptide
chr9	51763097	51781640	Arhgap20	Rho GTPase activating protein 20
chr4	15880335	15906914	Calb1	calbindin 1
chr15	59016633	59033746	Mtss1	metastasis suppressor 1
chr4	24965815	25008500	Gpr63	G protein-coupled receptor 63
chr6	63254687	63299184	Grid2	glutamate receptor, ionotropic, delta 2
chr11	84509915	84536544	Lhx1	LIM homeobox protein 1
chr7	139386804	139424726	Inpp5a	inositol polyphosphate-5-phosphatase A
chr7	79497537	79527384	Mir9-3hg	Mir9-3 host gene
chr11	24073611	24099200	Bcl11a	B cell CLL/lymphoma 11A (zinc finger protein)
chr13	83718383	83745844	C130071C03Rik	RIKEN cDNA C130071C03 gene
chr8	111720813	111746280	Bcar1	breast cancer anti-estrogen resistance 1

chr15	58971219	59004802	Ndufb9	NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9
chr15	86168217	86187160	Cerk	ceramide kinase
chr7	137295733	137316336	Ebf3	early B cell factor 3
chr12	51658821	51693562	Ap4s1	adaptor-related protein complex AP-4, sigma 1
chr11	31869833	31903898	Cpeb4	cytoplasmic polyadenylation element binding protein 4
chr6	108425981	108465693	Mir7661	microRNA 7661
chr9	122568186	122586218	9530059O14Rik	RIKEN cDNA 9530059O14 gene
chr7	126271173	126292655	Sbk1	SH3-binding kinase 1
chr9	62519411	62538329	Coro2b	coronin, actin binding protein, 2B
chr2	158666781	158682774	Ppp1r16b	protein phosphatase 1, regulatory (inhibitor) subunit 16B
chr11	44616001	44635818	Ebf1	early B cell factor 1

2.6 Discussion

We embarked on the difficult path to purify an extremely rare neuron at three developmental timepoints. We expanded the FANS-seq purification method from adult Purkinje nuclei to P0 and P7, and confirmed that it is suitable for downstream applications such as OxBS-Seq and ChIP-Seq. This allowed us to create an in-depth characterization of the chromatin landscape of developing Purkinje cells, including transcriptome profiling, methylation and hydroxymethylation of cytosine, accessible chromatin and certain histone tail modifications. These data allowed us to curate the expression profile into several classes of genes – constitutively expressed, never expressed, developmentally up- and down-regulated genes. Comparative analysis of these classes of genes spelled out several important rules for the regulation of gene expression.

Developmentally downregulated genes do not lose the previous accumulated 5-hydroxymethylcytosine over their gene bodies. This suggests that 5-hydroxymethylcytosine alone is not sufficient to promote nor repress transcription. Furthermore, we also do no observe methylation of the promoter regions of those genes, even though they have lost accessibility, which suggests that repression is occurring through other mechanisms. When exploring their patterns of methylation and hydroxymethylation in CpH context, developmentally downregulated genes fall into two groups – ones that accumulated modified CpH (modCHrich) and ones that do not (modCHpoor). ModCHrich genes are associated with even lower promoter accessibility and higher accumulation of H3K27me3 than the modCHpoor ones. This suggest that the accumulation of modified CpH nucleotides could lead to the binding of repressive proteins and complexes such and MeCP2 and Polycomb that mediate repression.

We identified multiple regions of the genome that are differentially accessible at different timepoints. These putative regulatory regions could be binding sites for activators or repressors, enhancers, alternative promoters or other unknown regulatory elements. Similar to silenced genes, sites that lose accessibility do not gain methylation, suggesting alternative mechanism of repression. Further studies focusing on the non-coding genome for regulation and refining the regions through multiple histone marks could shed more light on the precise control of the Purkinje cell fate. This could lead to the improvement of protocols for Purkinje cell differentiation in culture for more precises drug target selection and validation.

Developmentally upregulated genes follow the already established pattern of accumulating 5-hydroxymethylcytosine with active transcription. This was confirmed at all three developmental timepoints and begs the question whether the TET proteins are recruited to their targets through the transcriptional machinery. Surprising, a small subset of the most highly expressed genes does not accumulate high levels of 5-hydroxymethylcytosine, and in fact has very low levels of modified cytosines. This is not particularly novel, as transcription factors are known to lack any cytosine modification, but they are quite small (usually up to 10kb). This novel class of epigenetically regulated genes is quite large – one demethylated region is close to 200kb. We show that these genes undergo active demethylation between P0 and Adult. This is accompanied by broad regions of accessibility and H3K4me3 enrichment which increase from P0 to Adult. Why are these genes demethylated and extremely accessible? One possible explanation is that since these genes are of utter importance to the proper functioning of the Purkinje cells, they need to be expressed at very high and stable amounts. This accomplished by completely unraveling the gene body and binding as many transcriptional units as possible.

While we gathered an incredible amount of useful information about Purkinje cell differentiation from the three chosen timepoints, future studies would benefit from going even earlier in development to the birth of Purkinje cells. This would allow for the characterization of the chromatin landscape right after the last division and build an even better differentiation trajectory and the associated chromatin changes. It is interesting that we have observed DNA demethylation in cases when accessibility or transcription is required, but we do not see it in genes that have been silenced in order to establish repressive methylation. This could have implications long term, so further studies of older mice would be of great interest to the dynamics of 5hmC and whether it has any role in chromatin erosion and aberrant expression in aging.

CHAPTER 3. Purkinje neuron specific depletion of Tet1, Tet2 and Tet3.

3.1 Introduction

TET loss-of-function studies in mice have demonstrated that 5hmC is essential for normal development and maintenance of lymphocyte lineages, and that TET proteins are required to provide full accessibility of regulatory regions necessary for modulation of developmentally important genes. The only study in the cerebellum, in cultured granule neurons showed that a Tet1/Tet3 knock down impairs the expression of axon guidance and ion channel genes, affecting dendritic arborization and circuit formation (X. Zhu et al. 2016). Besides this, the knockout studies in the central nervous system are few and typically of only one gene at a time. Since the TET proteins are reported to have overlapping activity, only a conditional triple TET knockout can reveal the specific function of 5hmC in neurons. To continue our characterization of the function and dynamics of 5hmC we created a new conditional triple TET knockout in Purkinje cells and evaluated the phenotype in adult mice.

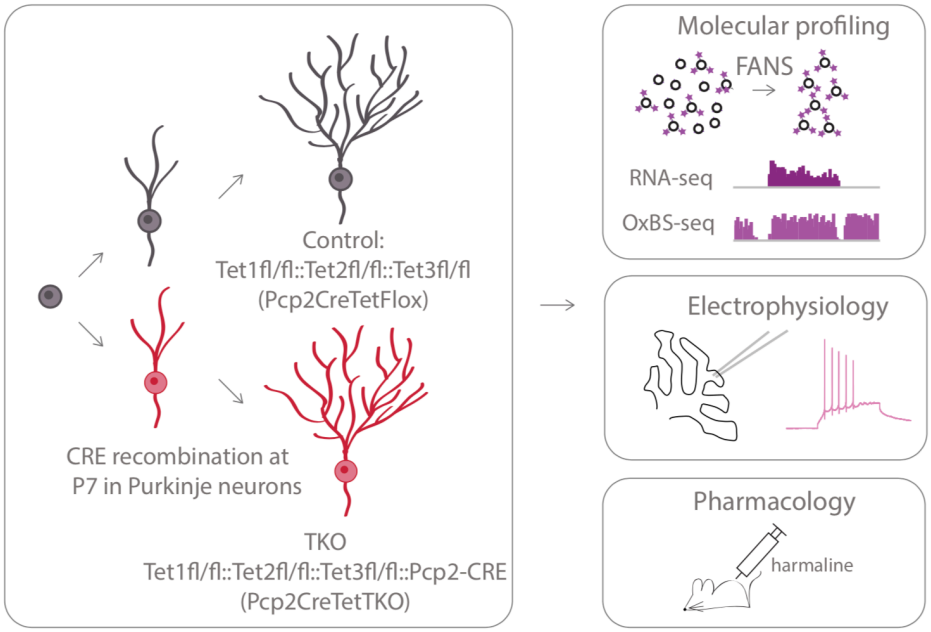


Figure 3.1. Probing the function of the TET proteins *in vivo*.
Schematic of experimental design to probe the triple Tet1,2,3 knock-out effect in adult PCs, using molecular biology, electrophysiology and pharmacology tools.

3.2 Purkinje neuron specific triple TET knockout mouse strain generation

To directly address the role of 5hmC in the transcriptional and epigenetic events that accompany Purkinje cell differentiation, we next generated PC specific Tet1/Tet2/Tet3 triple knockout (Pcp2CreTetTKO) mouse lines. We selected a Purkinje specific CRE driver construct (Pcp2-CRE), which was previously used by the Heintz lab to generate accurate mouse strains and its expression pattern was characterized in depth (Gong et al. 2007). Its expression is restricted to Purkinje cells and begins at P7 (**Figure 3.2A-B**). We received triple floxed Tet1, Tet2 and Tet3 mice as a generous gift from Anjana Rao. The engineered construct was introduced by pronuclear injection into ova from females carrying floxed alleles of all three Tet genes carried out at the Rockefeller University Transgenic and Reproductive Technology Center (**Figure 3.2C**). Out of 26 pups born, only 6 contained the Pcp2-CRE gene, and only 3 of them survived long enough to establish stable lines. One of them (Founder 10) had off target expression, causing aberrant deletion of the Tet proteins. Mice from that line consistently died at 8 weeks, and after postmortem examination it was discovered that they had leukemia with lymphoma. Because of these circumstances, this founder line was excluded from further analysis. This left us with two founder lines – 19 and 13 – to continue the investigation. PCR analysis of the floxed regions of each Tet gene in purified PC genomic DNA, and RNA-seq analysis confirmed the deletion of exons from all three TET proteins (**Figure 3.2D, Figure 3.5E**). To assess the expression pattern of the newly introduced CRE driver, we crossed the mice with ones carrying CRE-dependent GFP gene. We collected brains at P0, P7 and adult, cut thin sections and stained for GFP using immunofluorescence. We confirmed that the CRE recombination activity follows the endogenous protein expression, starting at P7 in a salt-and-pepper pattern, and developing robustly by adult (**Figure 3.3A-D**). Next, we investigated whether the lines had any obvious phenotype. We assessed their survival by tracking their life expectancy and found no significant difference between Pcp2CreTetTKO and CRE negative littermates for both founder lines (**Figure 3.4A**). We also found no significant difference between the body weights of the Pcp2CreTetTKO and CRE negative littermates (**Figure 3.4B**). Finally, since the cerebellum has a major role in movement, we tested the mice's locomotive ability. We found no significant different between Pcp2CreTetTKO and CRE negative littermates' performance on rotarod (**Figure 3.4C**). These lines, therefore, allow assessment of the consequences of loss of Tet activity during the rapid growth phase of postmitotic Purkinje neurons.

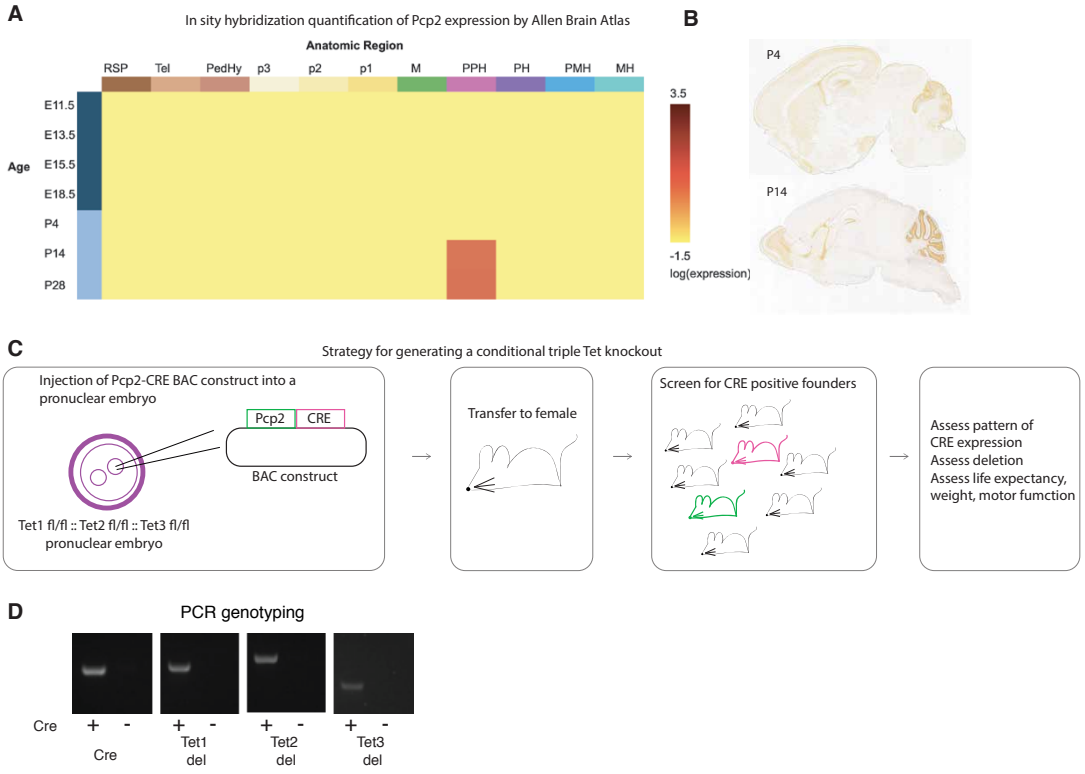


Figure 3.2. Generation of Pcp2 cre/wt :: Tet1 del/del :: Tet2 del/del :: Tet3 del/del (Pcp2CreTetTKO) mouse founder lines.
A-B. Data from Allen brain atlas showing Pcp2 expression among regions and ISH in P4 and P14 brains. **C**. Schematic of Pcp2CreTetTKO generation from pronuclear injection the BAC probe to screening the litters for the CRE gene. **D**. PCR genotyping for CRE, Tet1 deletion, Tet2 deletion and Tet3 deletion in adult Pcp2CreTetTKO Purkinje genomic DNA isolated post FACS purification. + refers to Cre positive, – refers to Cre negative.

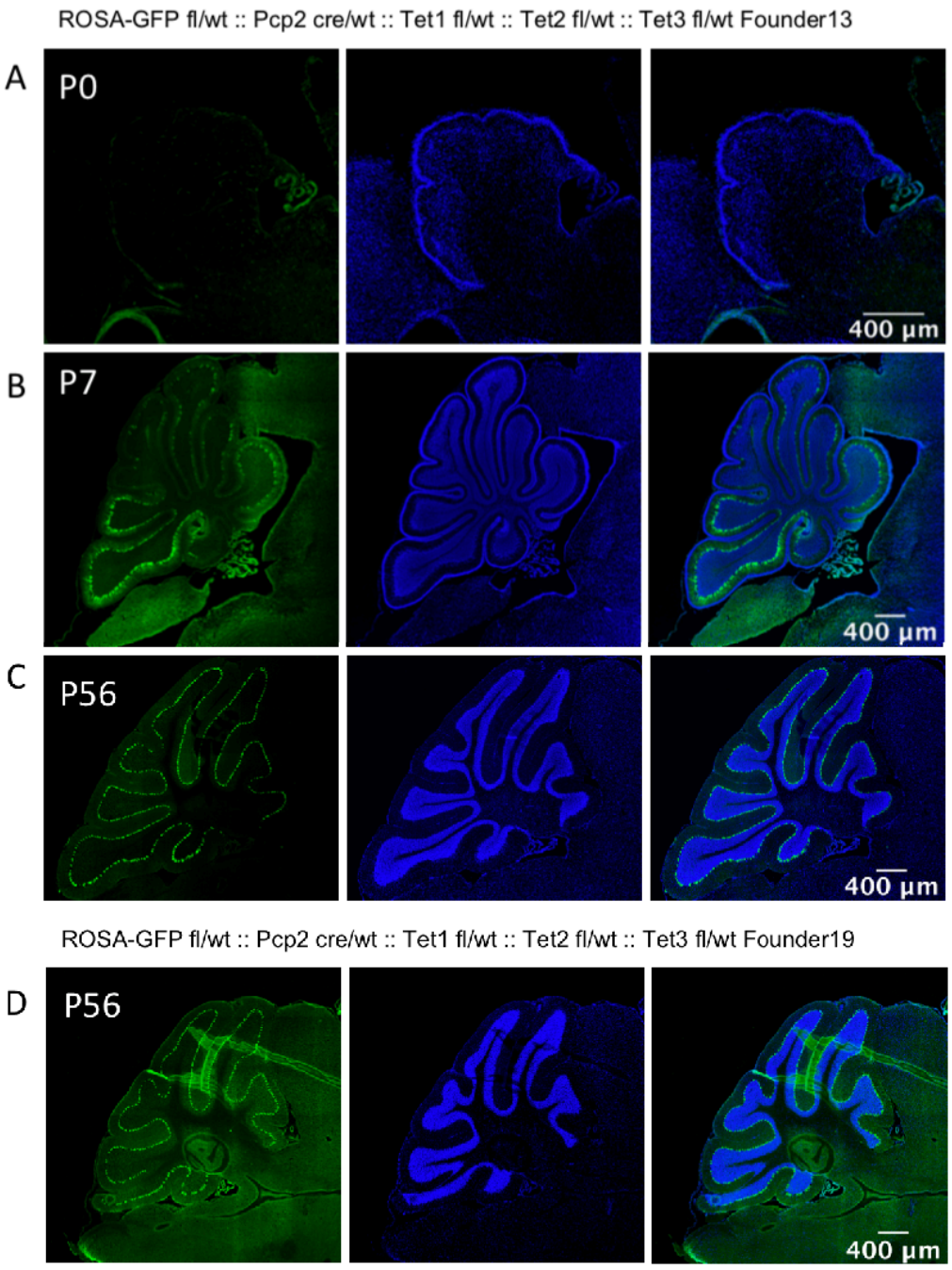


Figure 3.3. Characterization of Pcp2CreTetTKO mouse founder lines.
A-C. Pcp2 cre/wt :: Tet1 del/del :: Tet2 del/del :: Tet3 del/del mice were crossed to ROSA26-eGFP reporter line to evaluate CRE expression during development. Panels represent murine cerebella at P0 (**A**), P7 (**B**) and P56 (**C**) timepoints in FL13, stained with anti-GFP antibody. **D**. anti-GFP staining in P56 FL19.

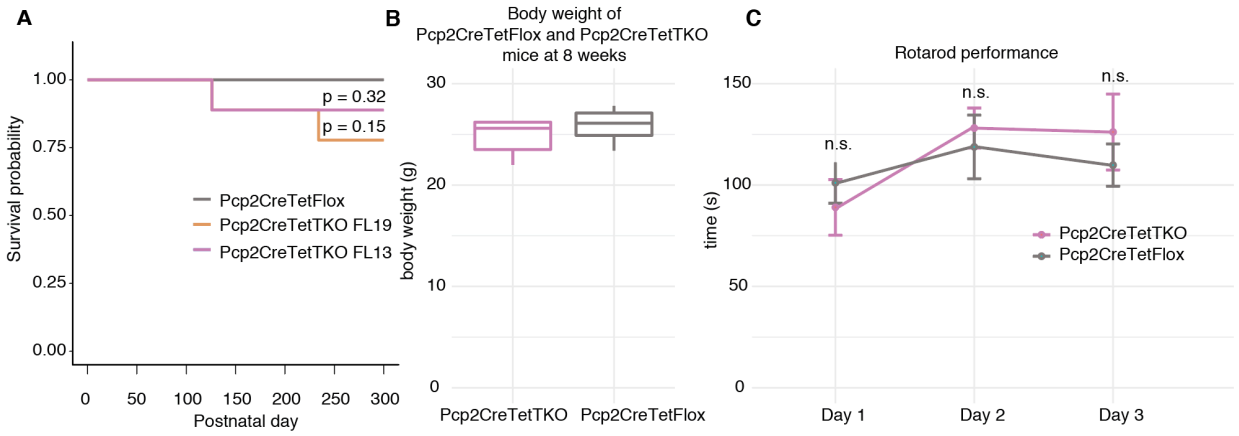


Figure 3.4. Gross phenotype evaluation of Pcp2CreTetTKO.
A. Kaplan-Meier curve representing the survival fraction of Pcp2CreTetFlox (n=8) and Pcp2CreTetTKO (Founder13 n=8, Founder19 n=8) mice since date of birth. **B**. Body weight of Pcp2CreTetFlox and Pcp2CreTetTKO at 8 weeks. **C**. Rotarod evaluation of motor skills in Pcp2CreTetFlox and Pcp2CreTetTKO at 8 weeks.

3.3 Loss of TET activity leads to altered gene expression regulation

We confirmed that the triple TET TKO did not have an effect on our Purkinje marker expression. Nuclei stained for Itpr1 and FACS plots look as expected compared to wild-type nuclei (**Figure 3.5A-B**). The Pearson correlation coefficient between replicates was over 0.98. The RNA-Seq datasets also showed the expected abundance of Purkinje specific genes, without any contamination from glia or cerebellar granule cells. To understand the consequences of loss of 5hmC in PCs, we compared the transcriptome of Pcp2CreTetTKO mice with floxed Cre negative littermates (Pcp2CreTetFlox) through a differential expression analysis. Surprisingly, we identified both genes with increased and decreased levels of expression, suggesting that potentially demethylation events are not solely responsible for activating gene expression. Furthermore, among those genes was a large number of ion channels and genes involved in synaptic transmission (**Figure 3.6A-B, Figure 3.7A**). Loss of function studies of Tet1, Tet2, Tet3 and combinations thereof have demonstrated that their essential functions in development of lymphocyte/hematopoietic lineage (C.-W. Lio et al. 2016; Tsagaratou et al. 2017; C.-W. J. Lio et al. 2019) and other cell types (Y. Xu et al. 2012, 1; T. Li et al. 2015; Rudenko et al. 2013, 1) reflects a direct role of TET proteins in transcriptional regulation as a consequence of replication dependent loss of 5mC in enhancers, promoters and gene bodies. Given the apparent demethylation of a subset of PC expressed genes during their postmitotic development, and the onset of Cre activity in PCs during this period, we were next interested in determining whether the changes in transcription in the Pcp2CreTetTKO are associated with altered methylation status of regulatory regions. Comparative analysis of P0 and adult ATAC-Seq data was first used to identify changes in chromatin accessibility that occur during PC differentiation, particularly becoming accessible in the Adult timepoint (**Figure 2.8A**). We then selected the regions which are associated with genes that are transcriptionally impacted in the Pcp2CreTetTKO analyzed. Accessible regions found only in adult PCs have significantly more 5mCG and 5hmCG in the knockout than they do in controls, similar to levels present in P7 PCs, suggesting that those regions have not reach the proper modification status required for correct levels of gene expression (**Figure 3.6C, Figure 3.7C**). Cytosine methylation and hydroxymethylation status over the promoters for these genes or their gene bodies is not altered (**Figure 3.7B**).

Genes that acquire broad DMVs late in development over the entire gene (Gpr63), or partially (Grid2), are no longer demethylated in the Pcp2CreTetTKO, whereas DMVs established early in development in transcriptionally silent (Hox cluster) or active (Calb1) genes are not affected (**Figure 3.6D-E**). These data demonstrate that ongoing oxidation of 5mCG to 5hmCG by Tet oxidases in PCs is required for execution of molecular and epigenetic events associated differentiation, and that demethylation of regulatory domains and gene bodies occurs in the absence of cell division.

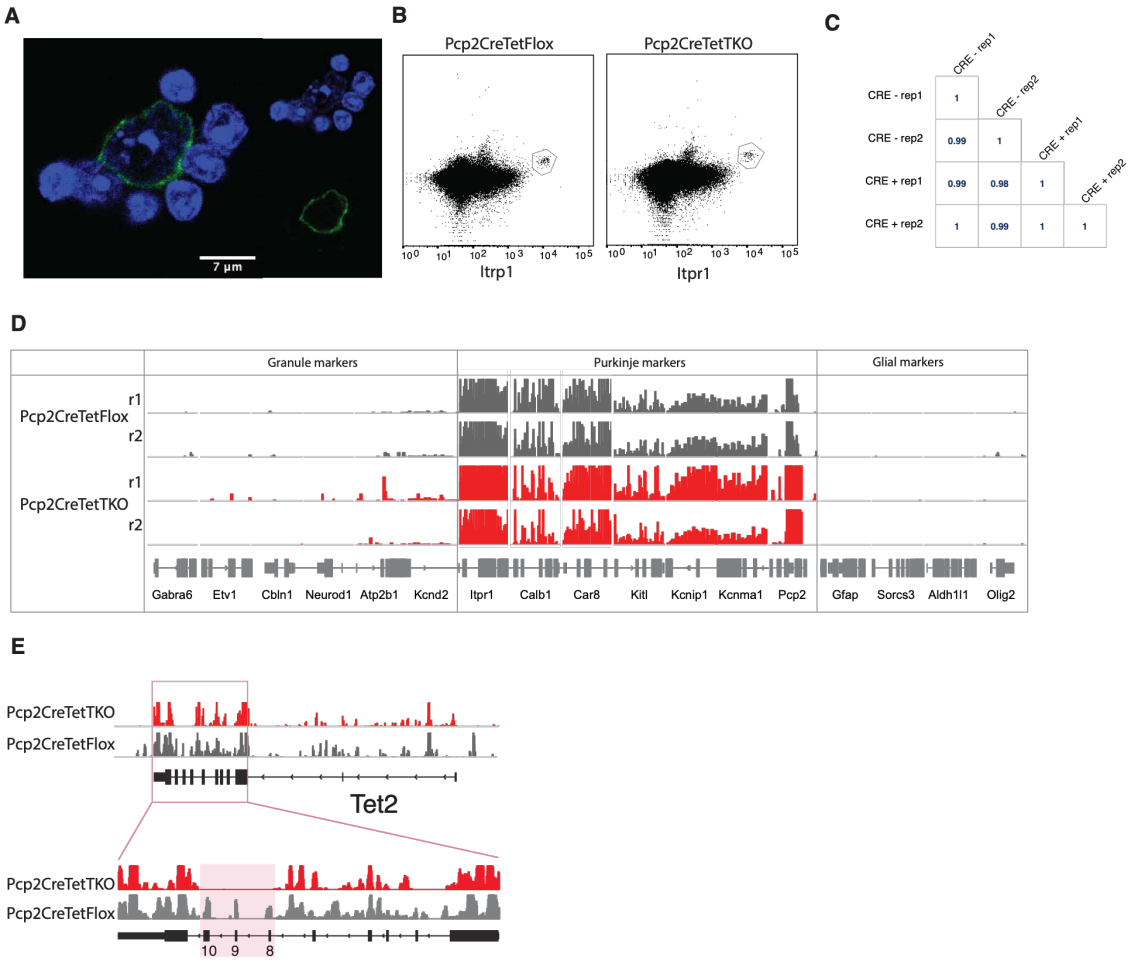


Figure 3.5. Quality control of TKO Purkinje cell isolation.
A. Example of Pcp2CreTetTKO nuclei stained with Itpr1 post dissociation and pre-sorting, counterstained with DAPI, a heterochromatin marker **B**. Example FANS plot of adult Pcp2CreTetTKO nuclei. **C**. Pearson correlation between TKO and control. **D**. IGV representation of Purkinje specific markers (Itpr1, Calb1, Car8, Kitl, Kcnip1, Kcnma1, Pcp2) enrichment and depletion of granule (Gabra6, Etv1, Cbln1, Neurod1, Atp2b1, Kcnd2) and glial (Gfap, Sorcs3, Aldh1l1, Olig2) markers in Pcp2CreTetTKO and Pcp2CreTetFlox. **E**. Deletion of Tet2 exons evident in nuclear RNA-Seq Pcp2CreTetTKO.

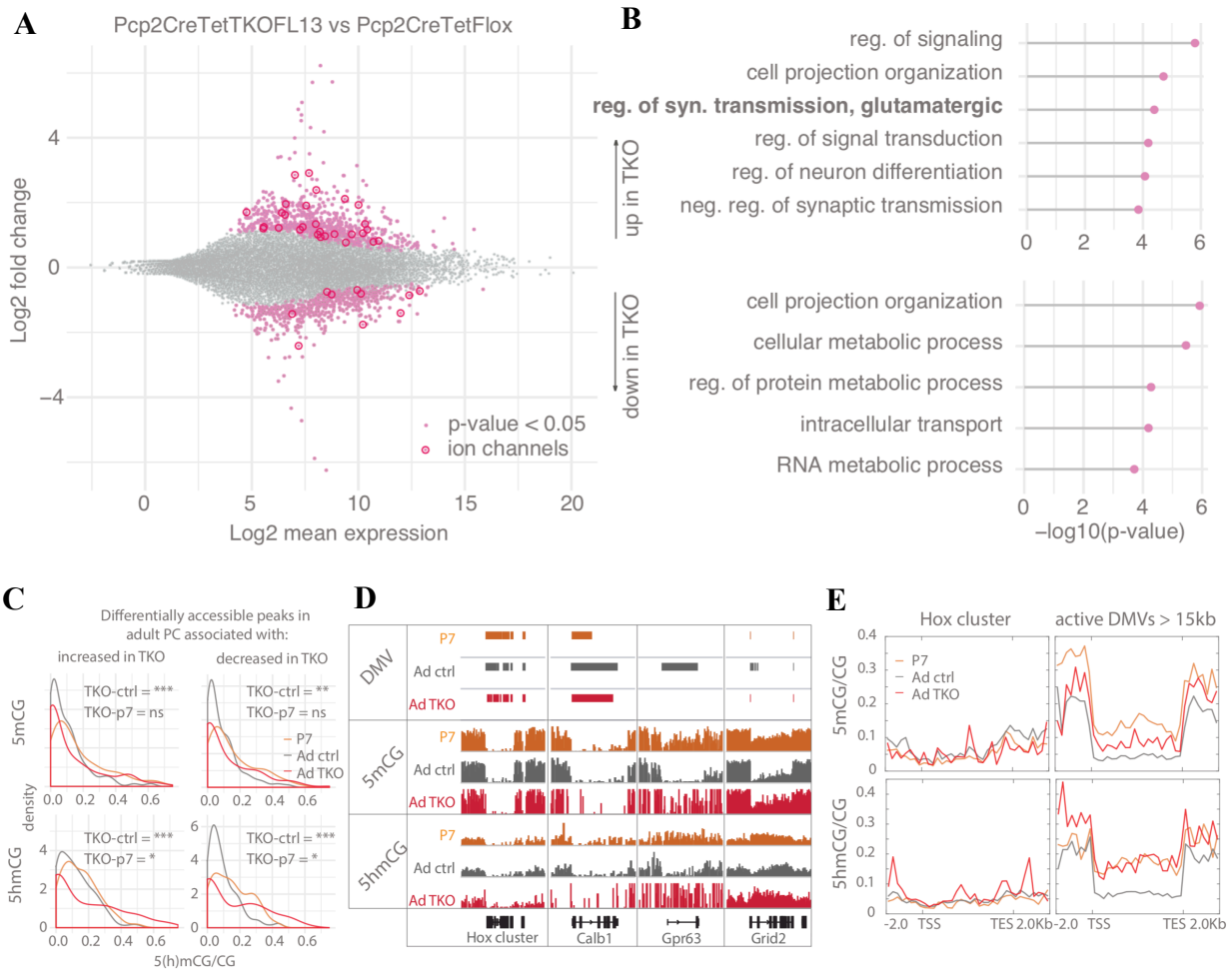
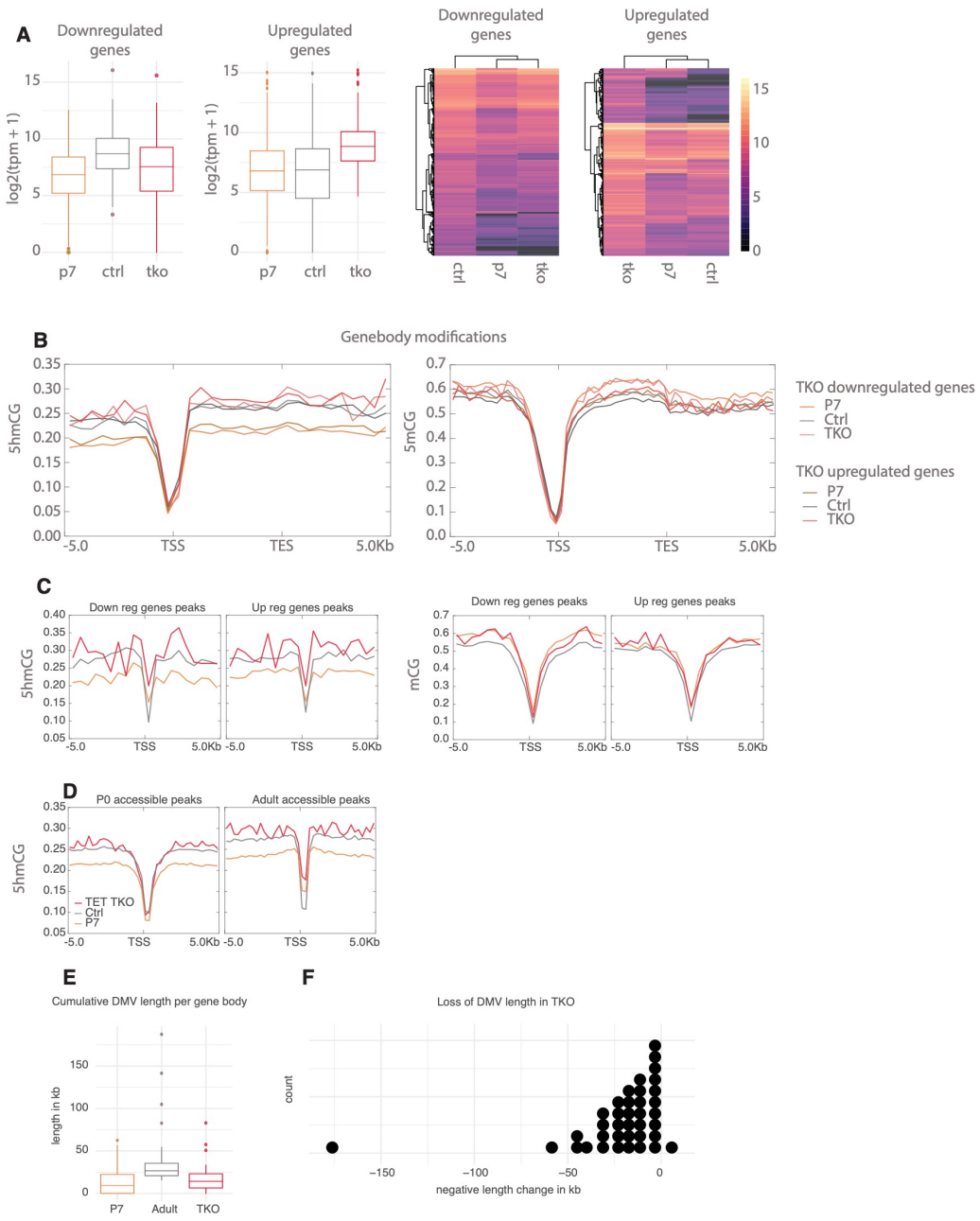


Figure 3.6. Loss of Tet activity leads to altered gene expression regulation in PCs.
A. MA plot representing differential gene expression analysis between Pcp2CreTetFlox and Pcp2CreTetTKO. Pink dots represent genes with p-value < 0.05, red dots - ion channels. **B**. Gene ontology analysis of statistically significant differentially expressed genes. Top panel shows categories of genes with increased expression in Pcp2CreTetTKO, bottom panel shows genes with decreased expression in Pcp2CreTetTKO. **C**. Density plots showing differential levels of 5hmCG/CG and 5mCG/CG over differentially accessible peaks only present in adult PCs. **D**. IGV representation of example DMVs affected by Pcp2CreTetTKO. Hox cluster and Calb1 show minor changes in length as they are established before the P7 onset of Cre. Gpr63 and Grid2 show significant reduction in length as they are established after P7. Solid bars represent DMVs identified at each condition. **E**. Quantification of 5hmCG/CG and 5mCG/CG over the two classes of DMVs – Hox cluster and large active DMVs.

Figure 3.7. Cytosine dynamics in Pcp2CreTetTKO.
A. Expression of genes with increased and decreased expression in Pcp2CreTetTKO (boxplots in left panel, heatmaps in right panel). **B**. Metagene plot of mean mCG/CG and hmCG/CG modifications over promoters and gene bodies of genes with increased or decreased expression in Pcp2CreTetTKO. **C**. Metagene plots of modifications in Pcp2CreTetTKO and Pcp2CreTetFlox accumulating over ATAC-Seq peaks found only in adult PCs associated with genes with increased or decreased expression. **D**. Metagene plots of modifications in Pcp2CreTetTKO and Pcp2CreTetFlox accumulating over ATAC-Seq peaks found only in adult PCs. **E**. Cumulative DMV length per gene body between Pcp2CreTetTKO and Pcp2CreTetFlox. **F**. Change in cumulative DMV length between adult Pcp2CreTetTKO and Pcp2CreTetFlox.


A
Downregulated genes
Upregulated genes
log2(tpm + 1)
p7
ctrl
tko
Downregulated genes
Upregulated genes
B
Genebody modifications
5hmCG
5mCG
-5.0
TSS
TES
5.0Kb
TKO downregulated genes
P7
Ctrl
TKO
TKO upregulated genes
P7
Ctrl
TKO
C
Down reg genes peaks
Up reg genes peaks
5hmCG
mCG
D
P0 accessible peaks
Adult accessible peaks
TET TKO
Ctrl
P7
E
Cumulative DMV length per gene body
length in kb
P7
Adult
TKO
F
Loss of DMV length in TKO
count
negative length change in kb

3.4 TET deficiency causes hyper-excitability and increased susceptibility to excitotoxic drugs

Given the large number of ion channel genes whose expression is altered in the Pcp2CreTetTKO, we were interested in determining whether these changes prevented PCs from attaining their mature electrophysiological properties. To address this issue, patch clamp recordings from Purkinje neurons in adult, acute slice preparations were performed in both knockout lines (FL13 and FL19) their respective heterozygous knockout (Pcp2CreTetTKO-het) controls, and Pcp2CreTetFlox control mice. We found no significant changes in resting membrane potential (**Figure 3.8C**) or input resistance (**Figure 3.8D**), indicating unaltered intrinsic electrophysiological and morphological properties. While we observed no changes in rheobase current (**Figure 3.8E**), a fixed current injection elicited dramatically higher rates of neuronal firing in Pcp2CreTetTKO PCs (**Figure 3.8A**). F-I curves demonstrated that homozygous knockouts have significantly higher PC firing rates than Pcp2CreTetTKO-het or Pcp2CreTetFlox (**Figure 3.8B**) across a series of escalating current injections. Taken together, these results indicate that Tet-mediated dysregulation of ion channels in Pcp2CreTetTKO results in elevated PC excitability.

The heightened excitability suggested that the Pcp2CreTetTKO mice might exhibit increased sensitivity to harmaline, a tremorgenic drug whose action elicits abnormal Purkinje cell activity. To assess this possibility, the sensitivity to harmaline was measured as the time required for tremor onset post injection. As predicted, a shorter onset interval was observed in the Pcp2CreTetTKO animals (**Figure 3.8F**). This demonstrates that the PC hyperexcitability can result in altered function of the cerebellar circuitry and motor behavior, at least in response to induced climbing fiber activation.

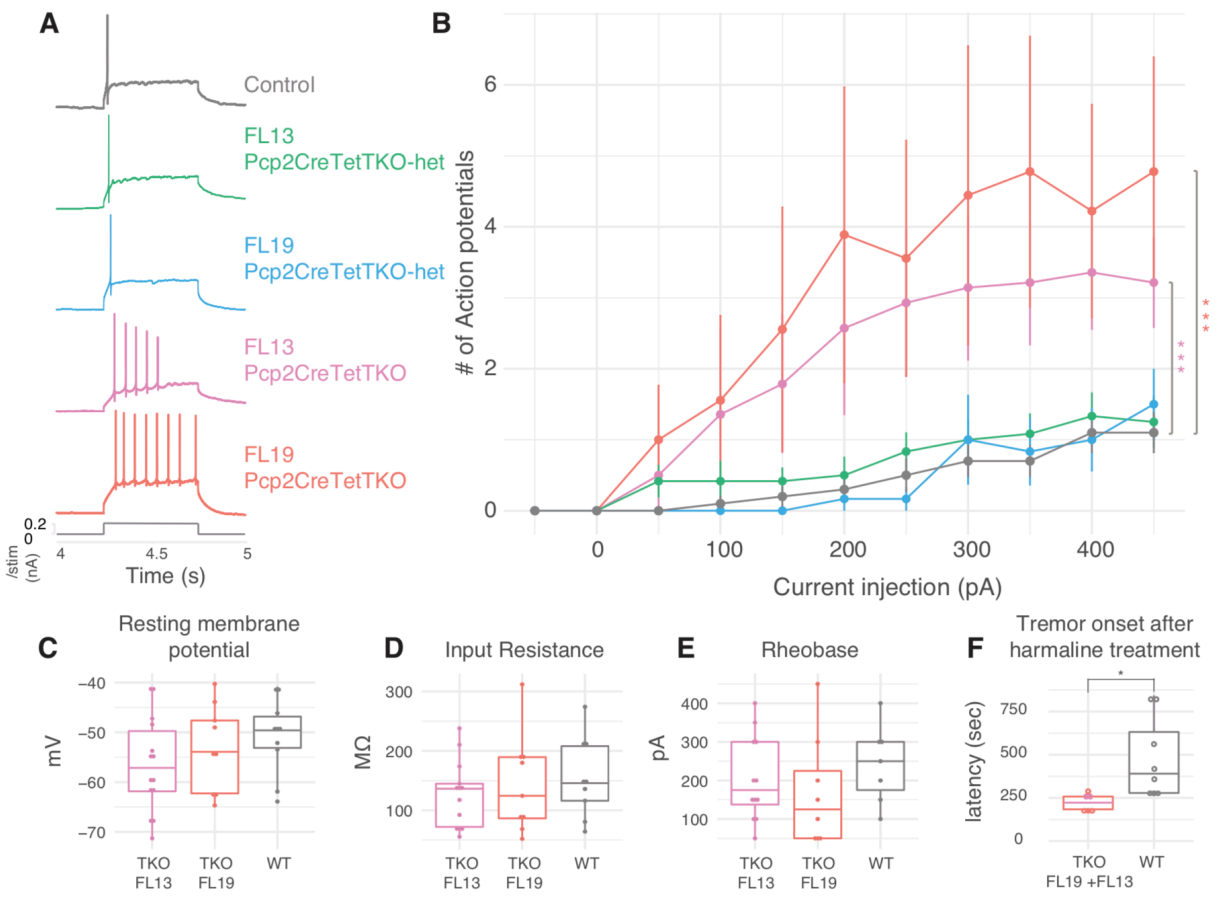


Figure 3.8. Conditional Tet 1,2,3 deficiency causes hyper-excitability and increased susceptibility to excitotoxic drugs.
A. Example whole cell current clamp recordings following 200pA current injection. **B**. F-I curve showing the frequency of action potentials elicited by a 50pA stepwise increase in injected current lasting 500msec (mean±se), p=1.26e-10 for Control and FL13 comparison, p=1.94e-13 for Control and FL19 comparison, n=6-14 cells per group, two-way ANOVA with Tukey HSD. **C**. Resting membrane potential, recorded after breaking the membrane with no current injected. **D**. Input resistance at -50pA current injection **E**. Rheobase current, the minimum current to elicit an action potential. **F**. Harmaline injection and onset of tremor induction, p-value = 0.04, student t-test. Data collected by Michael Riad.

3.5 Discussion

We generated a new mouse strain which has a Purkinje cell specific deletion of TET1, TET2 and TET3. This is the first study which executes a triple TET knockout in a postmitotic neuron and investigates its effect on the molecular and electrophysiological properties of the affected cells. We show that TET depletion causes aberrant gene expression, disrupting numerous ion channel and synaptic activity genes. After analyzing the DNA modification status of Adult specific regulatory regions associated with the affected genes, we show that Pcp2CreTetTKO has increased levels of both 5-hydroxymethylation and 5-methylcytosine, rendering the sites potentially inaccessible. This epigenetic dysregulation causes hyperexcitability and increased susceptibility to excitotoxic drugs. We show that the TET proteins are required for the proper chromatin remodeling during terminal differentiation of neurons.

In this study, we chose Pcp2 as the Cre driver, whose expression starts around P7, during the period of rapid granule cell expansion and synapse formation. However, a driver that recombines even earlier during development could elucidate further details about 5hmC and its role in gene expression regulation. Depletion of the TET proteins at E14.5 would prevent the cell from oxidizing any 5-methylcytosine and ever accumulating any 5-hydroxymethylcytosine, possibly revealing an even bigger phenotype. It also does not escape us that potentially a stronger phenotype could have developed in older animals (>8 weeks), which could reveal whether disruption 5-hydroxymethylcytosine accumulation has a role in aging.

The Simons Foundation Autism Research Initiative organized a study of 50,000 US families impacted by autism to identify potential targets for treatment (Feliciano et al. 2018; Abrahams et al. 2013). In the supporting database, DNMT3A, TET2 and TET3 have all been identified as having a potential role in the neurodevelopmental disorder. Recently, many studies looking into the root causes for multiple complex polygenic neuropsychiatric disorders have shown that there are certain common genes affected and have suggested that it could be due to regulatory regions in the non-coding genome. Other studies have implicated non-coding and loss-of-function mutations of TET2 in neurodegenerative disorders such as Alzheimer's disease, frontotemporal dementia and Parkinson's disease (Cochran et al. 2020, 2; Marshall et al. 2020, 2). Dysregulation of DNMT3A, TET2 and TET3 can affect the DNA modification status of repressive or activating transcription factor binding sites, and enhancers which could lead to abnormal gene expression patterns at any point of a cell's lifespan.

The rules and patterns of 5hmC that we have uncovered in Purkinje neurons can be extrapolated to other principal neurons, as they have similar developmental trajectories. Overall, the TET proteins and DNA hydroxymethylation are emerging as important players in development and disease and potential targets for treatment.

CHAPTER 4. 5-hydroxymethylcytosine dynamics in developing cerebellar granule cells

4.1 Introduction

To complement the Purkinje cell developmental characterization, we profiled granule cells at the exact same timepoints (P0, P7 and adult). In adult granule cells, oxidation of 5-methylcytosine is shown to refine the genomic occupancy of MeCP2, by replacing its high affinity binding sites with low affinity ones in transcribed genes (Mellén, Ayata, and Heintz 2017). The study further shows that hydroxymethylation in CpA context is the most abundant from the CpH nucleotides, and is present in active enhancers and promoters. However, unlike 5-hmCpG, 5-hmCpA has the same binding properties as 5-mC and does not cause functional demethylation. The complicated balance of methylation and hydroxymethylation in CpG or CpH refines and stabilizes the binding of MeCP2 in adult granule cells. In a separate developmental study profiling granule cells at earlier timepoints through immunoprecipitation methods, researchers show an enrichment of 5hmC exon start sites of expressed genes, but cannot distinguish the context (X. Zhu et al. 2016). The study also showed that a Tet1 and Tet3 RNAi knockdown in culture impairs proper dendrite formation and on a molecular level disrupts genes involved in circuit formation.

Besides these studies, not much is known about the developmental dynamics of 5hmC on a single nucleotide level and its relationship to chromatin accessibility in developing granule cells.

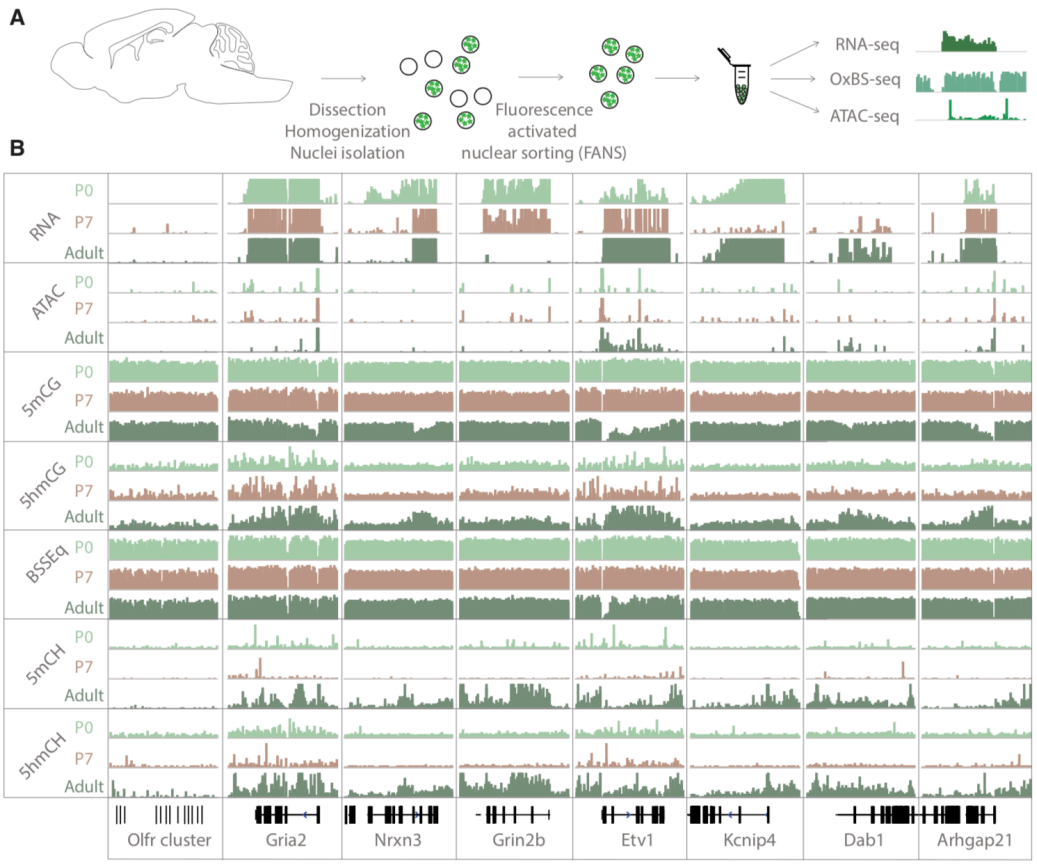


Figure 4.1. Chromatin landscape of the developing cerebellar granule cells
A. Schematic representation of granule cell purification and the downstream applications. **B**. Integrated genome viewer (IGV) representation of example regions of differentially regulated genes (Olfr cluster – always silent, Gria2 – always expressed, Grin2b– developmentally downregulated, Etv1, Dab1 and Arhgap21 – developmentally upregulated, Kcnip4 – expressed in P0 and Adult, but not P7, Nrxn3 – showing an isoform switch). Top tracks show RNA expression in RPKM (reads per kilobase per million mapped reads), mCG tracks show methylation level in CG context from 0 to 0.8, hmCG tracks show hydroxymethylation level in CG context from 0 to 0.45, mCH and hmCH show methylation and hydroxymethylation level in CH context (H = A, C or T) from 0 to 0.04. ATAC tracks show ATAC-seq read density in RPKM from 0 to 10.

4.2 Granule cell chromatin landscape dynamics

We chose to analyze cerebellar granule neurons at the same developmental stages as Purkinje cells: at the first postnatal day (P0), during the rapid phase of their division (P7) and as fully mature, differentiated neurons (adult). For the purification of granule cells, we relied on Neurod1-TRAP mice. TRAP mice have the ribosomal protein L10a fused to green fluorescent protein (GFP) and a cell type specific promoter. Because ribosomes assemble in the nucleolus, their nuclei are GFP-positive and therefore suitable for purification by FACS without the necessity for a cell type specific antibody (Kriaucionis and Heintz 2009). Furthermore, Neurod1 is expressed at all three timepoints, making it possible to use the same mouse line for this study (**Figure 4.1A**).

We generated high resolution data sets of the transcriptome, DNA methylation and hydroxymethylation, and open chromatin at all three timepoints (**Figure 4.1B**). We purified the nuclei using flow cytometry (**Figure 4.2A**). We started with the transcriptome profiling in order to confirm that our sorts are pure, without major contamination from other cell types such as glia and Purkinje cells (**Figure 4.2B**). We evaluated the quality of the datasets by performing Pearson correlation (**Figure 4.2C**) and showing scores of over 0.9 for the pairs of replicates. Then, we checked for expression of markers from contaminating cell types (**Figure 4.2B**). The granule datasets did not have contamination from glia or Purkinje neurons, and they show differential expression of genes and isoforms (**Figure 4.2D**). After we confirmed that the methodology is robust, we continued with the generation of OxBS-Seq and ATAC-Seq. We performed Pearson correlation between the values for methylation and hydroxymethylation over gene bodies in both CpG and CpH context, showing high scores for each time point (**Figure 4.2E-F**). The quality of the OxBS-Seq data was further validated by examining the oxidation rates for the C to T conversion, which were all within the expected range (**Figure 4.2G**).

Most studies that investigate DNA modifications use BS-Seq as the profiling method, completely ignoring the differential binding and functional properties of 5-methylcytosine and 5-hydroxymethylcytosine by not separating them. Here we show that not delineating the two modifications can be misleading about the underlying events (**Figure 4.3A**). At P0 and P7, the gene body of Arhgap21 does not accumulate much 5hmCG. At the adult timepoint however, it accumulates a significantly higher amount compared to P0 and P7 (**Figure 4.3B**, 5hmCG), and this is only evident from the OxBS-Seq data. From the BS-Seq data it seems that the DNA modifications do not change significantly from P0 to adult (**Figure 4.3B**, 5hmCG+5mCG).

From the transcriptomic data we identified multiple examples of gene expression dynamics. Similar to the Purkinje cells, we have the Olfr cluster, which is not expressed and heterochromatic (**Figure 4.1B**, Olfr); Gria2, which is constitutively expressed (**Figure 4.1B**, Gria2); Grin2b – developmentally downregulated (**Figure 4.1B**, Grin2b); Etv1, Dab1 and Arhgap21 – developmentally upregulated (**Figure 4.1B**, Etv1, Dab1, Arhgap21); Kcnip4 – expressed in P0 and Adult, but not P7 (**Figure 4.1B**, Kcnip4); Nrxn3 – showing an isoform switch (**Figure 4.1B**, Nrxn3). We used differential gene expression to computationally define two categories of genes – developmentally up- and down-regulated (filtering for significance of $p<0.01$ and log2 fold change of > 2 in either direction) (**Figure 4.4A-C**). We identified 578 developmentally upregulated genes (or genes up in Adult), and 344 developmentally downregulated genes (or genes that were up in P0). Similar analysis comparing P7 and adult showed 990 upregulated genes and 332 downregulated ones. A differential expression analysis

between P0 and P7 timepoints also yielded changing genes, although much less than the other two comparisons (upregulated: 34, downregulated: 147).

To explore the chromatin changes in developmentally downregulated genes, we used a similar comparison to what we did with the Purkinje cells in **Chapter 2**. We compared the developmentally dynamic genes to repressed ones (always off): Hox transcription factors, which were expressed earlier in cell lineage and olfactory and vomeronasal receptors, which are never expressed in the Purkinje lineage. Hox transcription factors lack cytosine methylation or hydroxymethylation, are not accessible. On the other hand, the olfactory and vomeronasal receptors are highly methylated, lack hydroxymethylation and are completely inaccessible as well (**Figure 4.4D-G**). Promoters of developmentally repressed genes become inaccessible, but this is not correlated with changes in DNA methylation. There is no accumulation of 5hmCG over gene bodies of expressed genes through P0 and P7, and in Adult the significant accumulation is only the constitutively expressed genes.

Overall, as granule cells are constantly dividing, they create a heterogeneous mix at the earlier developmental timepoints (P0 and P7), which results in fairly heterogeneous epigenetic profiles that are difficult to interpret.

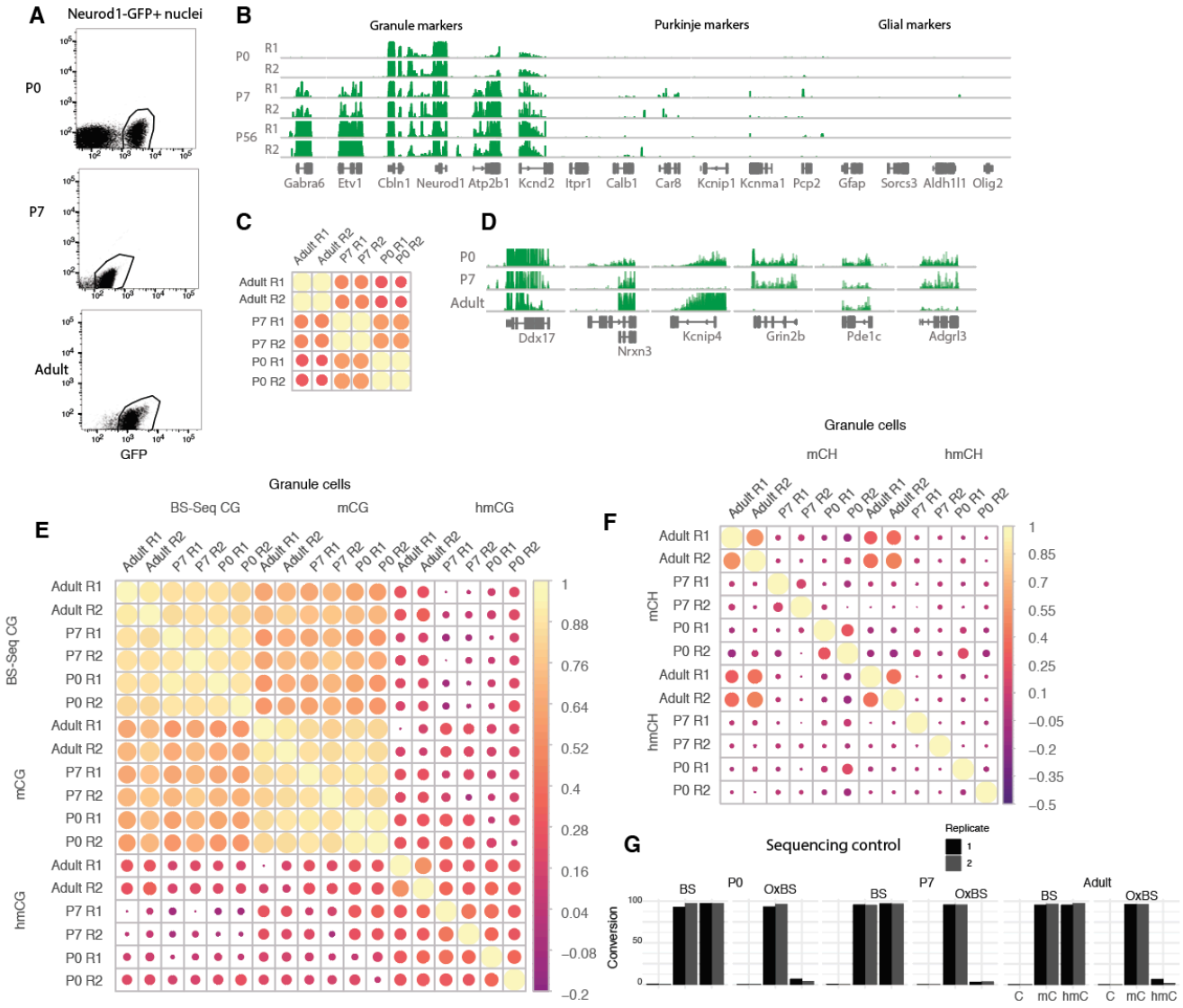


Figure 4.2. Quality control of sequencing datasets from granule cells

A. Example plots of fluorescence activated nuclear sorting of Granule cells at P0, P7 and adult timepoints with endogenous Neurod1-GFP expression. **B**. IGV representation of Granule specific markers (Itpr1, Calb1, Car8, Kcnip1, Kcnma1, Pcp2) enrichment and depletion of granule (Gabra6, Etv1, Cbln1, Neurod1, Atp2b1, Kcnd2) and glial (Gfap, Sorcs3, Aldh1l1, Olig2) markers. **C**. Pearson correlation of RNA-Seq datasets. **D.** IGV representation of Granule genes dynamics during differentiation. **E**. Pearson correlation of gene body accumulation of 5hmCG+5mCG (BS-Seq), 5hmCG and 5mCG (derived from maximum likelihood estimator model of BS-Seq and OxBS-Seq). **F**. Pearson correlation of gene body accumulation of 5hmCH and 5mCH (derived from maximum likelihood estimator model of BS-Seq and OxBS-Seq). **G**. Bisulfite conversion and oxidation efficiency in BS-Seq and OxBS-Seq datasets.

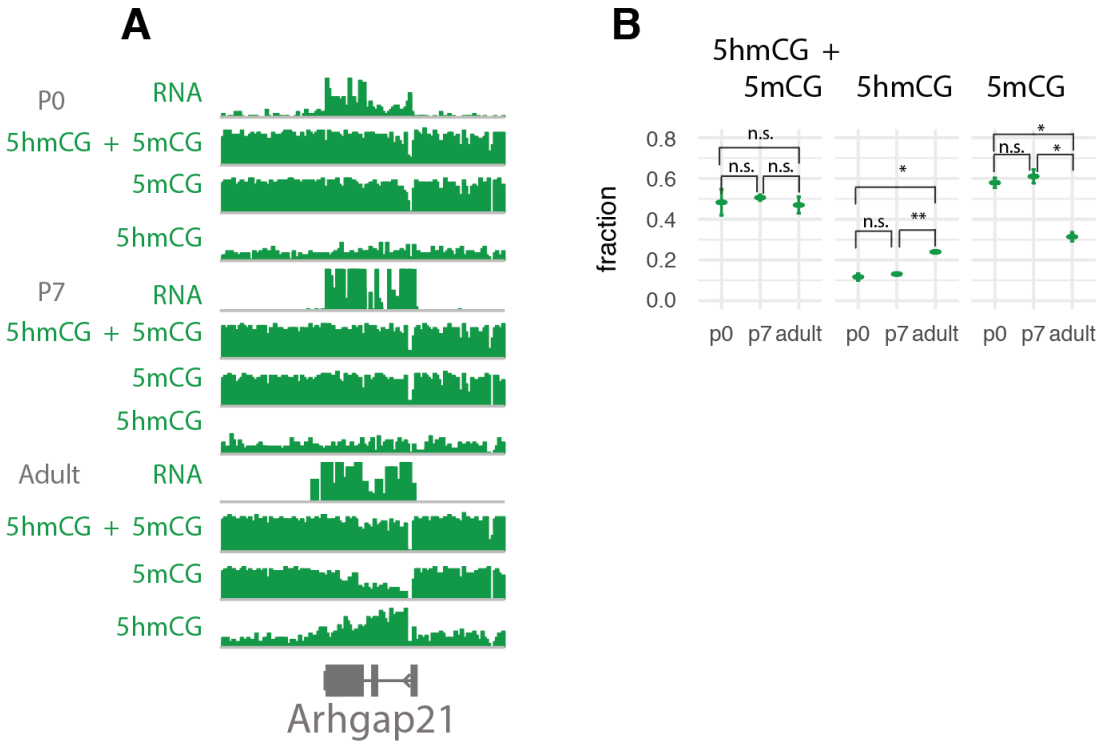


Figure 4.3. BS-Seq does not give proper resolution to 5-methylcystosine and 5-hydroxymethylcytosine.
A. IGV representation of BS-Seq (showing 5hmCG+5mCG signal together) and OxBS-Seq (showing 5hmCG+5mCG signal separately) at the three developmental timepoints **B**. Quantification of each modification. 5hmCG+5mCG refers to signal solely fr0m BS-Seq data. 5hmCG and 5mCG refers to computationally deriving those values from BS-Seq and OxBS-Seq data.

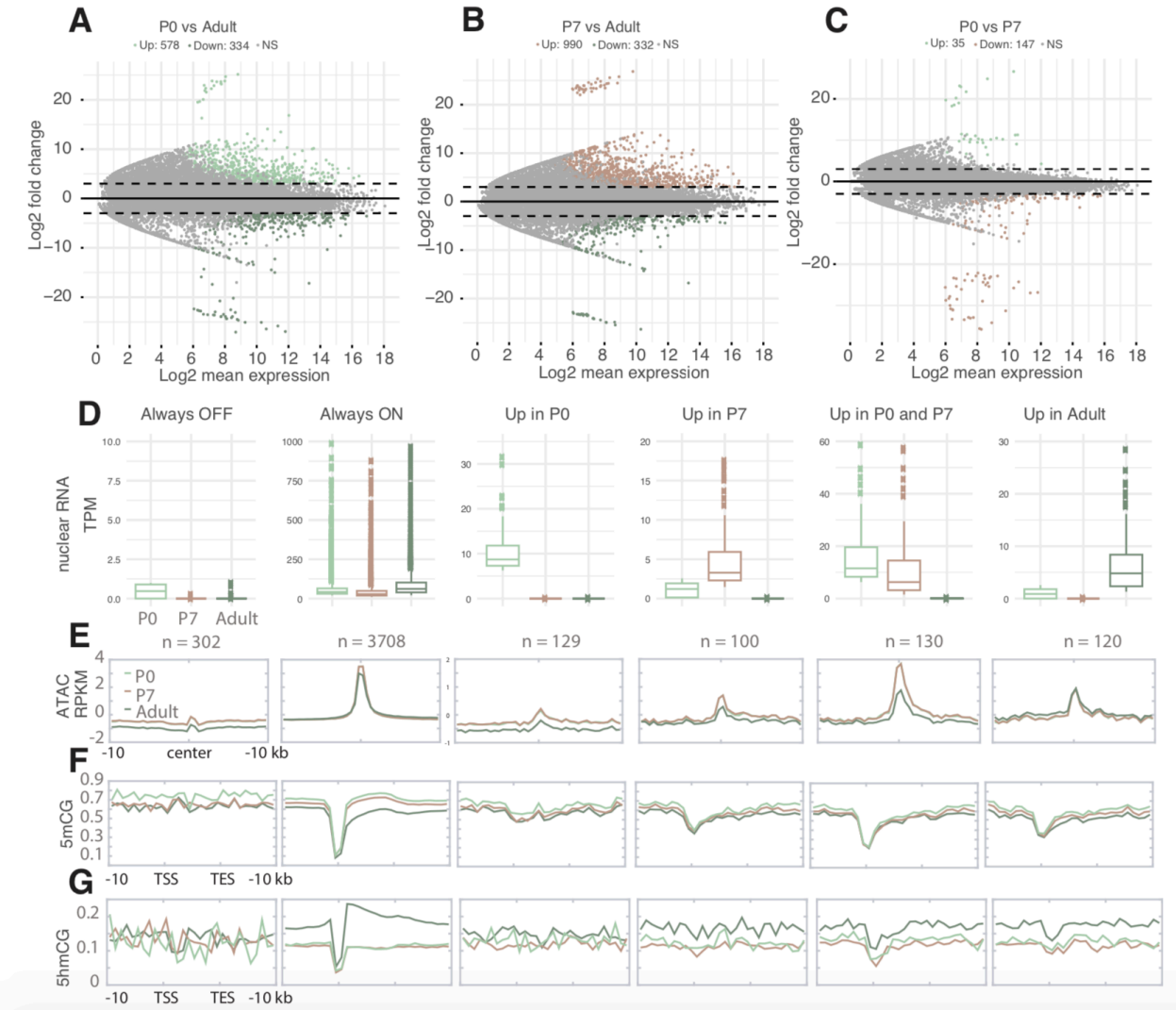


Figure 4.4. Differential expression and cytosine modification dynamics in granule cells

A-C. Differential expression analysis between P0 and Adult (**A**), P7 and Adult (**B**) and P0 and P7 (**C**). D. Quantification of the expression levels of each group of genes at the three developmental stages. **E-G**. Metagene representation of the average signal of the gene bodies of each group: ATAC-seq (**E**), 5mCG (**F**) and 5hmCG (**G**).

4.3 DNA hydroxymethylation differences between Purkinje and granule cells.

Since granule and Purkinje cells have such different developmental trajectories – one dividing up until P21, while the other postmitotic at E14.5 – we wanted to explore whether the rules for 5hydroxymethylcytosine accumulation are different. We plotted the expression level of genes versus the accumulation of either 5-hydroxymethylcytosine in each timepoint in each cell type (**Figure 4.3A-B**). P0 and P7 granule cells do not have the strong positive correlation between expression and 5hmC accumulation that the adult timepoint has. However, in Purkinje cells we see a strong positive correlation at each timepoint. At the P7 and adult timepoint, for genes with expression of log2(TPM+1) >=15, we see that correlation decrease, which could be explained by the active demethylation of the highest expressed genes discussed in Chapter 2.

To continue explore this relationship, we isolated two groups of genes – the top and bottom 20% genes ranked by expression (top 20% being the most expressed, bottom 20% the least). We then plotted the average modifications over their gene bodies using a metagene plot (**Figure 4.3C-D**). In granule cells, we clearly observe that only expressed genes in adult granule cells accumulate 5-hydroxymethylcytsoine over the gene bodies as previously reported. This could be explained by the passive demethylation occurring during mitosis. Since they are postmitotic, Purkinje cells have a steady accumulation of 5-hydorxymethylcytosine in expressed genes at each timepoint.

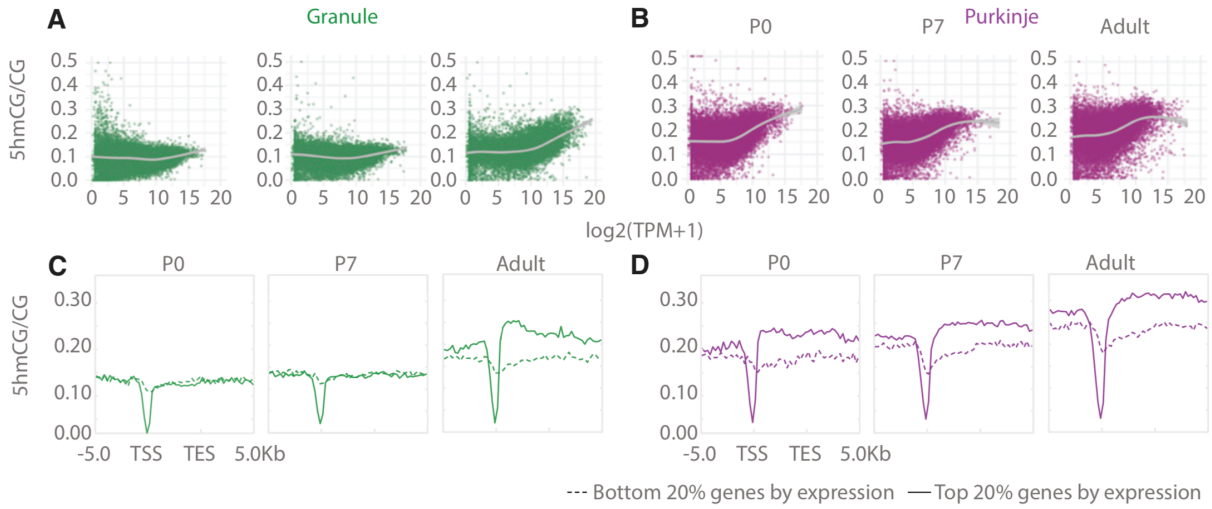


Figure 4.5. DNA hydroxymethylation accumulation differences between granule and Purkinje cells.

A-B. Scatter plots between expression in log2(TPM+1) on the x axis and 5hmCG on the y axis showing the correlation between them. Granule cells (left) in green, Purkinje cells (right) in purple. **C - D**. Metagene plot representation of the average 5hmCG over the top 20% (straight line) or bottom 20% (dashed line) of genes ranked by expression. Granule cells (left) in green, Purkinje cells (right) in purple.

4.4 Putative regulatory regions in differentiating granule cells

Similar to Purkinje cells, the open chromatin landscape of differentiating granule cells is very dynamic. To explore that, we computationally identified all the regions of the genome that have a statistically significant enrichment of ATAC-Seq reads at each time point. Following, we performed a differential accessibility analysis, which resulted in multiple regions that are changing between P0 and Adult (**Figure 4.4A**). Interestingly, the open chromatin landscape between P0 and P7 is very similar and it doesn't seem to be changing much. This could be due to the fact that between P0 and particularly at P7, the granule cells are actively dividing and migrating, creating a heterogeneous population of cell at different stages of development. Like the analysis we did for Purkinje cells, we divided the changing regions into two groups based on how significant their change is ($p<0.01$) and its magnitude (log2 fold change > 4) – regions that either "gained" (more accessible in adult compared to P0) or "lost" accessibility (less accessible in adult compared to P0) and explored the cytosine modification dynamics in them. In granule cells we observe the same pattern of loss of accessibility – the regions do not accumulate methylation or hydroxymethylation (**Figure 4.4B-C, Loss section**). This, again, points to some other mechanism mediating silencing that does not involved DNA modifications. Regions that gain accessibility unlike Purkinje cells do not lose hydroxymethylation, only methylation (**Figure 4.4B-C, Gain section)**. This is potentially due to the granule cells dividing up until P21, undergoing multiple rounds of passive demethylation. The center of the regions that gain accessibility are somehow protected from modification accumulation, possible through the binding of transcription factors, and we can see

Since we had two group of regions that are developmentally important, we decided to explore that transcription factor motifs we can identify in each. We found motifs of several transcription factors that have differential expression levels between P0 and Adult (**Figure 4.4D**).

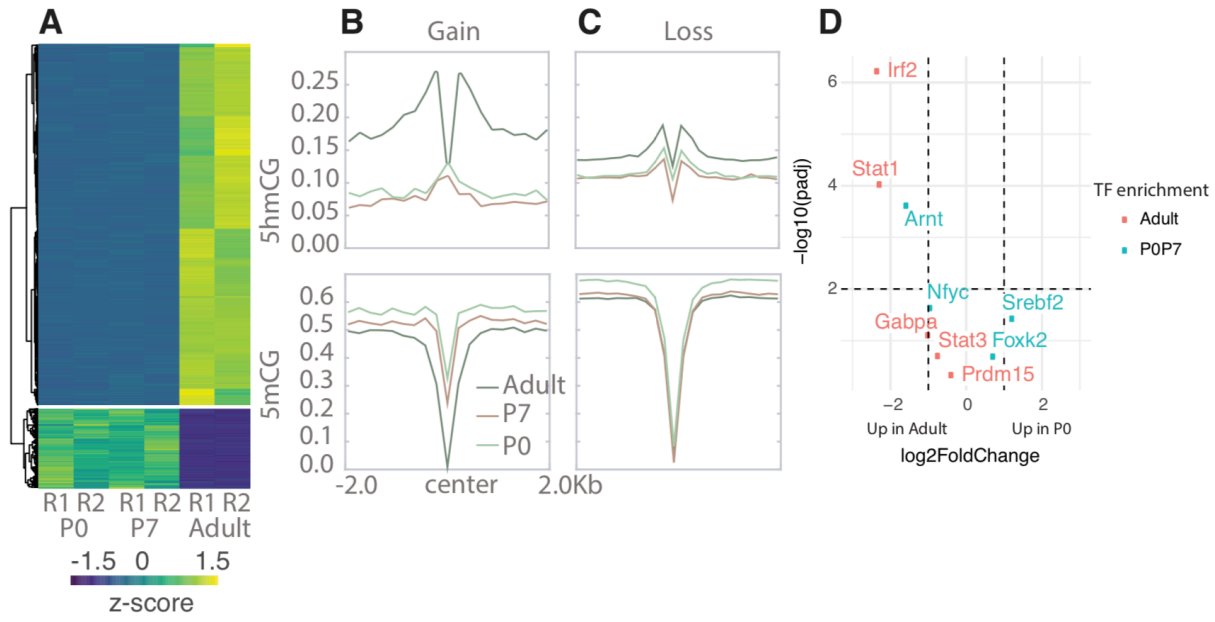


Figure 4.6. Open chromatin dynamics in granule cells
A. Heatmap representing the differentially accessible regions between P0 and adult GCs ($p < 0.01$, log2(fold change) > 4). **B-C**. Metagene plots representing the mean values of 5hmCG/CG (**B**) and 5mCG/CG (**C**) over the centers and flanking regions of peaks that gained or lost accessibility relative to adult GCs. **D**. Differential expression levels of transcription factors, whose motifs were present in the differentially accessible regions.

4.5 Discussion

We purified cerebellar granule cells at three developmental timepoints – P0, P7 and adult – using Neurod1-TRAP animals and flow cytometry. We profiled their chromatin landscape using transcriptomics, DNA methylation and hydroxymethylation, and chromatin accessibility. We show that granule cells undergo a lot of changes in gene expression, alongside non-coding regions that become accessible or not during development. Interestingly, we also show a dramatic difference in how Purkinje and granule cells accumulate 5-hydroxymethylcytosine. As PCs are postmitotic, they accumulate 5-hydroxymethylcytosine at each developmental stage, which is then maintained even if the genes are no longer expressed. In granule cells, we see the expected positive correlation between expression and 5-hydroxymethylcytosine only in the Adult timepoint, which is caused by passive dilution of the mark through division.

Because of granule cells' rapid division between P0 and P7, it is difficult to explore their genomic landscape and work through their heterogeneity. It would be interesting to continue these studies by focusing on the role of 5-hydroxymethylcytosine in aging, and whether similar to Purkinje cells, granule cells would require the Tet proteins to remodel regulatory regions after they have become postmitotic.

CHAPTER 5. Summary and Perspective

Summary

In **Chapter 2**, we investigated the developmental dynamics of 5-hydroxymethylcytosine in Purkinje cells. We show that PCs undergo dynamic transcriptional changes which are associate with dynamic chromatin accessibility. We confirm the positive correlation between 5-hydroxymethylcytosine, but we also show that genes that have accumulated higher levels of 5-hmC do not lose it once they are no longer expressed. This suggests that 5-hmC alone is not enough to promote nor repress transcription. Furthermore, we describe a novel class of epigenetically regulated genes, which are progressively demethylated during development, and are associated with broad regions of chromatin accessibility at H3K4me3 accumulation.

In **Chapter 3**, we tested the effect of a triple Tet depletion, by generating a conditional knockout in Purkinje cells. We show that the Tet proteins are required for proper expression of genes associated with ion channels, which is due to impaired remodeling of the regulatory regions associated with them. These molecular disruptions lead to impaired electrophysiological properties in adult Purkinje neurons.

In **Chapter 4**, we analyzed the developmental dynamics of 5-hydroxymethylcytosine in granule cells. We show that the positive correlation between expression and 5-hmC is only present at the postmitotic adult timepoint, while at the mitotic points at P0 and P7 there is no such relationship. We also identify multiple regions with dynamic accessibility status and transcription factor that could bind them.

The future of neuroscience at a cell type level

Around the time I started working on this project, there was an explosion of techniques targeting single pure population of neurons. They range from physically dissecting single cells through laser capture microscopy to high throughput sequencing approaches through ribosome affinity purification, and fluidics and plate based single cell sequencing. There are new methods for sequencing published daily. Massive organizations such as the Allen Brain Institute invest large amounts of money into curating atlases of every single neuronal type in the human and mouse brain. These tremendous efforts have vastly improved what we know about the brain, but they are hitting a dead end once we want to profile beyond the transcriptome or manipulate newly identified populations. Single cell sequencing in particular is not suited for oxidative bisulfite sequencing and attempts at chromatin immunoprecipitation and chromatin accessibility have had moderate success at best. Furthermore, very few techniques allow for the identified neuronal cell types to be mapped back physically in their location, making it difficult to target or experiment on newly discovered cell types. The biggest hurdle in molecular neuroscience lies in deriving transcriptomic, epigenetics, electrophysiological and morphological data from rare pure neuronal populations from the same source, without the need for transgenic tools.

This way of thinking about organs, as multiple different cell types and not homogeneous tissues, is revolutionizing the way we approach experiments beside academia. The biotechnology industrial complex also recognizes the need for a cell type specific profiling, as it makes for a more efficient drug screening, saving precious time, money and resources.

The importance of the non-coding genome in psychiatric and neurodevelopmental polygenic disorders, neurodegenerative disorders and aging

For years, complex polygenic neuropsychiatric disorders like schizophrenia, autism and obsessive-compulsive disorder have baffled scientists. They affect higher order cortical functions, such as mood and cognition, which are difficult to quantify and model, compared to more basic biological functions. Analysis of large cohorts of autism spectrum disorder (ASD) probands and their unaffected family members have identified dozes of ASD risk loci and genes, from copy number variations to insertions/deletions and single nucleotide polymorphisms. In particular, chromatin remodelers such as DNMT3A, TET2 and TET3 have been identified in ASD, which suggest a role for chromatin organization and methylation in disease progression. This, along with other studies, sparked a lot of interest in the non-coding genome. In has been shown that human regulatory regions are negatively selected, which implies that any mutations in those regions could be deleterious (The ENCODE Project Consortium 2012). Disruptions in common regulatory regions early in development or in aging can affect whole pathways at once, increasing the difficulty of modeling the disease or pinpointing a strong therapeutic target. It is feasible to imagine that aberrant methylation of regulatory regions can affect transcription factor binding and therefore downstream expression of multiple genes. If that is the case, the possibility of active demethylation in postmitotic neurons opens the door for new potential treatments. Addressing neurological disorders at the epigenetic level, where transcription patterns could be repaired for multiple genes at ones, could be a much more effective approach than the established pharmacological treatments of synaptic activity.

Materials and Methods

Animals

Wildtype C57BL6/J (RRID:IMSR_JAX:000664) were obtained from Jackson Laboratories. Animals were maintained on a 12hr light/12hr dark cycle with food and water ad libitum. Animal protocols were approved by the Rockefeller University Institutional Animal Care and Use Committee, in accordance with the US National Institutes of Health Guide for the Care and Use of Laboratory Animals.

Pcp2CreTetTKO murine strain generation

Tet1fl/fl::Tet2fl/fl::Tet3fl/fl animals were a gift from Anjana Rao. We used a previously characterized Pcp2-CRE BAC construct from the GENSAT project (Gong et al. 2003; 2007) to create a conditional TKO in Purkinje cells. Pronuclear injections of the BAC construct into zygotes were performed at the Transgenic and Reproductive Technology Center at the Rockefeller University.

Cell type specific nuclei isolation

Nuclei isolation: Nuclei isolation protocol was followed as described (X. Xu et al. 2018). Briefly, mouse cerebella were dissected and flash frozen using liquid nitrogen. To isolate nuclei, tissue was thawed on ice for 30 min and then transferred to 5mL of homogenization buffer (0.25 M sucrose, 150 mM KCl, 5 mM MgCl2, 20 mM Tricine pH 7.8, 0.15 mM spermine, 0.5 mM spermidine, EDTA-free protease inhibitor cocktail, 1mM DTT, 20U/mL Superase-In RNase inhibitor, 40U/mL RNasin ribonuclease inhibitor). Tissue were homogenized by 30 strokes of loose (A) followed by 30 strokes of tight (B) glass dounce. Homogenate was supplemented with 5mL of a 50% iodixanol solution (50% Iodixanol/Optiprep, 150 mM KCl, 5 mM MgCl2, 20 mM Tricine pH 7.8, 0.15 mM spermine, 0.5 mM spermidine, EDTA-free protease inhibitor cocktail, 1mM DTT, 20U/mL Superase-In RNase inhibitor, 40U/mL RNasin ribonuclease inhibitor), and laid on a 27% iodixanol cushion. Nuclei were pelleted by centrifugation 30 min, 10,000 rpm, 4°C in swinging bucket rotor (SW41) in a Beckman Coulter XL-70 ultracentrifuge. The nuclear pellet was resuspended in homogenization buffer.

Nuclei labeling and sorting: Nuclei were fixed with 1% formaldehyde for 8 minutes at room temperature (RT) with mild agitation. The crosslinking reactions was quenched with 0.125M glycine for 5 minutes at RT. Nuclei were pelleted at 1000g, 4 minutes, 4C, and then washed two times with Wash Buffer (PBS, 0.05% TritonX-100, 50ng/mL BSA, 1mM DTT, 10U/uL Superase-In RNase Inhibitor). Nuclei were blocked with Block Buffer (Wash buffer with an additional 50ng/mL BSA) for 30 minutes at RT, incubated with primary antibody for 1 hour at RT, and then washed three times with Wash Buffer with spins in between washes as described above. Nuclei were then incubated in secondary antibody for 30 minutes at RT and washed three times with Wash Buffer. Primary and secondary antibodies were diluted in Block Buffer. Secondary antibodies were purchased from Life Technologies or Jackson

Immunoresearch and were used at 1:500 dilution. Secondary antibodies from goat used: mouse Alexa488, rabbit Alexa594.

FACS: Nuclei were stained with DyeCycle Ruby to 20uM final concentration. Nuclei were sorted using a BD FACSAria cell sorter using the 488nm and 561nm lasers. First, samples were gated using DyeCycle Ruby to select only singlets, then appropriate populations were gated based on their separation. Analysis was performed using FlowJo software. Qiagen Buffer PKD was added to the sorted nuclei and the samples were stored at -80C.

Immunohistochemistry

Mice were decapitated at P0 and P7, and the brains were dissected, and immersion fixed in 4% formaldehyde (w/v) overnight at 4C. Adult mice were deeply anesthetized and then the brains were fixed by transcardiac perfusion with PBS followed by 4% formaldehyde. Brains were further fixed by immersion fixation in 4% formaldehyde overnight at 4C. All brains were cryoprotected in 30% sucrose in PBS, embedded in OCT and cut with a Leica CM3050 S cryostat into 20um sections. The sections were immediately mounted on slides and stored at -20C. Antigen retrieval using sodium citrate buffer (10mM sodium citrate, 0.05% Tween 20, pH 6.0) was performed by heating the slides to 95-100C and then 10-minute incubation in the microwave at the lowest power. The slides were cooled off to room temperature for 1 hour. The slides were washed in PBS and blocked with 3% BSA in PBS with 0.1% TritonX-100 for 30 minutes at room temperature. Primary antibody incubation was performed overnight at room temperature, washed with PBS, incubated with secondary antibody for 1 hour at room temperature, washed with PBS, stained with DAPI (1:10000) for 10 minutes at room temperature and washed three times with PBS. Slides were cover slipped with Prolong Diamond mounting media. Images were acquired using a Zeiss LSM700 confocal microscope using the same acquisition settings for all samples. Further image analysis was done using FIJI.

Sequencing methods

RNA-Seq: RNA and gDNA from fixed nuclei was purified using the Qiagen AllPrep FFPE kit with the following modifications from. After DNA/RNA separation spin, the RNA-containing supernatant was removed and incubated at 65oC for 30 minutes, 70oC for 30 minutes, and 80oC for 15 minutes and then proceeded with the manufacturer's protocol. gDNA was purified following the rest of the manufacturer's protocol. RNA quality was determined using Agilent 2100 Bioanalyzer. Purified RNA was converted to cDNA and amplified using the Nugen Ovation RNA-Seq System V2. cDNA was fragmented to an average size of 200bp using a Covaris C2 sonicator (intensity 5, duty cycle 10%, cycles per burst 200, treatment time 120 seconds). Libraries were prepared using the NEBNext Ultra DNA Library Prep Kit for Illumina with NEBNext Multiplex Oligos for Illumina. The quality of the libraries was assessed using the Agilent 2200 TapeStation system with D1000 High Sensitivity ScreenTape. Libraries were sequenced at The Rockefeller University Genomics Resource Center on the Illumina NextSeq 500 to obtain 75bp paired-end reads.

ATAC-Seq: ATAC-seq libraries were prepared as described (Buenrostro et al. 2013; Chen et al. 2016) with minor modifications. ~25k fixed nuclei were incubated for 10 minutes in

lysis buffer (10 mM Tris pH 7.5, 10 mM NaCl, 3mM MgCl2, 0.1% NP-40). Nuclei were then resuspended in 50 ul 1x TD Buffer containing 2.5 ul Tn5 enzyme from Nextera (Illumina) and incubated for 30 min at 37°C. 200ul of reverse-crosslinking buffer (50 mM Tris–Cl, 1 mM EDTA, 1% SDS, 0.2 M NaCl, 5 ng/ml proteinase K) was added to the samples and they were incubated overnight at 65C with shaking. The samples were then purified using the QiaQuick MinElute columns (Qiagen). Libraries were amplified by PCR using the Q5 High Fidelity Polymerase (NEB) for 12 cycles with barcoded primers. Libraries were size-selected with AMPure XP beads (Beckman Coulter). Libraries were sequenced at The Rockefeller University Genomics Resource Center on Illumina NextSeq 500 to yield 75bp paired-end reads.

OxBS-seq: DNA conversion and library preparation were performed using the CEGX TrueMethyl-Seq Whole Genome kit (Cambridge Epigenetix, Cambridge, UK) and after its acquisition from Tecan, the Ultralow Methyl-Seq with TrueMethyl oxBS Module (#0541 and #9513) following the manufacturer's instructions. Briefly, the DNA was sheared to 800bp using Covaris sonicator and treated with the oxidation agent and bisulfite following the manufacturer's protocol. Libraries were sequenced on Illumina NextSeq 500 to yield 75 bp paired-end reads. The efficiency of the DNA oxidation and conversion was assessed by interrogating the spike-in Digestion Control and Sequencing Control using the dockerized custom pipeline bsExpress from CEGX (https://bitbucket.org/cegx-bfx/cegx_bsexpress) and fell within the expected ranges (≥90% for hmC conversion in the OxBs reaction, <10% for hmC conversion in the BS reaction (hmC/BS over-conversion error rate), <5% for mC conversion in both the OxBs and BS reactions (over-conversion error rate)).

ChIP-seq: Chromatin immunoprecipitation was performed using the Low Cell ChIP-Seq kit (#53048, Active Motif) following the manufacturer's instructions. Libraries were prepared using the supplied library prep reagents. The samples were sequenced on Illumina NextSeq 500 to yield 75 bp paired-end reads.

Bioinformatic Data analysis

Most data analysis was done in the R/Bioconductor environment (Huber et al. 2015) in RStudio (https://www.R-project.org/, http://www.rstudio.com/). For general processing, data exploration and visualization we used the tidyverse array of packages, in particular ggplot and dplyr (Wickham et al. 2019).

RNA-seq

RNA-seq reads were aligned using STAR (Dobin et al. 2013) and genome assemblies from UCSC. In addition to default STAR parameters, we used the following for paired-end data (--outFilterMismatchNmax 999 --alignMatesGapMax 1000000 -- outFilterScoreMinOverLread 0 --outFilterMatchNminOverLread 0 --outFilterMatchNmin 60 -- outFilterMismatchNoverLmax 0.05). Aligned reads were converted to bigwigs for visualization in IGV (Robinson et al. 2011) using deepTools (Ramírez et al. 2016). UCSC gene model annotations for whole genes were downloaded using the UCSC Table Browser tool. Transcript level quantifications was performed using Salmon (v1.1.0) (Patro et al. 2017) and imported for differential expression analysis with tximport (v1.10.1) (Soneson, Love, and Robinson 2015) and DESeq2 (v1.22.2) (Love, Huber,

and Anders 2014). For up- and down-regulated genes we selected for a log2 fold change of 2 and p-adjusted value of 0.05. Gene ontology analysis was performed using GOrilla (http://cbl-gorilla.cs.technion.ac.il/) (Eden et al. 2009) or DAVID (https://david.ncifcrf.gov) (Huang, Sherman, and Lempicki 2009).

ATAC-seq

Reads were processed with trim_galore (Martin 2011) with parameters "--stringency 3 --fastqc --paired". Trimmed reads were mapped to mm10 using bowtie2 (version 2.1.0) (Langmead and Salzberg 2012) with parameters "-X 2000 --no-mixed --no-discordant". Duplicates were removed using samtools (H. Li et al. 2009). Reads were normalized to RPKM using deepTools bamCompare module, ignoring chrX, chrY, chrM and filtering reads for minimum mapping quality of 30. Metagene and heatmap profiles were generated using deepTools modules computeMatrix, plotProfile and plotHeatmap. Unique fragments under 100 nt were used to call peaks with macs2 (Y. Zhang et al. 2008) with parameters "--nomodel -q 0.01 --call-summits". Broad peaks were called with macs2 as well with parameters "--nomodel -f BAM --keep-dup all --broad -g mm -B -q 0.01". The peaks were filtered to remove chrY and chrM peaks and peaks that overlap with the mm10 blacklist from ENCODE (The ENCODE Project Consortium 2012). DiffBind (v2.10.0) (Ross-Innes et al. 2012) was used to identify differentially accessible chromatin regions using the DESeq2 method. Peaks were selected based on 4-fold difference and q-value of 0.05. Differential motif enrichment was performed using chromVAR (v1.4.1) (Schep et al. 2017).

OxBS-seq

Reads were processed with trim_galore with parameters "--stringency 3 --fastqc --paired --clip_R1 5 --clip_R2 10". Trimmed reads were mapped to mm10 using bismark (Krueger and Andrews 2011) with parameters "--bowtie2 -p 4 --multicore 4" (v0.20.0). Duplicates were removed using deduplicate_bismark. The following tools from methpipe (v3.4.2) (Q. Song et al. 2013) were used for downstream statistical estimation of the methylation and hydroxymethylation levels. First, bismark aligned reads were converted to the custom .mr format and sorted with default parameters. Methylation calls were extracted with methcounts with default parameters. Methylation and hydroxymethylation levels were estimated using the mlml tool with default parameters. Genome browser files were generated with bedGraphToBigWig. UMRs and LMRs were identified using MethylSeekR (v1.22.0) (Burger et al. 2013) with m = 0.5 and 5% FDR. DMVs were identified as UMRs ≥5 kb with mean.meth ≤15 and regions within 1kb of each other were merged (bedtools merge -d 1000). mm10 CpG island annotations were downloaded from the UCSC table browser. Unique Large DMVs were selected by excluding regions under 15kb. The regions were annotated using HOMER (Heinz et al. 2010). DMR analysis was performed using methylpy (v1.3.4) (Schultz et al. 2015) with default parameters.

ChIP-seq

Reads were processed with trim_galore with parameters "--stringency 3 --fastqc --paired". Trimmed reads were mapped to mm10 using bowtie2 (v2.1.0) with default parameters. Duplicates were removed using samtools. QC was performed with ChIPQC (Carroll et al. 2014)

and that estimated fragment size was used for broad peak calling with macs2 (--nomodel -f BAM --keep-dup all --broad -g mm -B -q 0.01).

Electrophysiology

At 6-8 weeks of age, mice were deeply anesthetized with ketamine/xylazine (100mg/kg, 10mg/kg b.w.) and perfused with ice cold dissection buffer (2.5 mM KCl, 0.5 mM $CaCl_2 \cdot 2H_20$, 7.0 mM $MgCl_2 \cdot 6H_20$, 25.0 mM NaHCO3, 1.25 mM NaH_2PO_4, 11.6 mM (+)-Sodium L-ascorbate, 3.1 mM sodium pyruvate, 110.0 mM choline chloride, and 25.0 mM glucose. Parasagittal cerebellar sections were collected at a thickness of 300μ on a Microslicer (DTK-1000N) and allowed to recover in artificial cerebrospinal fluid (aCSF) (2.5 mM KCl, 118 mM NaCl, 1.3 mM $MgCl_2$2.5 mM $CaCl_2$,, 26 mM $NaHCO_3$, 1 mM NaH_2PO_4, 10 mM glucose) at 32°C for 1 hour. Sections were transferred to room temperature and underwent whole-cell electrophysiological recordings. Dissection buffer and aCSF were kept gassed with carbogen. Slices were transferred to the recording chamber and visualized under an Olympus microscope using a (PCI extended camera). Sections were constantly perfused in aCSF gassed with carbogen, temperature controlled to 32°C (TC-324B). Electrophysiology signals were recorded using a HEKA EPC-10 dual patch clamp amplifier. 4-6MΩ glass capillary pipettes were pulled and filled with internal solution for current clamp (130 K-Gluconate, 5 KCl, 10 HEPES, 2.5 MgCl2. 4 Na2ATP, 0.4 Na3GTP, 10 Na-phosphocretine, 0.6 EGTA) or voltage clamp (115 CsMeSO3, 20 CsCl, 10 HEPES, 2.5 MgCl2, 4 Na2-ATP, 0.4 Na-GTP, 10 Na-phosphocreatine, and 0.6 EGTA). After a GΩ seal was achieved, the membrane was disrupted with short bursts of negative pressure to achieve a whole-cell configuration. Series resistance was measured and ranged between 1 to 20MΩ. In current clamp recordings, a 200ms -50pA hyperpolarizing pulse was used to measure membrane resistance. Rheobase current, the minimal current to elicit an action potential, was determined via stepwise injection of 450pA in 50pA increments. Frequency-current (FI) curves were generated across injected currents and quantified using Axograph (v1.7.6).

Harmaline injections

8wk old mice were administered 30mg/kg harmaline (Sigma-Aldrich, #H1392) through intraperitoneal injection as described (Brown et al. 2020). Harmaline tremor consistently developed in both WT and TKO mice. Behavior was scored using JWatcher (v1.0).

Antibody list

Antigen	Species	Vendor	Cat #	RRID	Dilution
Itpr1* Clone ID: S24-18	Mouse	Origene	TA326547		FACS 1:500 IF 1:500
Itpr1* S24-18	Mouse	Abcam	ab190239		FACS 1:500 IF 1:500
Calb1	Rabbit	Immunostar	24427	AB_572222	IF 1:500
GFP	Chicken	Abcam	ab13970		IF 1:500
H3K27me3	Rabbit	Active Motif	39155	AB_2561020	ChIP 4uL
H3K4me3	Rabbit	Active Motif	39159	AB_2615077	ChIP 4uL

*same clone, discontinued at Abcam

References

Abrahams, Brett S, Dan E Arking, Daniel B Campbell, Heather C Mefford, Eric M Morrow, Lauren A Weiss, Idan Menashe, Tim Wadkins, Sharmila Banerjee-Basu, and Alan Packer. 2013. "SFARI Gene 2.0: A Community-Driven Knowledgebase for the Autism Spectrum Disorders (ASDs)." *Molecular Autism* 4 (1): 36. https://doi.org/10.1186/2040-2392-4-36.

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.

Baylin, Stephen B., and Peter A. Jones. 2016. "Epigenetic Determinants of Cancer." *Cold Spring Harbor Perspectives in Biology* 8 (9): a019505. https://doi.org/10.1101/cshperspect.a019505.

Boyer, Laurie A., Kathrin Plath, Julia Zeitlinger, Tobias Brambrink, Lea A. Medeiros, Tong Ihn Lee, Stuart S. Levine, et al. 2006. "Polycomb Complexes Repress Developmental Regulators in Murine Embryonic Stem Cells." *Nature* 441 (7091): 349–53. https://doi.org/10.1038/nature04733.

Brown, Amanda M, Joshua J White, Meike E van der Heijden, Joy Zhou, Tao Lin, and Roy V Sillitoe. 2020. "Purkinje Cell Misfiring Generates High-Amplitude Action Tremors That Are Corrected by Cerebellar Deep Brain Stimulation." *ELife* 9 (March): e51928. https://doi.org/10.7554/eLife.51928.

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.

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.

Butts, T., M. J. Green, and R. J. T. Wingate. 2014. "Development of the Cerebellum: Simple Steps to Make a 'Little Brain.'" *Development* 141 (21): 4031–41. https://doi.org/10.1242/dev.106559.

Carroll, Thomas S., Ziwei Liang, Rafik Salama, Rory Stark, and Ines de Santiago. 2014. "Impact of Artifact Removal on ChIP Quality Metrics in ChIP-Seq and ChIP-Exo Data." *Frontiers in Genetics* 5 (April). https://doi.org/10.3389/fgene.2014.00075.

Chen, Xingqi, Ying Shen, Will Draper, Jason D. Buenrostro, Ulrike Litzenburger, Seung Woo Cho, Ansuman T. Satpathy, et al. 2016. "ATAC-See Reveals the Accessible Genome by Transposase-Mediated Imaging and Sequencing." *Nature Methods* 13 (12): 1013–20. https://doi.org/10.1038/nmeth.4031.

Cochran, J. Nicholas, Ethan G. Geier, Luke W. Bonham, J. Scott Newberry, Michelle D. Amaral, Michelle L. Thompson, Brittany N. Lasseigne, et al. 2020. "Non-Coding and Loss-of-Function Coding Variants in TET2 Are Associated with Multiple Neurodegenerative Diseases." *The American Journal of Human Genetics* 106 (5): 632–45. https://doi.org/10.1016/j.ajhg.2020.03.010.

Dawlaty, Meelad M., Achim Breiling, Thuc Le, Günter Raddatz, M. Inmaculada Barrasa, Albert W. Cheng, Qing Gao, et al. 2013. "Combined Deficiency of Tet1 and Tet2 Causes Epigenetic Abnormalities but Is Compatible with Postnatal Development." *Developmental Cell* 24 (3): 310–23. https://doi.org/10.1016/j.devcel.2012.12.015.

Dawlaty, Meelad M., Kibibi Ganz, Benjamin E. Powell, Yueh-Chiang Hu, Styliani Markoulaki, Albert W. Cheng, Qing Gao, et al. 2011. "Tet1 Is Dispensable for Maintaining Pluripotency and Its Loss Is Compatible with Embryonic and Postnatal Development." *Cell Stem Cell* 9 (2): 166–75. https://doi.org/10.1016/j.stem.2011.07.010.

Della Ragione, Floriana, Marcella Vacca, Salvatore Fioriniello, Giuseppe Pepe, and Maurizio D'Esposito. 2016. "MECP2, a Multi-Talented Modulator of Chromatin Architecture." *Briefings in Functional Genomics*, June, elw023. https://doi.org/10.1093/bfgp/elw023.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. "STAR: Ultrafast Universal RNA-Seq Aligner." *Bioinformatics* 29 (1): 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Doyle, Joseph P., Joseph D. Dougherty, Myriam Heiman, Eric F. Schmidt, Tanya R. Stevens, Guojun Ma, Sujata Bupp, et al. 2008. "Application of a Translational Profiling Approach for the Comparative Analysis of CNS Cell Types." *Cell* 135 (4): 749–62. https://doi.org/10.1016/j.cell.2008.10.029.

Eden, Eran, Roy Navon, Israel Steinfeld, Doron Lipson, and Zohar Yakhini. 2009. "GOrilla: A Tool for Discovery and Visualization of Enriched GO Terms in Ranked Gene Lists." *BMC Bioinformatics* 10 (1): 48. https://doi.org/10.1186/1471-2105-10-48.

Feliciano, Pamela, Amy M. Daniels, LeeAnne Green Snyder, Amy Beaumont, Alexies Camba, Amy Esler, Amanda G. Gulsrud, et al. 2018. "SPARK: A US Cohort of 50,000 Families to Accelerate Autism Research." *Neuron* 97 (3): 488–93. https://doi.org/10.1016/j.neuron.2018.01.015.

Gong, Shiaoching, Martin Doughty, Carroll R. Harbaugh, Alexander Cummins, Mary E. Hatten, Nathaniel Heintz, and Charles R. Gerfen. 2007. "Targeting Cre Recombinase to Specific Neuron Populations with Bacterial Artificial Chromosome Constructs." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 27 (37): 9817–23. https://doi.org/10.1523/JNEUROSCI.2707-07.2007.

Gong, Shiaoching, Chen Zheng, Martin L. Doughty, Kasia Losos, Nicholas Didkovsky, Uta B. Schambra, Norma J. Nowak, et al. 2003. "A Gene Expression Atlas of the Central Nervous System Based on Bacterial Artificial Chromosomes." *Nature* 425 (6961): 917–25. https://doi.org/10.1038/nature02033.

Gu, Tian-Peng, Fan Guo, Hui Yang, Hai-Ping Wu, Gui-Fang Xu, Wei Liu, Zhi-Guo Xie, et al. 2011. "The Role of Tet3 DNA Dioxygenase in Epigenetic Reprogramming by Oocytes." *Nature* 477 (7366): 606–10. https://doi.org/10.1038/nature10443.

Hackett, J. A., R. Sengupta, J. J. Zylicz, K. Murakami, C. Lee, T. A. Down, and M. A. Surani. 2013. "Germline DNA Demethylation Dynamics and Imprint Erasure Through 5-Hydroxymethylcytosine." *Science* 339 (6118): 448–52. https://doi.org/10.1126/science.1229277.

Hashimoto, Hideharu, Samuel Hong, Ashok S. Bhagwat, Xing Zhang, and Xiaodong Cheng. 2012. "Excision of 5-Hydroxymethyluracil and 5-Carboxylcytosine by the Thymine DNA Glycosylase Domain: Its Structural Basis and Implications for Active DNA

Demethylation." *Nucleic Acids Research* 40 (20): 10203–14. https://doi.org/10.1093/nar/gks845.

Hashimoto, Hideharu, Yiwei Liu, Anup K. Upadhyay, Yanqi Chang, Shelley B. Howerton, Paula M. Vertino, Xing Zhang, and Xiaodong Cheng. 2012. "Recognition and Potential Mechanisms for Replication and Erasure of Cytosine Hydroxymethylation." *Nucleic Acids Research* 40 (11): 4841–49. https://doi.org/10.1093/nar/gks155.

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* 333 (6047): 1303–7. https://doi.org/10.1126/science.1210944.

Heiman, Myriam, Anne Schaefer, Shiaoching Gong, Jayms D. Peterson, Michelle Day, Keri E. Ramsey, Mayte Suárez-Fariñas, et al. 2008. "A Translational Profiling Approach for the Molecular Characterization of CNS Cell Types." *Cell* 135 (4): 738–48. https://doi.org/10.1016/j.cell.2008.10.028.

Heinz, Sven, Christopher Benner, Nathanael Spann, Eric Bertolino, Yin C. Lin, Peter Laslo, Jason X. Cheng, Cornelis Murre, Harinder Singh, and Christopher K. Glass. 2010. "Simple Combinations of Lineage-Determining Transcription Factors Prime Cis-Regulatory Elements Required for Macrophage and B Cell Identities." *Molecular Cell* 38 (4): 576–89. https://doi.org/10.1016/j.molcel.2010.05.004.

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

Holliday, R., and G.W. Grigg. 1993. "DNA Methylation and Mutation." *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 285 (1): 61–67. https://doi.org/10.1016/0027-5107(93)90052-H.

Huang, Da Wei, Brad T Sherman, and Richard A Lempicki. 2009. "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.

Huber, Wolfgang, Vincent J Carey, Robert Gentleman, Simon Anders, Marc Carlson, Benilton S Carvalho, Hector Corrada Bravo, et al. 2015. "Orchestrating High-Throughput Genomic Analysis with Bioconductor." *Nature Methods* 12 (2): 115–21. https://doi.org/10.1038/nmeth.3252.

Hwang, Byungjin, Ji Hyun Lee, and Duhee Bang. 2018. "Single-Cell RNA Sequencing Technologies and Bioinformatics Pipelines." *Experimental & Molecular Medicine* 50 (8): 1–14. https://doi.org/10.1038/s12276-018-0071-8.

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* 333 (6047): 1300–1303. https://doi.org/10.1126/science.1210597.

Iyer, Lakshminarayan M., Mamta Tahiliani, Anjana Rao, and L. Aravind. 2009. "Prediction of Novel Families of Enzymes Involved in Oxidative and Other Complex Modifications of Bases in Nucleic Acids." *Cell Cycle* 8 (11): 1698–1710. https://doi.org/10.4161/cc.8.11.8580.

Jeong, Mira, Deqiang Sun, Min Luo, Yun Huang, Grant A. Challen, Benjamin Rodriguez, Xiaotian Zhang, et al. 2014. “Large Conserved Domains of Low DNA Methylation Maintained by Dnmt3a.” *Nature Genetics* 46 (1): 17–23. https://doi.org/10.1038/ng.2836.

Kaas, Garrett A., Chun Zhong, Dawn E. Eason, Daniel L. Ross, Raj V. Vachhani, Guo-li Ming, Jennifer R. King, Hongjun Song, and J. David Sweatt. 2013. “TET1 Controls CNS 5-Methylcytosine Hydroxylation, Active DNA Demethylation, Gene Transcription, and Memory Formation.” *Neuron* 79 (6): 1086–93. https://doi.org/10.1016/j.neuron.2013.08.032.

Katz, David M., Adrian Bird, Monica Coenraads, Steven J. Gray, Debashish U. Menon, Benjamin D. Philpot, and Daniel C. Tarquinio. 2016. “Rett Syndrome: Crossing the Threshold to Clinical Translation.” *Trends in Neurosciences* 39 (2): 100–113. https://doi.org/10.1016/j.tins.2015.12.008.

Kriaucionis, Skirmantas, and Nathaniel Heintz. 2009. “The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain.” *Science* 324 (5929): 929–30. https://doi.org/10.1126/science.1169786.

Krueger, Felix, and Simon R. Andrews. 2011. “Bismark: A Flexible Aligner and Methylation Caller for Bisulfite-Seq Applications.” *Bioinformatics* 27 (11): 1571–72. https://doi.org/10.1093/bioinformatics/btr167.

Langmead, Ben, and Steven L. Salzberg. 2012. “Fast Gapped-Read Alignment with Bowtie 2.” *Nature Methods* 9 (4): 357–59. https://doi.org/10.1038/nmeth.1923.

Li, Heng, Bob Handsaker, Alec Wysoker, Tim Fennell, Jue Ruan, Nils Homer, Gabor Marth, Goncalo Abecasis, and Richard Durbin. 2009. “The Sequence Alignment/Map Format and SAMtools.” *Bioinformatics* 25 (16): 2078–79. https://doi.org/10.1093/bioinformatics/btp352.

Li, Ting, Dehua Yang, Jia Li, Yu Tang, Juan Yang, and Weidong Le. 2015. “Critical Role of Tet3 in Neural Progenitor Cell Maintenance and Terminal Differentiation.” *Molecular Neurobiology* 51 (1): 142–54. https://doi.org/10.1007/s12035-014-8734-5.

Lio, Chan-Wang J., and Anjana Rao. 2019. “TET Enzymes and 5hmC in Adaptive and Innate Immune Systems.” *Frontiers in Immunology* 10 (February): 210. https://doi.org/10.3389/fimmu.2019.00210.

Lio, Chan-Wang J., Vipul Shukla, Daniela Samaniego-Castruita, Edahi González-Avalos, Abhijit Chakraborty, Xiaojing Yue, David G. Schatz, Ferhat Ay, and Anjana Rao. 2019. “TET Enzymes Augment Activation-Induced Deaminase (AID) Expression via 5-Hydroxymethylcytosine Modifications at the *Aicda* Superenhancer.” *Science Immunology* 4 (34): eaau7523. https://doi.org/10.1126/sciimmunol.aau7523.

Lio, Chan-Wang, Jiayuan Zhang, Edahí González-Avalos, Patrick G Hogan, Xing Chang, and Anjana Rao. 2016. “Tet2 and Tet3 Cooperate with B-Lineage Transcription Factors to Regulate DNA Modification and Chromatin Accessibility.” *ELife* 5 (November): e18290. https://doi.org/10.7554/eLife.18290.

Love, Michael I, Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550. https://doi.org/10.1186/s13059-014-0550-8.

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.” *Journal of Biological Chemistry* 286 (41): 35334–38. https://doi.org/10.1074/jbc.C111.284620.

Marshall, Lee L., Bryan A. Killinger, Elizabeth Ensink, Peipei Li, Katie X. Li, Wei Cui, Noah Lubben, et al. 2020. “Epigenomic Analysis of Parkinson’s Disease Neurons Identifies Tet2 Loss as Neuroprotective.” *Nature Neuroscience* 23 (10): 1203–14. https://doi.org/10.1038/s41593-020-0690-y.

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

Meehan, R. R., J. D. Lewis, S. McKay, E. L. Kleiner, and A. P. Bird. 1989. “Identification of a Mammalian Protein That Binds Specifically to DNA Containing Methylated CpGs.” *Cell* 58 (3): 499–507. https://doi.org/10.1016/0092-8674(89)90430-3.

Mellén, Marian, Pinar Ayata, Scott Dewell, Skirmantas Kriaucionis, and Nathaniel Heintz. 2012. “MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System.” *Cell* 151 (7): 1417–30. https://doi.org/10.1016/j.cell.2012.11.022.

Mellén, Marian, Pinar Ayata, and Nathaniel Heintz. 2017. “5-Hydroxymethylcytosine Accumulation in Postmitotic Neurons Results in Functional Demethylation of Expressed Genes.” *Proceedings of the National Academy of Sciences* 114 (37): E7812–21. https://doi.org/10.1073/pnas.1708044114.

Okano, Masaki, Daphne W Bell, Daniel A Haber, and En Li. 1999. “DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development.” *Cell* 99 (3): 247–57. https://doi.org/10.1016/S0092-8674(00)81656-6.

Pastor, William A., L. Aravind, and Anjana Rao. 2013. “TETonic Shift: Biological Roles of TET Proteins in DNA Demethylation and Transcription.” *Nature Reviews Molecular Cell Biology* 14 (6): 341–56. https://doi.org/10.1038/nrm3589.

Patro, Rob, Geet Duggal, Michael I. Love, Rafael A. Irizarry, and Carl Kingsford. 2017. “Salmon Provides Fast and Bias-Aware Quantification of Transcript Expression.” *Nature Methods* 14 (4): 417–19. https://doi.org/10.1038/nmeth.4197.

Ramírez, Fidel, Devon P Ryan, Björn Grüning, Vivek Bhardwaj, Fabian Kilpert, Andreas S Richter, Steffen Heyne, Friederike Dündar, and Thomas Manke. 2016. “DeepTools2: A next Generation Web Server for Deep-Sequencing Data Analysis.” *Nucleic Acids Research* 44 (W1): W160–65. https://doi.org/10.1093/nar/gkw257.

Ramón y Cajal, Santiago. 1992. *Textura del sistema nervioso del hombre y de los vertebrados: estudios sobre el plan estructural y composición histológica de los centros nerviosos adicionados de consideraciones fisiológicas fundadas en los nuevos descubrimientos*. Nachdr. Alicante: Instituto de Neurociencias.

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 (1): 24–26. https://doi.org/10.1038/nbt.1754.

Ross-Innes, Caryn S., Rory Stark, Andrew E. Teschendorff, Kelly A. Holmes, H. Raza Ali, Mark J. Dunning, Gordon D. Brown, et al. 2012. “Differential Oestrogen Receptor Binding Is Associated with Clinical Outcome in Breast Cancer.” *Nature* 481 (7381): 389–93. https://doi.org/10.1038/nature10730.

Rudenko, Andrii, Meelad M. Dawlaty, Jinsoo Seo, Albert W. Cheng, Jia Meng, Thuc Le, Kym F. Faull, Rudolf Jaenisch, and Li-Huei Tsai. 2013. “Tet1 Is Critical for Neuronal Activity-Regulated Gene Expression and Memory Extinction.” *Neuron* 79 (6): 1109–22. https://doi.org/10.1016/j.neuron.2013.08.003.

Sasidharan Nair, Varun, Mi Hye Song, and Kwon Ik Oh. 2016. “Vitamin C Facilitates Demethylation of the *Foxp3* Enhancer in a Tet-Dependent Manner.” *The Journal of Immunology* 196 (5): 2119–31. https://doi.org/10.4049/jimmunol.1502352.

Sathyanesan, Aaron, Joy Zhou, Joseph Scafidi, Detlef H. Heck, Roy V. Sillitoe, and Vittorio Gallo. 2019. “Emerging Connections between Cerebellar Development, Behaviour and Complex Brain Disorders.” *Nature Reviews Neuroscience* 20 (5): 298–313. https://doi.org/10.1038/s41583-019-0152-2.

Schep, Alicia N, Beijing Wu, Jason D Buenrostro, and William J Greenleaf. 2017. “ChromVAR: Inferring Transcription-Factor-Associated Accessibility from Single-Cell Epigenomic Data.” *Nature Methods* 14 (10): 975–78. https://doi.org/10.1038/nmeth.4401.

Schultz, Matthew D., Yupeng He, John W. Whitaker, Manoj Hariharan, Eran A. Mukamel, Danny Leung, Nisha Rajagopal, et al. 2015. “Human Body Epigenome Maps Reveal Noncanonical DNA Methylation Variation.” *Nature* 523 (7559): 212–16. https://doi.org/10.1038/nature14465.

Simon, Jeffrey A., and Robert E. Kingston. 2009. “Mechanisms of Polycomb Gene Silencing: Knowns and Unknowns.” *Nature Reviews Molecular Cell Biology* 10 (10): 697–708. https://doi.org/10.1038/nrm2763.

Smith, Zachary D., and Alexander Meissner. 2013. “DNA Methylation: Roles in Mammalian Development.” *Nature Reviews Genetics* 14 (3): 204–20. https://doi.org/10.1038/nrg3354.

Soneson, Charlotte, Michael I. Love, and Mark D. Robinson. 2015. “Differential Analyses for RNA-Seq: Transcript-Level Estimates Improve Gene-Level Inferences.” *F1000Research* 4 (December): 1521. https://doi.org/10.12688/f1000research.7563.1.

Song, Chun-Xiao, Keith E Szulwach, Ye Fu, Qing Dai, Chengqi Yi, Xuekun Li, Yujing Li, et al. 2011. “Selective Chemical Labeling Reveals the Genome-Wide Distribution of 5-Hydroxymethylcytosine.” *Nature Biotechnology* 29 (1): 68–72. https://doi.org/10.1038/nbt.1732.

Song, Qiang, Benjamin Decato, Elizabeth E. Hong, Meng Zhou, Fang Fang, Jianghan Qu, Tyler Garvin, Michael Kessler, Jun Zhou, and Andrew D. Smith. 2013. “A Reference Methylome Database and Analysis Pipeline to Facilitate Integrative and Comparative Epigenomics.” *PLOS ONE* 8 (12): e81148. https://doi.org/10.1371/journal.pone.0081148.

Spruijt, Cornelia G., Felix Gnerlich, Arne H. Smits, Toni Pfaffeneder, Pascal W.T.C. Jansen, Christina Bauer, Martin Münzel, et al. 2013. “Dynamic Readers for 5-(Hydroxy)Methylcytosine and Its Oxidized Derivatives.” *Cell* 152 (5): 1146–59. https://doi.org/10.1016/j.cell.2013.02.004.

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.

Tahiliani, M., K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, et al. 2009. “Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1.” *Science* 324 (5929): 930–35. https://doi.org/10.1126/science.1170116.

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.

Tsagaratou, Ageliki, Edahí González-Avalos, Sini Rautio, James P Scott-Browne, Susan Togher, William A Pastor, Ellen V Rothenberg, Lukas Chavez, Harri Lähdesmäki, and Anjana Rao. 2017. "TET Proteins Regulate the Lineage Specification and TCR-Mediated Expansion of INKT Cells." *Nature Immunology* 18 (1): 45–53. https://doi.org/10.1038/ni.3630.

Vincent, John J., Yun Huang, Pao-Yang Chen, Suhua Feng, Joseph H. Calvopiña, Kevin Nee, Serena A. Lee, et al. 2013. "Stage-Specific Roles for Tet1 and Tet2 in DNA Demethylation in Primordial Germ Cells." *Cell Stem Cell* 12 (4): 470–78. https://doi.org/10.1016/j.stem.2013.01.016.

Weber, Alain R., Claudia Krawczyk, Adam B. Robertson, Anna Kuśnierczyk, Cathrine B. Vågbø, David Schuermann, Arne Klungland, and Primo Schär. 2016. "Biochemical Reconstitution of TET1–TDG–BER-Dependent Active DNA Demethylation Reveals a Highly Coordinated Mechanism." *Nature Communications* 7 (1): 10806. https://doi.org/10.1038/ncomms10806.

Weng, Yi-Lan, Ran An, Jessica Cassin, Jessica Joseph, Ruifa Mi, Chen Wang, Chun Zhong, et al. 2017. "An Intrinsic Epigenetic Barrier for Functional Axon Regeneration." *Neuron* 94 (2): 337-346.e6. https://doi.org/10.1016/j.neuron.2017.03.034.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the Tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wossidlo, Mark, Toshinobu Nakamura, Konstantin Lepikhov, C. Joana Marques, Valeri Zakhartchenko, Michele Boiani, Julia Arand, Toru Nakano, Wolf Reik, and Jörn Walter. 2011. "5-Hydroxymethylcytosine in the Mammalian Zygote Is Linked with Epigenetic Reprogramming." *Nature Communications* 2 (1): 241. https://doi.org/10.1038/ncomms1240.

Wu, Xiaoji, and Yi Zhang. 2017. "TET-Mediated Active DNA Demethylation: Mechanism, Function and Beyond." *Nature Reviews Genetics* 18 (9): 517–34. https://doi.org/10.1038/nrg.2017.33.

Xu, Xiao, Elitsa I Stoyanova, Agata E Lemiesz, Jie Xing, Deborah C Mash, and Nathaniel Heintz. 2018. "Species and Cell-Type Properties of Classically Defined Human and Rodent Neurons and Glia." *ELife* 7 (October): e37551. https://doi.org/10.7554/eLife.37551.

Xu, Yufei, Chao Xu, Akiko Kato, Wolfram Tempel, Jose Garcia Abreu, Chuanbing Bian, Yeguang Hu, et al. 2012. "Tet3 CXXC Domain and Dioxygenase Activity Cooperatively Regulate Key Genes for Xenopus Eye and Neural Development." *Cell* 151 (6): 1200–1213. https://doi.org/10.1016/j.cell.2012.11.014.

Yamaguchi, Shinpei, Kwonho Hong, Rui Liu, Li Shen, Azusa Inoue, Dinh Diep, Kun Zhang, and Yi Zhang. 2012. "Tet1 Controls Meiosis by Regulating Meiotic Gene Expression." *Nature* 492 (7429): 443–47. https://doi.org/10.1038/nature11709.

Yin, Yimeng, Ekaterina Morgunova, Arttu Jolma, Eevi Kaasinen, Biswajyoti Sahu, Syed Khund-Sayeed, Pratyush K. Das, et al. 2017. "Impact of Cytosine Methylation on DNA Binding Specificities of Human Transcription Factors." *Science* 356 (6337): eaaj2239. https://doi.org/10.1126/science.aaj2239.

Yue, Xiaojing, Chan-Wang J. Lio, Daniela Samaniego-Castruita, Xiang Li, and Anjana Rao. 2019. "Loss of TET2 and TET3 in Regulatory T Cells Unleashes Effector Function." *Nature Communications* 10 (1): 1–14. https://doi.org/10.1038/s41467-019-09541-y.

Zemach, A., I. E. McDaniel, P. Silva, and D. Zilberman. 2010. “Genome-Wide Evolutionary Analysis of Eukaryotic DNA Methylation.” *Science* 328 (5980): 916–19. https://doi.org/10.1126/science.1186366.

Zhang, Liang, Xingyu Lu, Junyan Lu, Haihua Liang, Qing Dai, Guo-Liang Xu, Cheng Luo, Hualiang Jiang, and Chuan He. 2012. “Thymine DNA Glycosylase Specifically Recognizes 5-Carboxylcytosine-Modified DNA.” *Nature Chemical Biology* 8 (4): 328–30. https://doi.org/10.1038/nchembio.914.

Zhang, Yong, Tao Liu, Clifford A Meyer, Jérôme Eeckhoute, David S Johnson, Bradley E Bernstein, Chad Nussbaum, 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.

Zhu, Jian-Kang. 2009. “Active DNA Demethylation Mediated by DNA Glycosylases.” *Annual Review of Genetics* 43 (1): 143–66. https://doi.org/10.1146/annurev-genet-102108-134205.

Zhu, Xiaodong, David Girardo, Eve-Ellen Govek, Keisha John, Marian Mellén, Pablo Tamayo, Jill P. Mesirov, and Mary E. Hatten. 2016. “Role of Tet1/3 Genes and Chromatin Remodeling Genes in Cerebellar Circuit Formation.” *Neuron* 89 (1): 100–112. https://doi.org/10.1016/j.neuron.2015.11.030.