Summer 9-1-2014

Methods and Mechanisms of DNA Methylation in Development and Disease

Maximiliaan Schillebeeckx
Washington University in St. Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology & Biomedical Sciences
Molecular Genetics & Genomics

Dissertation Examination Committee:
Robi Mitra, Chair
James Huettner
Stephen Johnson
Shelly Sakiyama-Elbert
Ting Wang
David Wilson

Methods and Mechanisms of DNA methylation in Development and Disease

By

Maximiliaan A Schillebeeckx

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

August 2014
St. Louis, Missouri

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The pursuit of a PhD is fruitless without the involvement of countless people who provide support, mentorship, and guidance on the multiple facets of research and of life. The people mentioned herein are only those who come to mind in this moment. I apologize in advance for inadvertently omitting anyone that has helped me or been a source of creativity and energy during my time at Washington University in St. Louis. There are many people that deserve thanks; I only hope that I have been gracious enough during our interactions that you recognize the importance you have had in my training as a scientist and in my personal development.

First and foremost, I am incredibly indebted to my mentor, Rob Mitra. I am grateful for having been able to work under his mentorship. Rob held my place in lab during my year abroad in Barcelona when he easily could have accepted much more capable graduate students in my place. I also recognize how fortunate I am to have had a mentor who supported my interests outside of lab. I am grateful that I could be transparent and even share my excitement about The BALSA Group during my time with the organization. His trust in my abilities to get research done while pursuing other interests gave me the confidence to pursue them. Rob has instilled in me a sense of scientific rigor, transparency, and curiosity. He taught me to think big and gave me the courage to ask tough questions of others and of myself. I look up to Rob as a scientist whose creativity, energy, and vision never cease to amaze me. Mostly, I admire Rob's integrity and fairness as an individual.

I am grateful for the openness and support of my thesis committee: Stephen Johnson, David Wilson, Ting Wang, Shelly Sakiyama-Elbert, and James Huettner. Each member has directly contributed to my training as a scientist. Without hesitation, Jim became my advocate and gave me a broader understanding of neuron biology, especially in electrophysiology. Shelly instinctively opened her lab and resources to my scientific endeavors. I'm incredibly grateful for her willingness to collaborate and share the Hb9-puro stem cell line before it was published. Ting formally brought me into the DNA methylation community my second year of grad school when he included me in the final analysis of a high profile publication. I certainly was not deserving of this opportunity at the time and thank him for being a true mentor by making me part of this great work. Dave single-handedly ensured that I would complete my thesis in a reasonable time-frame. Without his collaborations, LCM-RRBS would not have been published

when it did. Dave is one of the most effective science communicators I know; I am incredibly grateful for his selfless mentorship, which has made me into a more well-rounded scientist. Last but not least, I can truthfully say that I would not be a graduate student at Wash U if it were not for Steve. We met before I had even applied. Our conversations during that meeting prompted me to rewrite my entire application to be more in line with Wash U's vision. During my time as a graduate student, Steve was always willing to spend hours carefully troubleshooting or strategizing experiments with me. I hope he knows how much those meetings meant to me. He has been a true champion of my work and one of a few individuals truly invested in me.

I would like to thank members of the Genetics Department for their continued support of me and the many other graduate students in the department. Jim Skeath and Melanie Relich are a big part of why there is such a great sense of community in the Molecular Genetics and Genomic program. I am grateful to have them as colleagues. Vanessa and I met at Jim Skeath's house; he's the reason we are together. I would also like to thank Debbie Peterson, Stephanie Amen, Marilyn Hummert, and especially Patricia Winkler for facilitating my research by handling all administrative aspects of my research.

The past and present members of the Mitra lab have been a source of inspiration throughout my PhD. I am indebted to KT Varley for her mentorship and support while I rotated in the lab. Yue Yun and Lee Tessler were great colleagues; I specifically remember all their help during my thesis proposal. Michael Brooks, Francesco Vallania, and David Mayhew were always willing to provide feedback, brainstorm, and trouble-shoot – the majority of what's done during a thesis. Scott Hidgon is probably the reason why I continued to show up to lab, always ready to provide comic relief and inspiration with neat zebrafish videos as well as scientific and emotional support. I'm glad to have gotten the opportunity to work with Tom Cohen, Justin Melendez, Zongtai Qi, Sumithra Sankararaman, Arjun Bahl, Gabe Bien-Willner, as well as the many rotations students, especially Mingchao Xie. Finally, I want thank Kay Tweedy, Xuhua Chen, and Jess Hoisington-Lopez for their dedication to the Mitra Lab and help with my research; they are pillars of the Mitra Lab and the CGS.

I am grateful for the Center for Genome Sciences and Systems Biology (CGS), and the incredible friendships and collaborations that are a result of this collegiate community. I appreciate the positive energy Jeffrey Gordon has instilled in the center and his unwavering openness to sharing resources and

time. I am especially grateful to members of the Cohen and Gordon lab: being able to look toward them for technical expertise has been invaluable to me. Well beyond their scientific help and support, they are the reason the CGS is fun and productive. I especially want to thank Jamie Kwasnieski, Philip Ahern, Joe Planer, Andy Kau, Shirley McKinney, and Antoniett Barnes. Thank you for your friendships and camaraderie.

I have had the tremendous privilege of being part of The BALSA Group while a PhD student. This organization and the people that are part of it have restored the idealism I once had in undergrad. Having a means to express my passion for science outside of the laboratory has strengthened my curiosity and persistence inside the laboratory. I have learned so much from my colleagues at The BALSA Group and from their friendship, feedback, and energy. I am especially grateful to Amy Peng, Paul Lu, and Brett Maricque who have been honest advocates of me and who have been instrumental in shaping the organization.

So many friends have been pivotal in completing my PhD. I was lucky enough to be part of a tight knit group of Molecular Genetics and Genomics students that became friends even before graduate school started at the interview weekends. They have been an incredible emotional and moral support system that has fostered my personal and professional development. I would not have completed my PhD without Anthony Tubbs, Carla Falasco, Nick Tran, Dave Grotsky, Laura Najemy, Diana Schorry, Kyle Brightman, Beth Tuck, Erica Schoeller, Victor Ridaura, and Andrew Metzger.

My family has been a constant source of inspiration. My father, Marc, has shown me what hard work means and has taught me that nothing is worth doing unless it is done well. My mother, Roselyne, has shown me what patience means and has taught me how to focus on and break down goals in order to accomplish them. My brothers, Titou, Ian, and Esa, have shown me true understanding and have taught me to seize and look for opportunities – not to wait for life to happen. I would not be the person I am without all of them.

My girlfriend, fiancée, wife, and love-of-my-life, Vanessa, has been a pillar in my life since we met at grad school orientation 7 years ago. She's helped me through the ups and downs of research and of life. Her patience, understanding, and unwavering presence have brought me to where I am now. I would

not have overcome significant times of adversity without her and depend on her as my conscious, champion, and critic. This body of work is dedicated to you. Te amo.

ABSTRACT OF THE DISSERTATION

Methods and Mechanisms of DNA methylation in Development and Disease

By

Maximiliaan A Schillebeeckx

Doctor of Philosophy in Biology and Biomedical Sciences

Molecular Genetics and Genomics

Washington University in St. Louis, 2014

Professor Robi D. Mitra, Chair

DNA methylation is a mechanism for long-term transcriptional regulation and is required for normal cellular differentiation. Failure to properly establish or maintain DNA methylation patterns leads to cell dysfunction and diseases such as cancer and neurological disorders. The goal of this thesis is to understand the role of DNA methylation in oncological cellular transformation and in normal development. To achieve this goal, I have developed a novel method for mapping genome-wide DNA methylation patterns and have applied the method to gonadectomy-induced adrenocortical neoplasms and to maturing motor neurons. The novel method, called Laser Capture Microdissected-Reduced Representation Bisulfite Sequencing (LCM-RRBS), accurately and reproducibly profiles genome-wide methylation of DNA extracted from microdissected fresh frozen or formalin-fixed paraffin-embedded tissue samples. Using this method, I find that significant DNA methylation changes, associated with attendant expression changes, occur in transformed adrenocortical cells. My work has also uncovered significant DNA methylation configuration in maturing motor neurons associated with dramatic expression changes. I show that demethylated regions are enriched for known neuron-specific transcription factor binding sites and that genetic disruption of the active demethylation machinery significantly inhibits motor neuron differentiation and maturation. Together, these experiments demonstrate that DNA methylation plays a role in the transformation of normal cells to cancer cells and that DNA methylation is critical to proper motor neuron formation. I conclude that aberrant DNA methylation controls gene expression in gonadectomy-induced adrenocortical neoplasms and that neuron-specific transcription factors could recruit demethylating enzymes to regions that lose DNA methylation in motor neurons upon maturation.

CHAPTER 1: INTRODUCTION

OVERVIEW

The genetic code of life is incredibly static and robust. From one cell, with two genome copies, results a multi-cellular organism made of over 100 distinct cell types (Gilbert 2013). Each cell contains an (almost) exact copy of that original genome present in the first cell, yet can exhibit unique cellular morphologies and functions (Gilbert 2013; Lodish 2013). What's more, scientists have observed that as a fertilized egg develops into a mature embryo, daughter cells of the first, totipotent cell gradually lose the ability to differentiate into any cell type. Indeed, as early cells differentiate toward one of three germ layers, which specialize into different cell types, cells lose the potential to revert to a different germ layer and become committed to the lineage. As a blastocyst develops, these intermediate progenitor cells further differentiate to the terminal states of fully formed, specialized adult cells. These seemingly contradictory observations have long perplexed scientists: How could such a diverse cadre of functionally and morphologically distinct cell types result from genetically identical cells? How does a dividing cell (e.g. liver) know to replicate into the same cell type (another liver cell)?

Conrad Waddington first introduced the concept of an "epigenetic landscape" to describe this phenomenon in which differentiating cells lose the ability to return to a more pluripotent state (Waddington 1940; Gilbert 2012). Waddington's use of "epigenesis" or "epigenetics" referred to the path cells take during development from genotype to phenotype and also incorporated the understanding that neighboring cells influence the differentiation of each other (Waddington 1939). The contributions of neighboring cells (e.g. Shh signaling), or 'extrinsic' factors, have since been expanded to include any environmental effects both *ex vivo* and *in vivo*. Modern definitions of "epigenetics" try to include the notion of non-genetic heritability, traversing either cell divisions or generations, of such factors or phenomenon to evoke the notion of cell identity "memory" (Hemberger et al. 2009).

In my opinion, the term "epigenetics" is highly sensationalized and used to describe observations that go beyond the intention of its original definition, which has led to confusion and over use. For these reasons, I avoid the use of the word "epigenetics," but instead describe my observations and understanding of the literature in the most concrete manner possible. On those occasions I do use the

term “epigenetics,” I broadly refer to any modification of DNA or DNA architecture that does not change the nucleic acid sequence. This work focuses on one of these modifications of DNA, the methylation of cytosine residues, which plays an important role in bringing a cell from “genotype to phenotype” and in determining cellular identity.

DNA METHYLTRANSFERASES CATALYZE THE METHYLATION OF CYTOSINE RESIDUES

The existence of methyl-modified cytosine (5mC) in nature was first discovered in 1925 by Johnson and Coghill as a structural unit of nucleic acids isolated from *tubercle bacillus* (Johnson and Coghill 1925) as anticipated by Wheeler and Johnson, the first to synthesize 5-methyl-cytosine in 1904 (Wheeler and Johnson 1904). More than two decades after Johnson and Coghill’s discovery, G. R. Wyatt showed that 5mC occurred in the nucleic acids of higher animals and plants (Wyatt 1950). Studies would later show that, unlike in plants (Gruenbaum et al. 1981), most animal 5mC occurs at the cytosine of a cytosine-guanidine dinucleotide known as a CpG (Grippo et al. 1968). The only mammalian 5mC occurring outside the CpG context was observed in embryonic stem cells and thought to be lost in adult tissues (Ramsahoye et al. 2000; Meissner et al. 2008) until methylation studies of the developing mouse and human brain revealed the presence of potentially biologically relevant levels of CpH (H = A, C, T) methylation in neuronal cell types (Xie et al. 2012; Lister et al. 2013). The role of non-CpG methylation, however, is still unclear. Mammalian genomes are depleted of CpG dinucleotides due to the propensity of 5mC to deaminate to thymine (Coulondre et al. 1978). This deamination has resulted in genomes containing unmethylated CpG-rich regions, known as CpG Islands, primarily located in gene promoters (Bird et al. 1985). CpGs outside of CGIs are typically methylated.

A family of proteins, known as the DNA methyltransferases (DNMTs), catalyzes the transfer of a methyl group from S-adenosylmethionine to a cytosine residue (Gold et al. 1963; Grippo et al. 1968). Four known DNMT proteins exist in mammals: *Dnmt1* (Bestor et al. 1988), *Dnmt3A/B* (Okano et al. 1998), and *Dnmt3l* (Aapola et al. 2000).

Work by Bessman et al. in 1958 characterizing the function of DNA polymerase showed that the enzyme cannot distinguish between the methylated and unmethylated cytosine nucleotide (Bessman et al. 1958) prompting the possibility for the existence of a methyltransferase responsible for propagating

5mC through DNA replication. *Dnmt1* serves this role due to its high affinity for hemi-methylated, newly synthesized DNA (Gruenbaum et al. 1982; Bestor and Ingram 1983) and is primarily responsible for faithfully copying the parental-strand methylation pattern onto the daughter strand after each round of DNA replication (Stein et al. 1982), though some evidence exists that suggests *Dnmt1* may have some *de novo* methylase activity (Vertino et al. 1996). The localization of *Dnmt1* to DNA replication foci (Leonhardt et al. 1992) and its interaction with PCNA (Chuang et al. 1997), a protein that stabilizes the binding of DNA polymerase to DNA, further strengthen the role it plays in maintaining DNA methylation states through cellular division.

Dntm3a and *Dnmt3b* are responsible for the *de novo* methylation of unmethylated DNA (Okano et al. 1998; Okano et al. 1999) and have both overlapping and disparate DNA sequence affinities. Although somatic tissues show very little expression of *Dnmt3a* or *Dnmt3b*, *Dnmt3a* is ubiquitously expressed throughout the early embryo while *Dnmt3b* expression is specific to the forebrain and eyes (Okano et al. 1999). The primarily role of *Dnmt3b* is to methylate minor satellite repeats of pericentric regions (Okano et al. 1999) while *Dnmt3a* has been shown to methylate all CpGs regardless of genomic context (Shirane et al. 2013). Beyond their role as *de novo* methylases, *Dnmt3a* and *Dnmt3b* seem to play a role in methylation maintenance. Early knockout studies showed that embryonic stem cells lacking *Dnmt3a* and *Dnmt3b* enzymes lose nearly all 5mC over progressive cell divisions, indicating *Dnmt1* is insufficient to fully maintain 5mC (Jackson et al. 2004). The most striking evidence of maintenance capability was shown in *Dnmt1* null colorectal carcinoma cells: even after 300 cell generations the cancer cells retained ~60% of their global 5mC (Rhee et al. 2000). A similar study using the same cell line but shRNA knockdown, however, showed a dramatic loss of global and gene-specific methylation (Robert et al. 2003).

Dnmt3l has no active methylase domain but seems to play a role in ensuring proper methylation of imprinted loci and transposable elements through the interaction with *Dnmt3a* and *Dnmt3b*. The expression of *Dnmt3l* is confined to germ cells where it regulates the methylation status of the maternally imprinted regions *Snrpn*, *Necdin*, *Zfp127*, *Kcnq1ot1*, and *Peg3* (Bourc'his et al. 2001). Furthermore, *Dnmt3l* has been shown to ensure methylation of transposons in male but not female germ cells (Bourc'his and Bestor 2004). *Dnmt3l* directly binds to each of the catalytic domains of *Dnmt3a* and

Dnmt3b, which increases their *de novo* methylation activity and ability to bind DNA (Gowher et al. 2005), and directly binds histone modifications to recruit Dnmt3a to imprinted regions (Ooi et al. 2007). More recently, *Dnmt3l* together with *Dnmt3a* have been shown to be responsible for CpG and non-CpG methylation in oocytes (Shirane et al. 2013)

DNA METHYLATION IS IMPORTANT FOR MAINTAINING CELLULAR IDENTITY

The symmetry of a CpG dinucleotide on the forward and reverse strand provides a natural means for ensuring 5mC states are faithfully inherited across DNA replication. Hence, DNA methylation is thought to be the predominant mechanism by which cells maintain their cellular identity through cellular divisions. The first evidence for a role of 5mC in maintaining cellular identity came two decades after Wyatt's discovery of vertebrae 5mC when Vanyushin et al. showed slight differences in the total 5mC content of DNA between various animal tissues (Vanyushin et al. 1970). Using methylation sensitive and insensitive enzymes, Waalwiik and Flavell showed significant 5mC difference between rabbit sperm, brain, and liver DNA at an intron of the beta-globin gene locus (Waalwijk and Flavell 1978). Ehlrich et al. would definitively demonstrate that global 5mC levels varied between tissue and cell types a few years later (Ehrlich et al. 1982). These studies laid down the foundation for understanding the role of DNA methylation in maintaining and determining cellular identify.

Several recent 5mC mapping studies at base pair resolution have demonstrated that cell types show tissue-specific 5mC and are characterized by unique 5mC profiles that overlap among tissues of similar origin. Varley et al. mapped the genome-wide distribution of 5mC across a diverse collection of 82 human cell lines and tissues and showed that individual cell types have very distinct 5mC profiles. They show *in vitro* propagation of cell lines can cause a divergence from primary tissue 5mC levels (Varley et al. 2013). This observation suggests loci that change 5mC levels in culture are dispensable, since these cultured cells maintain their cellular identity. Ziller et al. mapped the entire 5mC landscape of 30 cell and tissue types (Ziller et al. 2013); they find that roughly 22% of 5mC loci differ among normal cell types and that tissues originating of similar origin share similar 5mC signatures.

The turn of the century saw the ushering in of cell reprogramming and induced pluripotent stem (iPS) cells whereby somatic cells regain the capacity to differentiate into various germ cell types

(Robinton and Daley 2012). The ability to restore pluripotency to somatic cells has provided significant insight into the epigenetic mechanisms that underlie reprogramming and has shown that DNA methylation plays a part in fixing a cell's differentiated state and in conferring cellular memory (Kim et al. 2010; Polo et al. 2010). Kim et al. rigorously showed that the cell type of origin influences the differentiation potential of iPS cells (Kim et al. 2010). They showed that neural progenitor-derived iPS cells (NP-iPSCs) differentiate much less efficiently into cells of hematopoietic lineage than iPS cells derived from bone marrow cells (blood-derived; B-iPSCs), which are responsible for producing hematopoietic cells. The few hematopoietic progeny that were differentiated from NP-iPSCs were reprogrammed to iPS cells (NP-derived-blood-derived; B-NP-iPSCs), which subsequently were able to differentiate to hematopoietic lineages at a very high efficiency. The poor blood forming potential of NP-iPSCs suggests that NPs either have epigenetic marks that restrict the differentiation into blood fates or lack the epigenetic marks that enable blood formation. Furthermore, Kim et al. demonstrated that treatment during differentiation of NP-iPS cells with 5-azacytidine, a DNA methylation inhibitor, and TSA, a histone deacetylase inhibitor, dramatically enhanced hematopoietic lineage formation. Most interestingly, derived iPS cells retained residual DNA methylation that was indicative of their tissue of origin. Polo et al. also showed that iPSCs derived from various cell types exhibited distinguishing DNA methylation signatures representative of their cell type of origin (Polo et al. 2010). They expanded upon this understanding and showed that continuous passaging erased these DNA methylation signatures to attenuate cell of origin effects.

Loss-of-function experiments in mice further support the important role of 5mC and the methyltransferase machinery in maintaining cell identity (Broske et al. 2009; Sen et al. 2010; Dhawan et al. 2011b). A study of the epidermal progenitor population showed that *Dnmt1* is necessary for proper differentiation into keratinocytes and for maintaining a self-renewing state (Sen et al. 2010). *Dnmt1* knockdown with shRNAs in progenitors resulted in induced differentiation and a loss of self-renewal ability. Wild-type DNMT1, but not a catalytically inactive mutant, reversed the observed proliferation defects, suggesting that the methylation of DNA controls proliferation and differentiation induction in epidermal progenitors. Furthermore, differentiation induction also occurred in human muscle progenitors upon *Dnmt1* knockdown.

The functional disruption of *Dnmt1* in mice was shown to abrogate the differentiation of hematopoietic stem cells (HSCs) into myeloerythroid and lymphoid lineages (Broske et al. 2009). Broske et al. show that *Dnmt1* maintains high methylation levels at myeloerythroid-specific genes to repress transcription. Failure to maintain methylation at these genes in HSCs prevents the differentiation into lymphoid progeny. Indeed, *Dnmt1*-depleted HSCs differentiated only down the myeloid lineage suggesting hypermethylation of myeloerythroid genes is necessary for proper differentiation into lymphoid cells. Interesting, the removal of *Dnmt1* from committed lymphoid cells did not disrupt lymphoid identity or maturation, suggesting the maintenance of methylation is only necessary for lymphoid differentiation and not necessary for the maintenance of lymphoid cellular identity.

Most striking is a study demonstrating that pancreatic β cells, which produce insulin, undergo transdifferentiation to glucagon-producing α cells upon the conditional disruption of *Dnmt1* (Dhawan et al. 2011b). Using a genome-wide profiling method, they found that a short region upstream of *Arx*, a gene expressed specifically in α cells, lost methylation in *Dnmt1*-null β cells. Upon demethylation of Arx, β cells began to downregulate β cell markers and upregulate α cell markers. The inhibition of DNA methylation with 5-azacytidine also induced an upregulation of *Arx* de-repression, α cell marker activation, and glucagon production demonstrating that the methylation state of *Arx* determines the cellular identity of pancreatic β cells.

DNA METHYLATION AS A MECHANSIM OF STABLE GENE SILENCING

The mechanistic function of 5mC did not become apparent until 1977 when Christman et al. observed gene expression increases of globin genes in erythroleukemia cells cultured in the presence of L-ethionine (Christman et al. 1977). They showed that L-ethionine resulted in a loss of global 5mC. Using methylation sensitive endonucleases, McGhee and Ginder were the first to interrogate the 5mC levels of a specific locus to understand its correlation with gene expression (McGhee and Ginder 1979). By assessing 5mC and expression levels of the β *globin* gene in normal chicken tissues, they showed tissues that expressed or that had expressed β *globin* had sites at the ends of the gene sequence that were fully unmethylated; the same sites were at least partially methylated in tissues that did not express β *globin*. Desrosiers et al. documented a similar observation in virus producing and non-producing cells

(Desrosiers et al. 1979). Gene transfection experiments further showed that *in vitro* methylation of gene constructs inhibits their expression (Fradin et al. 1982) while many endogenous genes can be activated from a repressed state by treatment with 5-azacytidine, an inducer of demethylation (Jones and Taylor 1980). Global methylation and gene expression profiling would eventually show that the methylation of CpGs upstream of genes results in their down-regulation (Walsh et al. 1998; Bird and Wolffe 1999) and would establish 5mC as a transcription regulator.

The effects of DNA methylation, however, are strongly dependent on genomic context. Although 5mC in gene promoters and enhancers are correlated with gene silencing, gene body methylation is associated with transcriptionally active genes (Hellman and Chess 2007; Ball et al. 2009). Many genome-wide studies have shown poor overall correlations of 5mC with gene expression (Eckhardt et al. 2006; Zilberman et al. 2007; Brenet et al. 2011; Hartung et al. 2012). These inconsistencies are most likely a result of our inability to accurately identify the regulatory regions for all genes, as most can occur very distal to transcription start sites.

The covalent bond attaching the methyl group to the 5` carbon of a cytosine ensures methyl markers are stable and not easily lost. The durable nature of DNA methylation allows for the long-term silencing of genes and other genomic elements within specific cellular contexts. Dormant transposable elements pose a threat to the stability of the genome if reactivated. DNA methylation functions to silence transposable elements and maintain chromosomal integrity (Slotkin and Martienssen 2007) by preventing the reactivation of endoparasitic sequence that cause translocations and gene disruptions (Esteller 2007). DNA methylation is also responsible for silencing the inactive X chromosome during female X inactivation (Mohandas et al. 1981) and for ensuring allele-specific expression from paternal or maternal copies at imprinted regions (Reik et al. 1987; Sapienza et al. 1987).

The mechanism by which DNA methylation represses transcription can be direct, by excluding the binding of proteins that affect transcription (Watt and Molloy 1988), or indirect, by recruiting methyl-CpG-binding proteins and their associated repressive chromatin remodelers (Robertson 2005).
The chromatin boundary element binding protein, CTCF, acts to insulate the effects of enhancer regions from gene promoters by binding between these genomic elements. Its DNA binding sequence contains a CpG which when methylated inhibits (Ohlsson et al. 2001) CTCF from binding DNA. The imprinted gene

H19 is repressed on the paternal chromosome through the methylation of an imprinted-control region 2 kilobases from its transcription start site. The same imprinted-control region is unmethylated on the maternal chromosome where CTCF binds to block the effects of a *H19* enhancer on the *Igf2* gene and maternally repress the *Igf2* gene (Hark et al. 2000). Other methylation sensitive and insensitive binding proteins abound, but their role in regulating expression is unclear (Tate and Bird 1993). While CTCF is inhibited by 5mC, several proteins explicitly bind 5mC. *Mecp-1* (now known as *Mecp2*) was the first protein shown to bind specifically to methylated DNA sequences (Boyes and Bird 1991), but several others have since been described (Wade 2001). Methyl-binding proteins mediate gene silencing by recruiting histone-modifying factors that shift the chromatin architecture to a closed, inaccessible, and, therefore, repressed state (Jones et al. 1998; Nan et al. 1998). MECP2, for example, recruits the arginine methyltransferase, PRMT6, among others, to add repressive chromatin marks and remove activating chromatin marks at the *Arx* locus to repress transcription (Dhawan et al. 2011b).

DNA METHYLATION AND NEURONAL DEVELOPMENT AND FUNCTION

The past decade has accumulated significant insights into the role DNA methylation plays in neuronal development and postnatal function. Lister et al have shown that the mouse and human brain undergo significant methylation reconfiguring during development and highlight the potential importance of non-CpG methylation (Lister et al. 2013). They show that the fetal cortex is practically devoid of CpH methylation but rapidly gains global methylation at CpH dinucleotides shortly after birth during a time of high synaptogenesis. CpH methylation in gene bodies was inversely correlated with gene expression which is consistent with the role of 5mC in repressing transcription. Large, non-centromeric megabase regions of the genome, however, were resistant to CpH methylation gains and contained genes that encode receptors required for sensory neuron function and immune function. This study suggests neurons of the brain undergo significant gene repression via 5mC gains in the CpH context between fetal and adult stages. Several studies characterizing 5mC levels show that the human cerebral cortex undergoes significant 5mC changes throughout an individual's lifespan showing gradual as well as sharp gains in 5mC levels at various promoters (Siegmund et al. 2007). Many 5mC changes were associated with transcriptional decline, though the consequences, if any, of these changes are not clear (Hernandez

et al. 2011). Contrary to these studies, Lister et al showed CpG and CpH methylation declines with age post-adolescence. The discordance of these studies most likely lies in the fact that Lister et al interrogated the whole genome while the previous studies were much less comprehensive.

The adult brain can undergo significant neurogenesis in response to external stimuli resulting in advanced structural plasticity. The methylation of DNA plays an important role in this observed plasticity, including in synaptic formation (Levenson et al. 2006), learning and memory (Day and Sweatt 2010; Miller et al. 2010; Zovkic et al. 2013), emotional behavior (Lubin et al. 2008; LaPlant et al. 2010), adult neurogenesis (Ma et al. 2009), and age-related cognitive decline (Oliveira et al. 2012). Most interestingly, is the role of 5mC in neural plasticity and neurogenesis, as various studies have shown that neuronal activity leads to significant gains and losses of 5mC in the brain (Nelson et al. 2008; Miller et al. 2010). For example, the promoter of *Bdnf*, a gene important for adult neurogenesis, is regulated in an activity-dependent manner (Martinowich et al. 2003; Ma et al. 2009). Upon synchronized electro-convulsion of the hippocampus, the *Bdnf* promoter is demethylated by *Gadd45b* within 4 hours, which causes a release of the MECP2 repressor complex and increased expression of *Bdnf*. The study further shows that *Gadd45b* is essential for neural progenitor proliferation and activity-induced dendritic development (Ma et al. 2009). Global profiling studies have revealed the plasticity of 5mC in the brain across several genes (Guo et al. 2011a). Persistent activity in neurons, as occurs during electroconvulsive stimulation or exercise, leads to subtle yet significant 5mC losses at *Per2*, *Crebbp*, and *Grip1* and also 5mC gains at *Zfhx2* and *Ccdc44*. Most of these 5mC changes occur at exonic and intronic regions, which may explain why these 5mC alternations do not correlate well with gene expression. These studies showing that significant 5mC configuration occurs after neuronal activity are helping to expel the belief that DNA methylation is highly stable in terminally differentiated cells.

Glial cells (i.e. astrocytes, oligodendrocytes, and microglia) and neurons are both derived from the same neuron progenitor cell population. Proper DNA methylation establishment seems to play an important role in the fate determination of neural progenitor cells into neurons and glial cells. Neuron and non-neuron cells display unique global methylation profiles, which ensures the regulation of neuronal- and glial-specific transcription (Iwamoto et al. 2011). Glial cells support the viability of neurons and therefore arise after the formation of neurons from the same neuronal progenitor cells (NPCs) (Qian et al. 2000).

Indeed, during development, NPCs gradually acquire competence for gliogenesis (Sauvageot and Stiles 2002). Once NPCs are gliogenic, they simultaneously lose the ability to differentiate into neurons, suggesting the existence of a neurogenic to gliogenic switch that occurs during central nervous system development. Indeed, neurons appear at embryo day 11.5 (E11.5) while astrocytes begin to appear at E14.5. This switch in differentiation ability is regulated by 5mC. Before E14.5, the promoter of *Gfap*, an astrocyte marker, is methylated at a *Stat3* binding site in neuroepithelial (progenitor) cells (Takizawa et al. 2001). At E14.5, neuroprogenitor cells lose methylation at this site which leads to *Stat3* binding, *Gfap* activation, and astrocyte differentiation. Studies of conditional *Dnmt1* knockout mice further validate the role of 5mC in neuron/glial differentiation. Fan et al. show that *Dnmt1*-deficient NPCs are marked by hypomethylation at the *Gfap* promoter and spontaneously differentiate into glial cells (Fan et al. 2005). Upon *Dnmt1* inactivation, glia-associated transcription factors, but not neuronal-specific genes are activated.

Direct experimental findings have shown that the DNA methyltransferases *Dnmt1* and *Dnmt3a* are important for learning, memory, and synaptic plasticity. Conditional mutant mice lacking *Dnmt3a* in the entire central nervous system develop normally but have significant neuromuscular defects, reduced number of motor neurons in the brain stem, and shortened life spans (Nguyen et al. 2007). Because the knockout induction occurrs in neuron progenitor cells, these experiments highlight the importance of *Dnmt3a* in neuron differentiation. Exclusive conditional knockout of *Dnmt1* and *Dnmt3a* in post-mitotic neurons of mice results in learning and memory deficits (Feng et al. 2010). Here, *Dnmt* gene deletion only occurs in mature, post-mitotic neurons of the central nervous system allowing researchers to understand their role in mature neuron function. *Dnmt1* and *Dnmt3a* single knockout mice did not show an appreciable phenotype while double knockout animals had normal lifespans but smaller hippocampi. Furthermore, double knockout mice showed an appreciable loss of neuronal DNA methylation, induction of immune genes, and impaired synaptic plasticity. As we continue to profile the methylation and transcriptional landscapes of neurons under various conditions, the role of 5mC in neuron function will only become more evident. Furthermore, the reversibility and ability to regulate gene expression makes DNA methylation an attractive target for therapies that may be able to activate aberrant gene silencing or silence aberrant expression profiles.

DNA METHYLATION IN HUMAN HEALTH

The importance of DNA methylation in development has been well describing using transgenic and knockout mice. Deletion of *Dnmt1* is lethal to embryonic development (Li et al. 1992); *Dnmt3b*-deficient mice do not reach full term while mice lacking *Dnmt3a* die within 4 weeks of birth (Okano et al. 1999). *Dnmt3l*-null mice are viable and develop normally (Bourc'his et al. 2001); The heterozygous offspring of female mice lacking *Dnmt3l*, however, die before the mid-gestation period from the biallelic expression of imprinted genes normally methylated and silenced on the allele of maternal origin while male mice lacking *Dnmt3l* are sterile due to a lack of germ cells in the adult. The importance of DNA methylation is further emphasized by the growing number of human diseases that are known to occur when 5mC is not properly established and/or maintained (Robertson 2005; Jakovcevski and Akbarian 2012). Indeed, aberrant DNA methylation has been broadly implicated in cancer and in numerous neurological diseases.

Cancer

DNA methylation was first implicated in cancer by Feinberg and Vogelstein in 1983 when they observed significant 5mC differences at specific loci between normal and diseased tissues of four cancer types (Feinberg and Vogelstein 1983). They also found the lowest level of 5mC in the metastasis of a lung tumor, first demonstrating that cancer cells lose methylation as they progress. The aberrant methylation patterns observed in cancers have since been well documented (Portela and Esteller 2010; Sharma et al. 2010; Taby and Issa 2010). Tumors are often characterized by global hypomethylation (Goelz et al. 1985; Wilson et al. 2007) as well as by locus-specific hypermethylation of genes involved in the main cellular pathways (Portela and Esteller 2010). The global hypomethylation, which can occur in large megabase regions and unmethylates repetitive regions (Hansen et al. 2011), is thought to destabilize the genome and result in transposon-mediated rearrangements. Loss of methylation also causes the aberrant activation of growth promoting genes (Wilson et al. 2007) and the loss of imprinting (Reik and Lewis 2005). Hypermethylation primarily at CGIs of promoters, on the other hand, contributes to tumorigenesis by silencing tumor suppressor and DNA repair genes and could thus serve as the second hit in Knudson’s two-hit model (Suzuki et al. 2004; Weinberg 2007). Roughly 5-10% of CGIs that

are constitutively unmethylated in adult tissues are hypermethylated in cancer (Bird 2002; Weinberg 2007). The aberrant methylation of normally unmethylated tumor-suppressor genes can causes loss of heterozygosity and uncontrolled cell division and proliferation in countless cancers (Jones and Baylin 2002; Chen et al. 2003). Today, hypermethylation events are being used as biomarkers to screen for colorectal cancer, among others (Lofton-Day et al. 2008). Abnormal 5mC in conjunction with disruptions in histone modifications may be key initiating events in some forms of cancer (Feinberg et al. 2006a).

The onset of whole genome and exome sequencing efforts have revealed the presence of mutations in chromatin modifying proteins as well as DNA methyltransferases across many tumor types suggesting the aforementioned methylation aberrations could be downstream of genetic abnormalities. To date, over 15% of hematopoietic and lymphoid cancers catalogued in the COSMIC database show mutations in the *de novo* methyltransferase *DNMT3A*, with enrichment in patients with *de novo* acute myeloid leukemia (AML) (Ley et al. 2010). Mutations in *DNMT3A* were correlated with poor patient survival, suggesting *DNMT3A* is involved in AML disease progression. Despite most mutations occurring in the methylase domain of *DNMT3A*, genome-wide 5mC analysis of AML genomes with and without *DNMT3A* mutations did not reveal an appreciable difference in 5mC between the two samples. Furthermore, the few differences that were observed did not result in an expression change of the nearest genes. These observations suggest that the function of *DNMT3A* in cancer progression may not be due to altered DNA methylation, but rather related to its ability to bind DNA and recruit histone modifiers (Ley et al. 2010). Few mutations are observed in *DNMT3B* for any cancer type, highlighting the primarily role of *DNMT3B* in early development. Similarly, very few mutations have been observed in *DNMT3L* of tumors catalogues in the COSMIC database. Over 50% of lung cancers, however, show copy number variations of *DNMT1* and most non-small cell lung cancer tumors show an over-expression of DNMT1, DNMT3A, and DNMT3B protein and show increases in 5mC at tumor suppressor genes (Lin et al. 2007).

The observation that pharmacological DNA methylation inhibitors, such as 5-azacytidine, results in gene activation (Venolia et al. 1982) demonstrated that 5mC is reversible and has since ushered in DNA methylation inhibitors as potent cancer therapies. Several DNA methyltransferase inhibitors have been approved in the past decade by the FDA for the treatment of patients with AML or myelodysplastic syndrome (Kaminskas et al. 2005). The response rates to the drugs, however, are low and highly variable

among patients. Furthermore, it is unclear whether the beneficial effects of treatment are associated with DNA methylation (Silverman et al. 2002; Figueroa et al. 2009).

Neurological Diseases

DNA methylation has been strongly implicated in diseases affecting the nervous system either directly, by the aberrant methylation of DNA (Kumari and Usdin 2009; Buiting 2010), or indirectly, through functional loss of the DNA methyltransferases machinery (Jin et al. 2008; Chestnut et al. 2011; de Greef et al. 2011; Klein et al. 2011) or through the disruption of methyl-binding proteins (Amir et al. 1999). Fragile X Syndrome, a form of mental retardation, is caused by inactivation of the *FMR1* gene through aberrant methylation of its promoter sequence. Aberrant methylation is also responsible for loss of imprinting in Prader-Willi Syndrome and Angelman Syndrome, which disrupts cognitive development in childhood. Mutations in *DNMT3B* and *DNMT1* and the resultant disruption of proper methylation is the cause of two rare neuropathies Immunodeficiency, Centromere instability and Facial anomalies Syndrome 1/2 (ICF1/2) and Hereditary Sensory and Autonomic Neuropathy type 1, respectively, while all three methyltransferases have recently been implicated in Amyotrophic Lateral Sclerosis (ALS), the motor neuron degenerative disease. Finally, Rett Syndrome, one of the most studied neurological diseases, is caused by the disruption of the methyl-CpG-binding protein, MECP2, and characterized by early-stage cognitive decline.

Several neurological diseases are directly caused by the aberrant loss or gain of 5mC at a single locus including repeat-instability disease (Kumari and Usdin 2009) and imprinting disorders (Robertson 2005). Fragile X Syndrome, a common form of mental retardation (Martin and Bell 1943), is caused when the CGG trinucleotide-repeat within the 5`-UTR of the *FMR1* gene extends beyond 200 repeats. At this length, the CGG-repeat becomes methylated and the *FMR1* gene is silenced leading to improper synthesis of neuron-specific proteins (Penagarikano et al. 2007). Recent work has shown that repression occurs because *FMR1* mRNA hybridizes to the complementary CGG-repeat portion of the *FMR1* gene (Colak et al. 2014). Future work will reveal how the DNA methyltransferase machinery is recruited to the RNA·DNA duplex and whether the hypermethylation can be permanently reversed. Two forms of mental impairment can occur when a single imprinted locus is not properly methylated and regulated. Prader-

Willi Syndrome and Angelman Syndrome are caused by a loss of imprinting either through chromosomal loss or by aberrant methylation at the 15q11q13 chromosomal region (Buiting 2010). Prader-Willi occurs when imprinting is lost on the paternal allele while Angelman occurs when imprinting is lost from the maternal allele (Glenn et al. 1993).

Hereditary sensory and autonomic neuropathy type 1 (HSAN1) is the only disease known to be caused by mutations in the DNA targeting domain of *DNMT1* (Klein et al. 2011). Patients with HSAN1 suffer from late onset dementia, hearing loss, and loss of pain sensation due to peripheral and central neuron degeneration. The disruption of this domain results in premature protein degradation, reduced methyltransferase activity, and impaired heterochromatin binding during the G2 cell cycle phase which leads to a moderate global loss of methylation, primarily in repetitive satellite elements, and modest site-specific hypermethylation (Klein et al. 2011). Although *DNMT1* is primarily considered to only function in dividing cells, these studies ascribe a novel function to DNMT1 in post-mitotic neurons, namely in ensuring the proper methylation of pericentric and other condensed portions of the genome (Easwaran et al. 2004).

The ICF syndromes affect multiple organs and are characterized by mental retardation and defective brain development. ICF1 is caused by partial loss-of-function mutations in the catalytic domain of *DNMT3B* and the attendant loss of 5mC at pericentric repeats and gene promoters, which leads to dysregulation of genes involved in neurogenesis, immune function, and development (Jin et al. 2008). Interestingly, patients with ICF2, a syndrome with identical pathologies as ICF1, have completely functional DNA methyltransferase proteins but lack functional copies of the putative transcriptional repressor, *ZBTB24* (de Greef et al. 2011). Hypomethylation defects are also present in ICF2 patients but occur at slightly different pericentric satellites, which implicates *ZBTB24* in regulating DNA methylation.

ALS is one of a handful of neurodegenerative diseases that results as a consequence of motor neuron axonal degradation. Unlike the previously discussed neurological disorders, ALS is an affliction of the upper and lower somatic motor neurons located in the spinal cord that enervate skeletal muscles to control movement and locomotion. Although less well established, DNA methylation may be associated with ALS. One recent study showed that DNA methyltransferases play a role in regulating apoptosis in motor neurons, which suggests a novel function for the methyltransferase machinery (Chestnut et al.

2011). They show that *Dnmt3a* mediates apoptosis in chemically or physically challenged motor neurons through interactions with the mitochondria and through increasing genomic and mitochondrion 5mC. Inhibition of the methyltransferases prevents the induction of apoptosis demonstrating their role in maintaining motor neuron viability. Remarkably, motor neurons affected by ALS showed a significant overexpression of *Dnmt3a* as well as hypermethylation of genomic and mitochondrion DNA. This study alludes to the possibility of treating ALS with methyltransferase inhibitors.

The early onset autism spectrum disorder, Rett Syndrome, is caused by dysfunction of the aminergic neurons of the brainstem. The X-linked gene methyl-CpG-binding protein, *MECP2*, is responsible for regulating a cadre of genes involved in neurogenesis and neuron function. *MECP2* has been shown to act by binding 5mC and recruiting chromatin-modifying factors that reorganize the chromatin into a closed and repressive state (Amir et al. 1999) although various studies have suggested an activating role of *MECP2* (Li et al. 2011). Mutations in *MECP2* preclude effective binding to methylated cytosines and recruitment of these repressive factors (Chahrour et al. 2008). Encouragingly, the restoration of *Mecp2* in the post-mitotic neurons of *Mecp2*-deficient mice recovers most Rett syndrome symptoms (Guy et al. 2007) demonstrating that the lack of Mecp2 does not irreversibly damage neurons and that Rett is not strictly a neurodevelopment disorder.

The occurrence of DNA methylation-associated neurological diseases highlights an under-appreciated role of DNA methylation, namely in ensuring proper function of post-mitotic cells. Previously, DNA methylation was only thought of in the context of cell division and early development. These diseases – and the ability to recover normal function in the case of Rett Syndrome – have shifted the notion of 5mC as a static, permanent mark, important only in development, to a dynamic mark of significant relevance to the human postnatal and adult brain. As we continue to explore the role of 5mC in post-mitotic neurons, we will undoubtedly uncover new regulatory functions of the DNA methylation machinery and associated binding partners.

TECHNOLOGICAL CHALLENGES AND LIMITATIONS FOR MAPPING DNA METHYLATION

Current technologies that measure DNA methylation are limited by cost, sample throughput, the requirement for a large amount of starting material, and the inability to analyze DNA methylation of

complex tissues. Because PCR amplification erases DNA methylation patterns, a methylation-dependent treatment is required before methylation readout. The two main approaches to determining methylation patterns or profiles are affinity enrichment and bisulfite treatment. Where previous technologies limited DNA methylation analysis to a single locus, the introduction of microarray hybridization experiments and next-generation sequencing has permitted the profiling of 5mC at the genomic scale (Laird 2010a). Many of the earliest techniques used to measure DNA methylation have been supplanted by next-generation sequencing approaches (Esteller 2007). Today, discovery and profiling studies use genome-wide approaches while validation and targeted experiments use locus-specific approaches to determine the methylation status of the genome in disease and development.

Genome-wide approaches

Current genome-wide sequencing-based approaches for DNA methylation analysis vary in resolution, coverage, and accuracy (Bock et al. 2010; Harris et al. 2010). Affinity enrichment-based techniques (e.g. MeDIP-seq, MethylCap-seq, and MBD-seq) can theoretically cover 100% of the genome but at a 100-1,000 base pair resolution. The actual 1x and 10x coverage of affinity enrichment published experiments is about 60-67% and 10-20% respectively, which indicates that increased coverage would require significant sequencing to reach saturation (Beck 2010). Furthermore, 70-80% of reads of whole genome sequencing provide no information of CpG methylation due to the non-uniform distribution of CpG dinucleotides (Ziller et al. 2013). Recently developed algorithms when coupled with the methylation-depletion technique MRE-seq have increased the resolution of these tools to CpG resolution (Stevens et al. 2013). The benefit of these methods is the unbiased representation of the genome to include inter- and intragenic regions, promoters, and repetitive regions. An alternative whole genome approach to methyl-affinity enrichment is whole genome bisulfite shotgun sequencing (MethylC-seq)(Lister et al. 2009). The treatment of DNA with sodium bisulfite transforms the epigenetic methylation mark into a genetic mark by only converting unmethylated Cs to Ts (by uracil)(Frommer et al. 1992) and can therefore be translated using sequencing. Bisulfite approaches have quickly become the gold standard for methylation analysis because of its single base pair resolution, independence of CpG methylation status,

and ability to interrogate almost any region in the genome. The most inclusive platform, MethylC-seq, however costs an estimated $30,000 per sample (personal calculation).

A more economical alternative to MethylC-seq is reduced representation bisulfite sequencing (RRBS), which capitalizes on the non-uniform distribution of CpGs across the genome (Meissner et al. 2008). Most CpGs are concentrated in CpG Islands (CGI). CGI appear in 60% of promoters and often show aberrant methylation in cancer (Jones and Baylin 2007). RRBS enriches for these regions with high CpG content, so that, instead of covering the entire genome, RRBS effectively reduces the portion of the genome that is sequenced to ~10% (Bock et al. 2010). Furthermore, when detecting differentially methylated regions (DMR) between two embryonic stem cell lines, RRBS was able to detect 1,000 DMRs with only 15 million reads per sample whereas MeDIP-seq required more than 30 million reads to detect 1,000 DMRs (Bock et al. 2010). RRBS is an attractive choice for analyzing promoter and CGI methylation, which play an important role in regulating genes in human cancers (Jones and Baylin 2007).

Until recently, genome-scale DNAme profiling technologies requires a significant amount of starting DNA. Targeted padlock probes require >200 ng of bisulfite treated DNA (Deng et al. 2009); MeDIP-seq has been optimized for 30-120 ng (Borgel et al. 2010); and MethylC-seq requires several micrograms of DNA (Lister et al. 2009). New innovations in Illumina library preparations using transposon-mediated adapter insertions, however, have allowed for whole genome bisulfite sequencing on 10 ng of DNA (Adey and Shendure 2012). At present, RRBS can be reproducibly performed on <1 ng of purified input DNA (Smith et al. 2012). Furthermore, LCM-RRBS, presented in this thesis, functions on <2 ng of DNA extracted from laser capture microdissected tissues (Schillebeeckx et al. 2013).

Locus-specific approaches

Low-throughput validation assays are common for interrogating the 5mC status of a limited number of genomic loci. Whole genome or genome-wide discovery studies identify differentially methylated regions, which must be validated independently by orthogonal methods. Various locus-specific methods exist that vary in complexity, labor, quantitative resolution, and number of CpGs interrogated. The most useful and most commonly used techniques fall into three categories:

endonuclease techniques, methylation-specific PCR techniques, and bisulfite sequencing techniques. Each category has many variations of the basic assay; the basic principals are described here.

Methods leveraging the specificity of endonucleases are able to interrogate the qualitative level of 5mC at one specific CpG or CpH (Xiong and Laird 1997). In this method, broadly referred to as COBRA, bisulfite-treated DNA is amplified using primers specific to converted DNA. The DNA is then digested with the endonucleases, *Bst*UI and *Taq*I, which contain a converted or unconverted CG sequence in their binding sites. Methylated DNA would prevent bisulfite conversion and therefore eliminate the binding site whereby preventing digestion, whereas unmethylated DNA would allow digestion. Various methods employ different endonucleases depending on the sequence of the site of interest. The HELP assay does not rely on bisulfite conversion (Khulan et al. 2006). Instead, the methylation sensitive and insensitive enzyme, *HpaII* and *MspI*, are used to digest DNA. The *HpaII*-digested DNA fraction represents unmethylated DNA while the *MspI*-digested fraction serves as the background, "input" control. Digested DNA can either be amplified with site-specific primers or ligated with universal oligonucleotides and hybrized to microarrays. COBRA provides a fairly quick qualitative assessment of 5mC at one or two specific sites, but is limited by interrogating 5mC at sites with an endonuclease's sequence recognition motif.

Methylation-specific PCR techniques also rely on bisulfite conversion to translate 5mC into a nucleic acid mark. One such assay, called MethylLight, is quantitative and highly sensitive, capable of detecting methylated alleles in the presence of 10,000-fold excess unmethylated alleles (Eads et al. 2000). In the MethylLight assay, genomic DNA is first bisulfite treated to create methylation-dependent sequence differences. Quantitative PCR is next performed using primers that overlap CpG regions and only anneal if methylated (unconverted). The PCR reaction can also be done under non-quantitative conditions and visualized on an agarose gel to qualitatively assess methylation levels. Like restriction-based assays, methylation-specific PCR only interrogates one CpG per reaction, which significantly limits the assay's throughput and utility.

The most common locus-specific methylation mapping method, bisulfite-specific PCR, consists of bisulfite conversion, PCR amplification, and sequencing (Sasaki et al. 2003). Bisulfite conversion of DNA can result in the creation of abasic DNA backbones resulting in brittle DNA that easily degrades or

sheers. Large amplicons are, therefore, difficult to amplify by bisulfite-specific PCR. Nonetheless, 600-1,000 base pair amplicons can readily be amplified from bisulfite converted DNA using primers designed to anneal to converted DNA. Here, the primers do not contain any CpG sequence which allows for unbiased, methylation-agnostic amplification. Amplified products can be sequenced using next-gen (e.g. Illumina) sequencing, Sanger sequencing, or pyrosequencing. For Illumina sequencing, primers are typically designed with Illumina adapters in the tails; amplicons can also be sequenced using standard library preparation methods. Percent methylation is calculated by determining the fraction of total reads that are methylated. Sanger sequencing requires that amplicons be ligated into a vector and individual clones sequenced. Because each clone represents an individual molecule, 10-15 clones must be sequenced to accurately quantify the percent methylation. Finally, if a biotinylated primer is used during the PCR amplification, the amplicon can be sequenced using a pyrosequencing machine. Here, additional primers designed roughly 30 base pairs from a CpG are annealed to complete the sequencing reaction. Because pyrosequencing is inherently quantitative, percent methylation can be directly calculated from the reaction. Successful bisulfite-specific PCR on a single locus has been performed on as little as 8 cells collected from an embryo and laser capture microdissected tissue samples. (Millar et al. 2002). Bisulfite-specific PCR is robust, sensitive, and allows for the interrogation of all CpGs within an individual amplicon.

SCOPE OF THESIS WORK

Although often considered homogenous, tumors consist of histologically diverse tissues (Weinberg 2007) that display intratumor heterogeneity in gene expression, genotype, and metastatic and proliferative potential (Fidler et al. 1979; Heppner 1984). This heterogeneity impedes the investigation and treatment of tumors because tissue samples may not be representative of the entire tumor (Michor and Polyak 2010). Therefore, the ability to profile methylation genome-wide in a small number of cells is crucial to analyzing 5mC in complex tissues like tumors. Such a tool would enable the use of laser capture microdissection to separate heterogeneous cell types from one another. This thesis describes a novel method, called Laser Capture Microdissection-Reduced Representation Bisulfite Sequencing (LCM-RRBS), that accurately maps genome-wide 5mC of cells isolated using laser capture microdissection.

The method was applied to understand what 5mC changes, if any, occur in gonadectomy-induced adrenocortical neoplasias in mice. Furthermore, LCM-RRBS was applied to pure motor neuron cultures at early and late stages to elucidate the role of 5mC in motor neuron maturation.

Chapter 2 addresses the need for a novel technique that can map 5mC genome-wide of very few cells. I show that LCM-RRBS accurately and reproducibly profiles genome-wide methylation of DNA extracted from microdissected fresh frozen or formalin-fixed paraffin-embedded tissue samples. The LCM-RRBS method adapts Illumina library preparation techniques to minimize sample loss and to pool multiple samples together and uses laser capture microdissection (LCM) to isolate specific populations of cells. I applied LCM-RRBS to character the methylation of adrenocortical neoplasias and was able to map genome-wide 5mC of isolated neoplastic and adjacent normal tissue using less than 10 mm^2 of fresh frozen tissue. In one multiplexed experiment, we interrogated the methylation status of >13,000 promoters and >13,000 CpG Islands of 6 tissue samples (3 neoplastic and 3 normal representing 3 different mice) across an average of >800,000 CpGs per sample. We identified 37 and 8 promoters of genes implicated in adrenal and gonadal function that showed significant loss or gain of 5mC, respectively, in neoplastic relative to normal tissues. This chapter is the proof-of-concept for LCM-RRBS and demonstrates the method's utility.

Chapter 3 is a follow-up to the study described in Chapter 2. We hypothesized that adrenocortical neoplasms that result from gonadectomy develop a gonadal phenotype. To test this hypothesis, we took several approaches. First, we show that *Igfbp6* and *Foxs1*, which are differentially methylated, are also differentially expressed at the transcript and protein level between neoplastic and normal adrenal tissues. Second, we applied transcriptome profiling to identify three novel gonadal markers that were upregulated in adrenocortical neoplasms. Finally, enrichment analysis demonstrated that genes upregulated in the adrenal glands of gonadectomized mice were more likely to be highly expressed in ovary or testis but not in the non-steroidogenic tissue brain. Furthermore, genes downregulated in the adrenal glands of gonadectomized mice were more likely to be highly expressed in the normal adrenal tissue. Together, these findings suggest that adrenal neoplasms exhibit mixed characteristics of male and female gonadal somatic cells and, therefore, represent a possible example of cell type transformation involving 5mC configuration.

In Chapter 4, I describe the 5mC profiles of early and late stage maturing motor neurons using LCM-RRBS. The Hb9-puro RW4 mouse embryonic stem cells have a puromycin transgene driven by the *Hb9* enhancer. *Hb9* is exclusively expressed in motor neuron precursors and early differentiated motor neurons allowing for the selection of pure populations of motor neurons in culture. We find that motor neurons lose and gain 5mC at 2,894 and 1,147 CpGs, respectively, between 24 hours and four days after selection with puromycin. Many of the hypomethylated CpGs were near genes important for motor neuron function that showed a gain in expression over time. I show that hypomethylated regions are enriched for known motor neuron-specific transcription factors like *Isl1*, *Lhx3*, *Phox2a*, and *Tbx20*. Furthermore, we show that motor neurons gain 5-hydroxymethylation (5hmC) as they mature. Because Hb9-expressing cells are post-mitotic, this configuration in 5mC is due to active pathways suggesting a role for *Dnmt3a, Dnmt3b*, and the Tet family of proteins. To test this hypothesis, we derived *Tet1*, *Tet2*, and *Tet3* single, double, and triple knockout Hb9-puro lines using the CRISPR/Cas9 genome engineering system. We find that individual *Tet* genes are dispensable for motor neuron differentiation and maturation and, therefore, play redundant roles. Knockout of *Tet1* and *Tet2* together significantly reduced differentiation potential but resulted in mature motor neurons and normal 5hmC gains over time. Finally, knockout of all three *Tet* genes prevented efficient differentiation and resulted in the differentiation of very few mature motor neurons. Together, our findings highlight a new role for 5mC and 5hmC in motor neurons suggesting 5mC and 5hmC configuration during motor neuron development is necessary for proper functional maturation.

This body of work makes several contributions to the field of DNA methylation. First, the development of LCM-RRBS allows for genome-wide 5mC mapping of laser capture microdissected tissues and of limited DNA input. Second, this study provides the first detailed examination of 5mC patterns in gonadectomy-induced adrenocortical neoplasms. Finally, this study provides the first evaluation of 5mC and 5hmC in maturing motor neurons and provides the first evidence of an important role for the Tet demethylation machinery in normal motor neuron maturation.

CHAPTER 2: LASER CAPTURE MICRODISSECTION-REDUCED REPRESENTATION BISULFITE SEQUENCING (LCM-RRBS) MAPS CHANGES IN DNA METHYLATION ASSOCIATED WITH GONADECTOMY-INDUCED ADRENOCORTICAL NEOPLASIA IN THE MOUSE

Published in Nucleic Acids Research (2013).

ABSTRACT

DNA methylation is a mechanism for long-term transcriptional regulation and is required for normal cellular differentiation. Failure to properly establish or maintain DNA methylation patterns leads to cell dysfunction and diseases such as cancer. Identifying DNA methylation signatures in complex tissues can be challenging due to inaccurate cell enrichment methods and low DNA yields. We have developed a technique called Laser Capture Microdissection-Reduced Representation Bisulfite Sequencing (LCM-RRBS) for the multiplexed interrogation of the DNA methylation status of CpG Islands and promoters. LCM-RRBS accurately and reproducibly profiles genome-wide methylation of DNA extracted from microdissected fresh frozen or formalin-fixed paraffin-embedded tissue samples. To demonstrate the utility of LCM-RRBS, we characterized changes in DNA methylation associated with gonadectomy-induced adrenocortical neoplasia in the mouse. Compared to adjacent normal tissue, the adrenocortical tumors showed reproducible gains and losses of DNA methylation at genes involved in cell differentiation and organ development. LCM-RRBS is a rapid, cost-effective, and sensitive technique for analyzing DNA methylation in heterogeneous tissues and will facilitate the investigation of DNA methylation in cancer and organ development.

INTRODUCTION

DNA methylation has long been recognized to play a role in normal cellular differentiation and development. Methylation most often occurs at the cytosine of a cytosine-guanine dinucleotide (CpG) and acts to down-regulate gene expression (Bird and Wolffe 1999). Disruption of the DNA methylation machinery can lead to imprinting disorders (Robertson 2005), repeat-instability disease (Kumari and Usdin 2009), and neurological defects (Chahrour et al. 2008; Jin et al. 2008).

DNA methylation has been shown to play an important role in cancer progression. Tumors often display a global loss of methylation, or hypomethylation, at repetitive elements, which is thought to destabilize the genome through transposon-mediated rearrangements (Goelz et al. 1985; Wilson et al. 2007), activate growth promoting oncogenes (Wilson et al. 2007), and cause de-differentiation through the loss of imprinting (Reik and Lewis 2005). An abnormal gain of methylation, or hypermethylation, at gene regulatory elements also contributes to tumorigenesis by silencing tumor suppressor genes involved in DNA damage repair, cell cycle control and other processes (Jones and Baylin 2002). This aberrant methylation may be due, at least in part, to recurring mutations in genes that are involved in epigenetic regulation (Stratton 2011; Ryan and Bernstein 2012), such as DNA methyltransferases, which are commonly mutated in acute myeloid leukemia (Ley et al. 2010), and chromatin remodeling enzymes, which are frequently mutated in renal carcinomas and pancreatic neuroendocrine tumors (Dalgliesh et al. 2010; Jiao et al. 2011).

Accurate analysis of DNA methylation is complicated by the heterogeneous nature of normal and diseased tissues. Normal tissues contain cells at different stages of differentiation/maturity. Tumors also consist of histologically diverse cell types (Weinberg 2007; Shackleton et al. 2009) and display intratumor heterogeneity in gene expression (Dalerba et al. 2011), genotype (Navin et al. 2010; Navin et al. 2011), and metastatic and proliferative potential (Fidler et al. 1979; Heppner 1984). Therefore, the analysis of gross tumor samples often obscures the diverse cell types that comprise the entire tumor (Michor and Polyak 2010). To assess cell type specific DNA methylation of complex tissues, cell isolation techniques must be used. Laser capture microdissection (LCM) has enabled researchers to separate specific cell types from heterogeneous tissues (Espina et al. 2006). DNA yields from such samples, however, are too small to use with current methods for genome-wide DNA methylation analysis. Moreover, clinical samples

are typically fixed in formalin and embedded in paraffin, further compromising DNA quality. For these reasons, the genome-wide mapping of DNA methylation in LCM samples has not been previously demonstrated.

Current DNA methylation analysis methods are limited by the number of loci interrogated, quantity and quality of DNA input required, and sample throughput (Laird 2010b). Methods that function on a very small number of cells interrogate only a few genomic loci and are challenging to implement (Dietrich et al. 2009; Herrmann et al. 2011). Furthermore, few methods function on clinical samples that are formalin-fixed and paraffin-embedded (FFPE) (Dietrich et al. 2009). Genome-wide DNA methylation methods are limited by the input of DNA. Affinity enrichment techniques like MeDIP-Seq (Taiwo et al. 2012), MDB-seq (Serre et al. 2010), MethylCap-seq (Brinkman et al. 2010) require 0.16 – 5 μg of DNA input and are limited to a 150- to 200-bp resolution. Other global methods, like CHARM (Irizarry et al. 2008) and padlock probes (Ball et al. 2009; Diep et al. 2012) also require large DNA inputs. MethylC-seq, the only truly whole genome approach (Lister et al. 2009), is prohibitively expensive when many samples need to be analyzed. Reduced Representation Bisulfite Sequencing (RRBS) can map genome-wide DNA methylation of limited DNA samples (Smith et al. 2012), but has not been demonstrated to function on small amounts of DNA recovered from FFPE samples or on samples collected by LCM.

Here, we describe a new technique termed Laser Capture Microdissection-Reduced Representation Bisulfite Sequencing (LCM-RRBS) that can interrogate genome-wide DNA methylation patterns in samples collected from complex heterogeneous tissues. As a proof of principle, we have used LCM-RRBS to analyze global DNA methylation changes associated with adrenocortical neoplasia in the mouse. In response to gonadectomy (GDX) and the ensuing rise in serum gonadotropin levels, sex steroid-producing neoplasms accumulate in the subcapsular region of the adrenal cortex of certain strains of mice, including DBA/2J (Bielinska et al. 2006b). This phenomenon is thought to reflect gonadotropin-induced metaplasia of stem/progenitor cells in the adrenal cortex, although the term *neoplasia* is used more often than *metaplasia* to describe the process (Bielinska et al. 2006b).The molecular basis of GDX-induced adrenocortical neoplasia is unknown (Bernichtein et al. 2008b; Krachulec et al. 2012), but it has been hypothesized that DNA methylation and other epigenetic modifications may impact the phenotypic plasticity of adrenocortical stem/progenitor cells allowing them to respond to the rise in circulating

gonadotropins (Bielinska et al. 2009). GDX-induced adrenocortical neoplasia in the mouse is an ideal phenomenon to study using LCM-RRBS because of the limited amounts of tissue that can be collected.

RESULTS

LCM-RRBS

Reduced Representation Bisulfite Sequencing (RRBS) is an established methylation analysis method that can interrogate most CGIs and promoters across the entire genome. Current RRBS protocols however, do not allow for multiplexing and have not been demonstrated on samples isolated by LCM. Furthermore, although RRBS has been used to analyze small amounts of high quality DNA (Smith et al. 2012), it has not been applied to less than 1 µg of DNA extracted from formalin-fixed paraffin-embedded (FFPE) samples (Gu et al. 2010), the most common method used to preserve clinical tissue samples. We therefore sought to develop a method to analyze small amounts (1 ng) of DNA from laser capture microdissected samples and from FFPE-preserved samples and to interrogate multiple samples in parallel.

Our protocol removes most clean-up steps, which ensures DNA samples are not lost, and leverages the capabilities of Illumina indexing to pool samples prior to size selection and sequencing, thus dramatically increasing the number of samples that can be processed in parallel (Figure 2.1). LCM-RRBS digests genomic DNA with MspI to create fragments with a 5` CpG end. Digested fragments are blunted, adenylated, and ligated with methylated sequencing adapters then column purified to remove excess adapters. To convert the epigenetic methylation mark into a genetic mark that can be read through genomic sequencing, adapter-ligated fragments are treated with bisulfite. At this stage, converted DNA is amplified with a low-cycle PCR to introduce sample-specific indexes. Once each sample is 'indexed,' samples are pooled prior to gel electrophoreses and the isolation of 40-220 base pair fragments. The purified, pooled library is PCR enriched using universal primers and sequenced on the Illumina platform to generate 42 base pair reads. Using our modified method, we can interrogate CpGs genome-wide from laser capture microdissected samples freshly frozen or previously preserved through formalin-fixing and paraffin-embedding.

LCM-RRBS accurately measures genome-wide DNA methylation of fresh frozen and formalin-fixed paraffin-embedded samples

To evaluate the performance of LCM-RRBS, we benchmarked it against RRBS. Using 1 ng and 400 ng for LCM-RRBS and RRBS, respectively, we compared the genome-wide DNA methylation status of DNA isolated from human blood leukocytes. We did not do LCM on the 1 ng of purified DNA; instead, we only applied the downstream library preparation of the LCM-RRBS protocol (Figure 2.1). LCM-RRBS was able to interrogate >75% of CGIs and >65% of gene promoters, results that were similar to those obtained by RRBS (Figure 2.S1). LCM-RRBS was able to accurately measure the DNA methylation levels of CGIs and core promoters (Figure 2.2A) as well as individual CpG dinucleotides (Figure 2.S2). Increasing the required coverage for each CpG considered for CGI methylation did not significantly alter the concordance between RRBS and LCM-RRBS (Figure 2.S3). For CGIs (n = 18,448) and promoters (n = 8,500) having at least 50 methylation measurements, we observed a Pearson correlation of 0.98 and 0.94, respectively, between 1 ng and 400 ng (Figure 2.2A). Most CGIs are either highly methylated (80-100%) or highly unmethylated (0-20%). We therefore sought to test how well LCM-RRBS could call a CGI as methylated or unmethylated. For CGIs with at least 50 high-quality CpG measurements, LCM-RRBS identified methylated CGIs with 91% sensitivity and 99% specificity and unmethylated CGIs with 97% sensitivity and 94% specificity when compared to the RRBS dataset. We therefore conclude that LCM-RRBS functions on as little as 1 ng of genomic DNA, interrogates most CGIs and promoters, and is very accurate.

Human clinical samples are usually stored as either fresh frozen or FFPE specimens. Mapping DNA methylation in the latter, however, can be challenging because formalin fixation degrades DNA. Most DNA methylation techniques that have been used on FFPE samples require greater than 1 µg of DNA or can only interrogate a few loci. To validate the reproducibility of LCM-RRBS and demonstrate its universal clinical applicability, we performed methylation profiling on 1 ng samples of DNA from a primary endometrial carcinoma, half of which was fresh frozen and the remainder which was formalin-fixed and paraffin-embedded. Methylation of CGIs and promoters was highly concordant between fresh frozen and FFPE samples (Pearson correlation 0.98 and 0.97, respectively; Figure 2.2B). CGI and promoter methylation correlated strongly between FFPE technical replicates (Pearson correlation 0.97 and 0.95,

respectively; Figure 2.S4). We also observed high concordance across individual CpGs between fresh frozen and FFPE tumor samples with a Pearson correlation of 0.96 (Figure 2.S2). LCM-RRBS, therefore, can accurately interrogate genome-wide methylation of 1 ng extracted from FFPE samples.

LCM-RRBS is robust across fresh frozen and formalin-fixed paraffin-embedded laser capture microdissected samples

In situ analysis of DNA methylation is challenging due to the heterogeneous nature of complex tissues. In order to interrogate only cells of interest in biological and clinical samples, LCM techniques must be used to enrich for a specific cell type. Current genome-wide DNA methylation methods, however, have not been demonstrated to function on LCM-collected samples. We therefore set out to evaluate the performance of LCM-RRBS on fresh frozen and FFPE samples collected by LCM.

Because the cellular architecture of a normal liver is homogeneous, the methylation state should be very similar throughout the organ. Thus, the liver serves as the ideal tissue for benchmarking the LCM-RRBS method against the RRBS gold standard, since each microdissected region should have a very similar methylation pattern to that of the bulk tissue. We harvested the liver of C57BL/6J mice and prepared the liver using standard preservation techniques. Separate regions of the liver were either directly snap frozen, preserved in Tissue-Tek O.C.T. compound and then snap frozen, or preserved using formalin-fixation and paraffin-embedding. Bulk DNA was extracted from the snap frozen sample and used for downstream RRBS analysis. To determine whether the process of LCM alters the methylation of DNA and to assess the lower limits of LCM-RRBS, we applied LCM-RRBS to samples collected from 20 mm^2, 10 mm^2, 5 mm^2, and 2 mm^2 of the fresh frozen and FFPE mouse liver and compared the DNA methylation patterns to those determined by performing RRBS on 400 ng of DNA extracted from the bulk fresh frozen liver tissue. As observed with 1 ng of DNA, most CGIs and promoters were represented even when only 2 mm^2 of tissue was collected (Figure 2.S5). Furthermore, CGI (mean Pearson = 0.98) and promoter (mean Pearson = 0.96) methylation showed high concordance across all fresh frozen samples (Figure 2.3A) when compared to 400 ng of DNA. For samples collected from 2 mm^2 of microdissected tissue, LCM-RRBS identified unmethylated (0-20%) CGIs with 88% sensitivity and 99% specificity and methylated (80-100%) CGIs with 99% specificity and 87% sensitivity. While the interrogation of DNA

methylation of 2 mm^2 of fresh frozen tissue was robust, 20 mm^2 of FFPE tissue was required for accurate analysis. Using 20 mm^2 of FFPE starting material, LCM-RRBS showed 79% sensitivity and 99% specificity for calling unmethylated (0-20%) CGIs and 99% sensitivity and 78% specificity for calling methylated (80-100%) CGIs with a Pearson correlation of 0.95 (Figure 2.3B). Samples collected from less than 20 mm^2 of FFPE tissue, however, showed poor CGI, promoter, and CpG correlations as compared to 400 ng (data not shown). We conclude that the process of LCM does not alter DNA methylation and that LCM-RRBS accurately determines methylation patterns from as little as 2 mm^2 of fresh frozen tissue. FFPE tissue is more problematic requiring an area of at least 20 mm^2 to achieve acceptable, but not exceptional performance.

Evaluation of PCR Bias

PCR amplification of small amounts of bisulfite-treated DNA can result in PCR bias and inaccurate DNA methylation calling (Warnecke et al. 1997). In DNA samples obtained from females, the X chromosome serves as a good internal control for assessing PCR bias, since X inactivation methylates one copy of the X chromosome at most loci. To determine if LCM-RRBS suffers from PCR bias, we analyzed the fraction of molecules that were methylated at loci known to be affected by X inactivation. As expected, the majority (>70%) of CGIs on the X chromosome showed an intermediate level (30-70%) of DNA methylation across all fresh frozen samples and the 20 mm^2 FFPE sample (Figure 2.S6), demonstrating that LCM-RRBS shows little PCR bias. We conclude that LCM-RRBS shows little PCR bias across 2 mm^2 of fresh frozen tissue and 20 mm^2 of FFPE tissue.

Analysis of GDX-induced adrenocortical neoplasia in the mouse using LCM-RRBS

To demonstrate the utility of LCM-RRBS in a biological setting, we applied the method to analyze the DNA methylation of neoplasms that arise in the adrenal cortex of DBA/2J mice after GDX. Although genetic factors have been identified that influence susceptibility to GDX (Bernichtein et al. 2008b; Krachulec et al. 2012), little is known about the role DNA methylation plays in the formation of GDX-induced neoplasia. Molecular characterization is further complicated because mice adrenal glands are only 0.1 cm^2 in size and neoplasms arise among normal tissues requiring LCM enrichment methods for

tissue isolation. We therefore applied LCM-RRBS to neoplastic and adjacent normal mouse adrenal tissues (Figure 2.4).

In one multiplexed experiment, we interrogated the methylation status of >13,000 promoters and >13,000 CGIs of 6 tissue samples (3 neoplastic and 3 normal representing 3 different mice) across an average of >800,000 CpGs per sample (Figure 2.4 and Table 2.S1). Using a threshold difference of at least 10%, 37 promoters were significantly hypomethylated and 8 promoters were significantly hypermethylated ($P < 0.05$, FDR adjusted; Table 2.1 and Table 2.S2) in the neoplasms compared to adjacent normal tissue. Many of the top hypo- and hypermethylated genes have been implicated in cell fate determination and differentiation, including adrenocortical formation (*Tinagl1*), gonad development (*Foxs1*, *Wdr63*, *Tmem184a*), pancreas development (*Nsmce1*), kidney development (*Hoxc10*, *Dpep1*), prostate development (*Il17rc*, *Ano7*), and muscle and skeletal development (*Myo18b*, *Trim63*, *Lmod3*, *Meox1*). The observed methylation changes suggest the neoplastic tissue may arise due to aberrant gene expression of genes normally silent in adrenocortical cells or the silencing of adrenal-specific markers. To validate our findings, we performed bisulfite-specific polymerase chain reaction (BSP) followed by sequencing of 3 hypomethylated promoters and 1 hypermethylated promoter on neoplastic and normal tissues isolated by LCM. For all promoters tested, BSP showed a significant difference (Fisher's exact test, $P < 10^{-15}$) in DNA methylation between the neoplasia and normal tissue as predicted by LCM-RRBS (Figure 2.5). Taken together, these results demonstrate that LCM-RRBS can identify differentially methylated genes in a complex tissue and reveal functionally relevant epigenetic effects.

DISCUSSION

Current DNA methylation mapping techniques are limited by input and the number of loci interrogated. RRBS, a genome-wide DNA methylation mapping technique, was recently shown to function on 0.5 to 10 ng of genomic DNA isolated from mouse embryos (Smith et al. 2012). RRBS, however, has not been demonstrated to function on LCM samples collected from FFPE tissue nor is it amenable to large scale sample processing. We have developed a new method, LCM-RRBS, which can accurately profile genome-wide DNA methylation of many LCM samples in parallel at single base pair resolution.

Our method can be implemented in 3-4 days, and the bulk of the protocol can be automated for high-throughput 96-well experiments. While traditional RRBS requires each processed sample to undergo gel extraction, a laborious process when processing more than a few samples, our method pools all samples together prior to gel extraction, reducing the required number of gel extractions to one. Thus, a large number of samples can be easily processed at a single time. Furthermore, because high DNA loss results from gel extraction, pooling samples prior to gel extraction allows the use of low (1 ng) DNA inputs.

The LCM-RRBS protocol affords a significant reduction in sequencing costs compared to whole genome bisulfite sequencing. We typically collect 1.5 gigabases (GB) per sample, which is considerably less than the ~60 GB needed for 20x coverage of a whole genome bisulfite library. The sequencing cost per sample can be reduced further if fewer CpGs are interrogated. For example, if a smaller size fraction is isolated during gel extraction, only about 0.75 GB are required per sample.

We found that although formalin-fixation and paraffin-embedding does not alter DNA methylation *per se*, at least 20 mm^2 of tissue must be isolated for accurate DNA methylation profiling. We were able to create LCM-RRBS libraries from 10 mm^2, 5 mm^2, and 2 mm^2 of FFPE tissue and obtained similar numbers of sequencing reads as with fresh frozen samples, but overall mapping quality was very low (~30% aligned) in the FFPE samples, precluding an accurate analysis of DNA methylation. On the other hand, LCM-RRBS generated high quality methylation maps from 2 mm^2 of microdissected fresh frozen tissue, as demonstrated by our analysis of mouse liver.

To demonstrate the utility of LCM-RRBS, we analyzed the DNA methylation patterns of GDX-induced adrenocortical neoplasms using an average of 5.5 mm^2 of fresh frozen tissue. We hypothesized that aberrant DNA methylation changes could be involved in the formation of these neoplastic tissues. Indeed, recent studies have shown that altered DNA methylation can redirect cell fate in endocrine tissues (Dhawan et al. 2011a). Conditional mutagenesis of the mouse *Dnmt1* gene, which encodes the maintenance DNA methyltransferase, converts insulin-producing pancreatic β-cells into glucagon-producing α-cells (Dhawan et al. 2011a). It is thought that because of a common developmental origin, β- and α-cells share general epigenetic programs that provide a compatible environment for cell fate conversions (Akerman et al. 2011). GDX-induced adrenocortical neoplasia may be another example of

DNA methylation-regulated cell fate conversion in an endocrine tissue; in this case, adrenocorticoid-producing cells become sex-steroid producing cells (Bielinska et al. 2003b; Bielinska et al. 2005; Johnsen et al. 2006b). The changes in DNA methylation we observe around the transcription start site (TSS) could lead to changes in gene expression (Brenet et al. 2011; Hartung et al. 2012). Several of the genes we found to be differentially methylated in GDX-induced adrenocortical neoplasms have established roles in adrenocortical or gonadal development. For example, *Tinagl1*, a gene implicated in adrenal zonation (Mukai et al. 2003; Li et al. 2007), showed a gain in DNA methylation, which could lead to down-regulation. *Wdr63, Foxs1, and Tmem184a,* genes involved in gonadal development (Svingen et al. 2007; Best et al. 2008; Bonilla and Xu 2008; Sato et al. 2008), showed a loss of DNA methylation, which could lead to the aberrant expression of these gonadal-like markers in the adrenal cortex. Furthermore, *Srd5a3*, a gene involved in the biosynthesis of the potent androgen 5α-dihydrotestosterone (Uemura et al. 2008), showed a loss of DNA methylation, which could enhance the ectopic production of sex steroids in the adrenal gland (Payne and Hales 2004). Future studies will explore the role of these methylation changes in the pathogenesis of GDX-induced adrenocortical neoplasia.

In conclusion, LCM-RRBS is a robust, cost-effective method for the DNA methylation analysis of heterogeneous tissues. This technique allows the study of tumor evolution and epigenetic heterogeneity *in situ* of less than 1 ng (~150 cells) and can also be applied to investigate the role of DNA methylation in cell fate specification during tissue development. LCM-RRBS is an important milestone toward highly parallel *in situ* analysis of single cells. We anticipate that this protocol will greatly facilitate the analysis of any sample that contains multiple cell types.

ACKNOWLEDGEMENTS

We thank Amy Schmidt and Paul Goodfellow for kindly providing endometrial samples and members of the Mitra Lab and Gordon Lab for thoughtful feedback and suggestions throughout this work. We thank Dietmar Spengler for providing the image of the mouse brain section. We thank Vanessa Ridaura, Constantino Schillebeeckx, and Charles Higdon for assistance in creating the figures. We thank Frank Schoettler of the Vision Research Community Morphology and Imaging Core for assistance with LCM.

AUTHOR CONTRIBUTIONS

This work was done in collaboration with Anja Schrade, Ann-Kathrin Löbs, Marjut Pihlajoki, Dave Wilson, and Rob Mitra. Rob Mitra, Dave Wilson, and I conceived this project and designed the experiments. Anja Schrade, Ann-Kathrin Löbs, and Marjut Pihlajoki performed mouse dissections and laser capture microdissection on adrenal glands. I designed and developed LCM-RRBS, performed LCM on liver tissue, performed bisulfite PCR validation experiments, and analyzed all data. Rob Mitra, Dave Wilson, and I wrote the paper. This chapter is a manuscript published in 2013 in the journal *Nucleic Acids Research*. This work was supported by the National Institutes of Health grants 1R01DA025744, to RDM, and 1R01DK075618, to DBW].

MATERIALS AND METHODS

Experimental mice

Procedures involving mice were approved by an institutional committee for laboratory animal care and were conducted in accordance with NIH guidelines for the care and use of experimental animals. C57BL/6J and DBA/2J mice were purchased from Jackson Laboratories (Bar Harbor, ME). Mice were anesthetized and ovariectomized at 3-4 weeks of age (Bielinska et al. 2005).

DNA extraction

Human tumor specimens were collected under an Institutional Review Board (IRB)-approved protocol. Immediately after surgery, a human endometrial tumor was divided in half. One half was fresh frozen while the other was formalin-fixed and paraffin-embedded (FFPE). 50 mg of fresh frozen endometrial tumor was cut into small pieces with a sterile scalpel blade and DNA extracted using the Maxwell 16 Tissue DNA Purification Kit (AS1030, Promega). The formalin-fixed paraffin-embedded preserved portion was cut into 6-µm sections onto microscope slides. Four 4 mm^2 slices were scratched off the slide with a sterile scalpel blade, combined in 80 µl buffer and proteinase K (740901.50, Clontech), and incubated overnight at 65°C. Liver tissue was harvested from C57BL/6J mice and divided in half. One half was preserved in Tissue-tek optimal cutting temperature (O.C.T.) compound (25608-930, VWR) and snap frozen, while the other half was formalin-fixed and paraffin-embedded for downstream bulk DNA extraction and LCM. Fresh frozen, FFPE, and LCM samples were purified using NucleoSpin Tissue XS columns (740901.5, Clontech) following the protocol for laser-microdissected tissue and eluted in 20 µl of nuclease-free water. Genomic DNA was quantified using the Quant-it dsDNA High Sensitivity kit (Invitrogen) and the Qubit fluorometer (Invitrogen).

Laser Capture Microdissection

Adrenal glands were harvested from mice 3 months after ovariectomy. Liver and adrenal cryosections (10 µm) were collected on membrane slides (PEN-Membrane 2.0 µm; Leica) designed to free the dissectate from the remainder of the tissue section. Adrenal tissue sections were fixed in

acetone (5 sec, -20°C), stained with hematoxylin and eosin (H&E) or crystal violet, and dehydrated by passage through successively higher concentrations of ethanol followed by xylene. FFPE livers sections were deparaffinised and dehydrated using standard methods. LCM was performed using a Leica LMD6000 microscope. Dissectates were collected in SDS/proteinase K for genomic DNA isolation (740901.5, Clontech).

RRBS and LCM-RRBS

RRBS was performed on 400 ng of commercially purchased leukocyte genomic DNA (D1234148, Amsbio) and genomic DNA extracted from a mouse liver as previously described (Gertz et al. 2011). For LCM-RRBS, leukocyte genomic DNA (1 ng), extracted endometrial tumor genomic DNA (1 ng), and LCM DNA samples were incubated overnight at 37°C with 20 U of the methylation insensitive restriction enzyme MspI (R0106S, NEB) and 2 µl of 10 x NEBuffer 2 in a 18 µl reaction. Without subsequent purification, fragment ends were filled in and an adenosine added with 10 U of Klenow Fragment (3` → 5` *exo*$^-$, M0212L, NEB), 0.04 mM dGTP, 0.04 mM dCTP, 0.4 mM dATP, and 1x NEB Buffer 2 in a final volume of 22.4 µl. The reaction was incubated at 30°C for 20 min, 37°C for 20 min, and heat inactivated at 75°C for 20 min. Pre-annealed methylated paired-end Illumina indexing adapters (Adap1: ACACTCTTTCCCTACACGACGCTCTTCCGATCT, Adap2: P-GATCGGAAGAGCACACGTCTGAACTCCAGTCAC; P = phosphate) at a concentration of 0.26 mM were ligated overnight at 16°C to the ends of the DNA fragments using 1200 U of T4 DNA Ligase (M0202L, NEB) in 1x Ligase Buffer at a final volume of 28.9 µl. These adapter oligonucleotides are only complementary at 13 bases which, after annealing, form a “Y” structure. Because excess adapters prevent the complete conversion of CpGs at the MspI digestion site, adapter-ligated fragments are purified using MinElute columns (Qiagen) and eluted twice with 11 µl of warm EB buffer. The purified products were treated using the EZ DNA Methylation Gold Kit (D5005, Zymo). Samples were eluted in 11 µl of M-Elution buffer. To incorporate the sample-specific index, 3 µl of each bisulfite treated sample was amplified in triplicate with 0.2 µM of indexed primers (PCR1: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT; PCR2: CAAGCAGAAGACGGCATACGAGATNNNNNGTGACTGGAGTTCAGACGTGTGCTCTTCCGA,

N=index), 5 U of Platinum Taq Polymerase (10966-034, Invitrogen), 1x PCR buffer, 2 mM $MgCl_2$, 0.5 M betaine (B0300, Sigma), 1 mM dNTP, in a 10 µl reaction using the following cycling conditions: 98°C for 2 min, 12 cycles of 98°C for 30 sec and 65°C for 2 min. All PCR products and replicates were pooled and analyzed by electrophoresis on a 3% 1X Tris-acetate-EDTA (TAE) NuSieve agarose gel (50090, Lonza) using a voltage of 5 V/cm until the blue loading dye was 6-7 cm away from the loading well. Fragments between 150 bp and 350 bp were extracted and purified using MinElute columns (Qiagen) and 15 µl of warmed EB buffer. Prior to final library PCR enrichment, the minimum cycle number must be determined to ensure no PCR bias. Using 2 µl of eluted product and 0.2 µM universal primers (Pool1: CAAGCAGAAGACGGCATACGAGAT, Pool2: AATGATACGGCGACCACCGAGATCT), multiple PCR reactions with a final volume of 50 µl are set up using the previous conditions but varying the cycle number from 10 – 16 cycles. 10 µl of each PCR product is analyzed through electrophoresis on a 4-20% Precast TBE gel (3450059, BioRad) and stained with Sybr Gold (S-11494, Invitrogen) for 15 min and imaged. In order to minimize PCR bias, the final PCR library is amplified in quadruplicate using the previous PCR conditions and the minimum cycle number (typically ~14) that shows amplification only within the 150-350 bp range on the Sybr Gold stained TBE gel. The four replicates are pooled and gel extracted as previously mentioned to remove remaining adapter dimers and primers, then purified, and sequenced on Illumina HiSeq 2000 machines.

Bisulfite-specific PCR

Neoplastic and adjacent normal mouse adrenocortical tissue was collected and bisulfite treated using the EZ DNA Methylation-Direct kit following the manufacturer's instructions for LCM samples eluting with 15 µl of Nuclease-free water (D5020, Zymo). Bisulfite-specific PCR primers were designed using MethPrimer (http://www.urogene.org/methprimer/). To amplify the promoter regions of interest, 2 µl of bisulfite-treated DNA was combined with 2.5 U of Jumpstart Taq (Sigma), 1x PCR buffer, 1 M betaine (B0300, Sigma), 0.2 mM dNTP, and 0.4 µM of each primer in a total reaction volume of 25 µl. The reaction was incubated at 95°C for 5 min, followed by 5 cycles of 94°C for 30 sec, 60°C for 30 sec, and 72°C for 90 sec, followed by 5 cycles of 94°C for 30 sec, 55°C for 30 sec, and 72°C for 90 sec, followed by 30-33 cycles of 94°C for 30 sec, 50°C for 30 sec, and 72°C for 90 sec, followed by 72°C for 5 min.

PCR products were prepared for sequencing on the MiSeq following the manufacturer's protocol (Illumina).

Sequence alignment and methylation calling

All analysis was performed using the February 2009 (GrCh37/hg19) build of the human genome and the July 2007 (NCBI37/mm9) build of the mouse genome. On average 25 million single-end 42-bp raw high quality reads per sample were either aligned to the cytosine-poor strand reference using the bisulfite mode of MAQ (Li et al. 2008a) or aligned to the reduced reference using RRBSMAP (Xi et al. 2012) filtering against reads that contain adapter sequence. Reads that showed less than 90% bisulfite conversion (~1 unconverted non-CpG cytosine per read) were filtered to remove those that resulted from incomplete bisulfite converted molecules. Aligned reads with a mapping quality of zero were also discarded. The resulting high quality uniquely mapped reads were used for methylation calling. We identified the genomic coordinates of all CpGs in the reference sequence and assessed percent DNA methylation by calculating the fraction of reads that had an unconverted cytosine at the CpG position relative to the total reads. We required that each read have either a "TG" or "CG" dinucleotide at the expected CpG coordinate to be considered for analysis.

Genomic feature annotation and statistical analysis

Cytosine methylation levels were determined for two classes of genomic features downloaded from the UCSC genome browser (Rhead et al. 2010). CpG Islands (CGIs) were defined as a region greater than 200 base pairs with a GC content of 50% or greater and observed-to-expected ratio of CG dinucleotides greater than 0.6 (Gardiner-Garden and Frommer 1987). Promoters were defined as a 2 kb region centered on the annotated transcription start site of RefSeq genes (Pruitt et al. 2012). For LCM-RRBS, RRBS, fresh frozen, and FFPE comparisons, only genomic features with at least 100 methylation measurements in each pairwise comparison were considered for analysis.

To identify differentially methylated promoters in adrenocortical neoplasia and normal samples, the DNA methylation status of all CpGs within a 2 kb region of all RefSeq annotated transcription start sites were compared. Those promoters with at least 50 methylation measurements that showed greater

than 10% methylation difference were considered for statistical analysis. Promoters were considered statistically significant with a $P < 0.05$ using Student's t-test after p-values were adjusted using a false-discovery rate (FDR) of 5%. Statistical significance across bisulfite-specific PCR samples was determined using the Fisher's exact test. All statistical analysis was performed using R.

Data Release

The DNA methylation data generated for this study can be found under the NCBI Gene Expression Omnibus (GEO) accession number GSE45361. DNA methylation and raw sequence data are also publically available at the Center for Genome Sciences (www.cgs.wustl.edu/~maxim/).

FIGURE 2.1

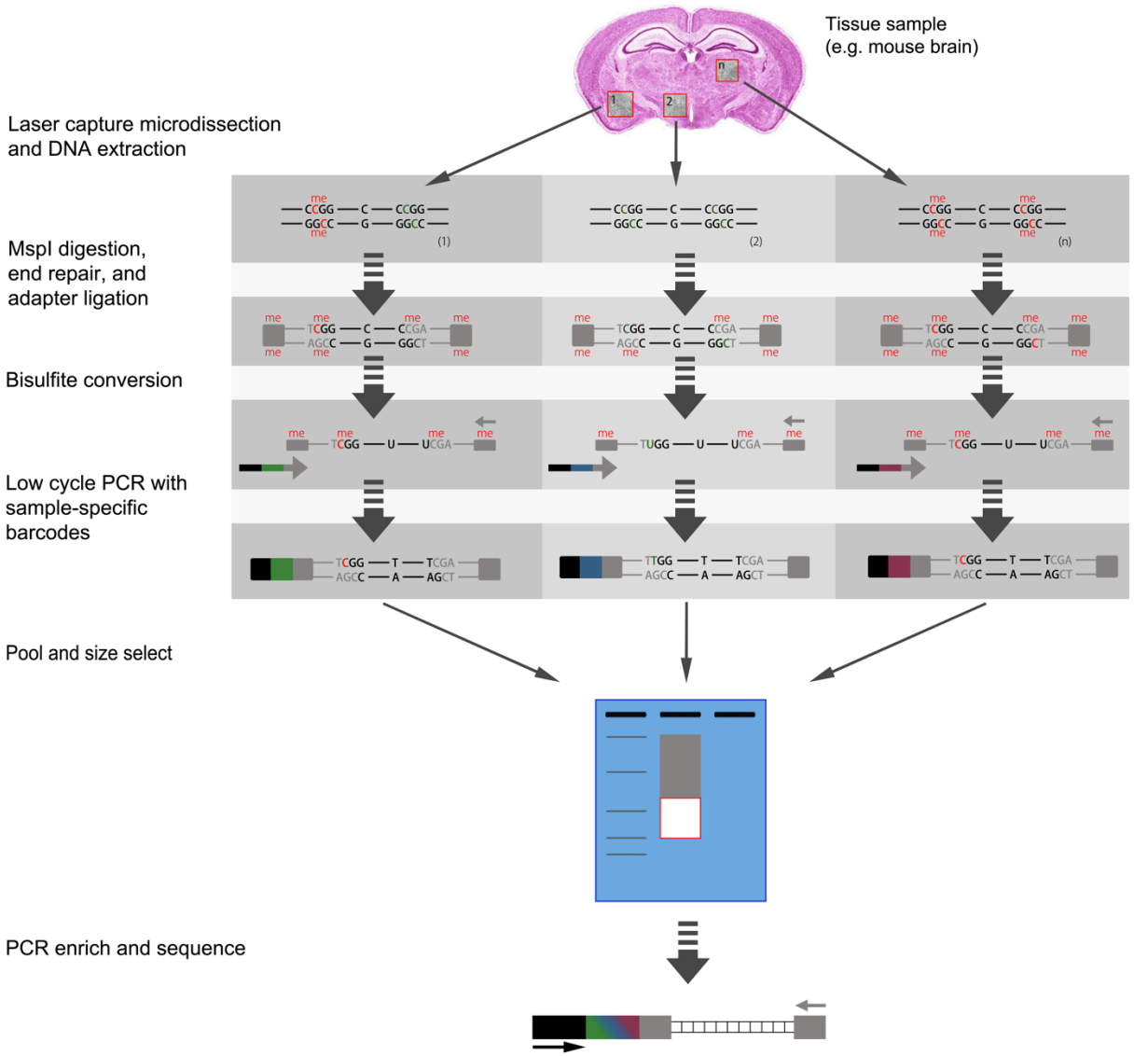

Figure 2.1 LCM-RRBS workflow. A complex tissue is dissected using LCM. Extracted DNA is digested by the methylation insensitive enzyme MspI, end repaired, and ligated with methylated Illumina adapters. After bisulfite conversion, each sample is 'barcoded' by introducing a sample-specific index (shown as green, blue, or violet boxes) through low cycle PCR. Samples are pooled and loaded onto a high percentage gel for fragment separation and size selection. Using universal primers, the final library is amplified and sequenced on the Illumina platform.

FIGURE 2.2

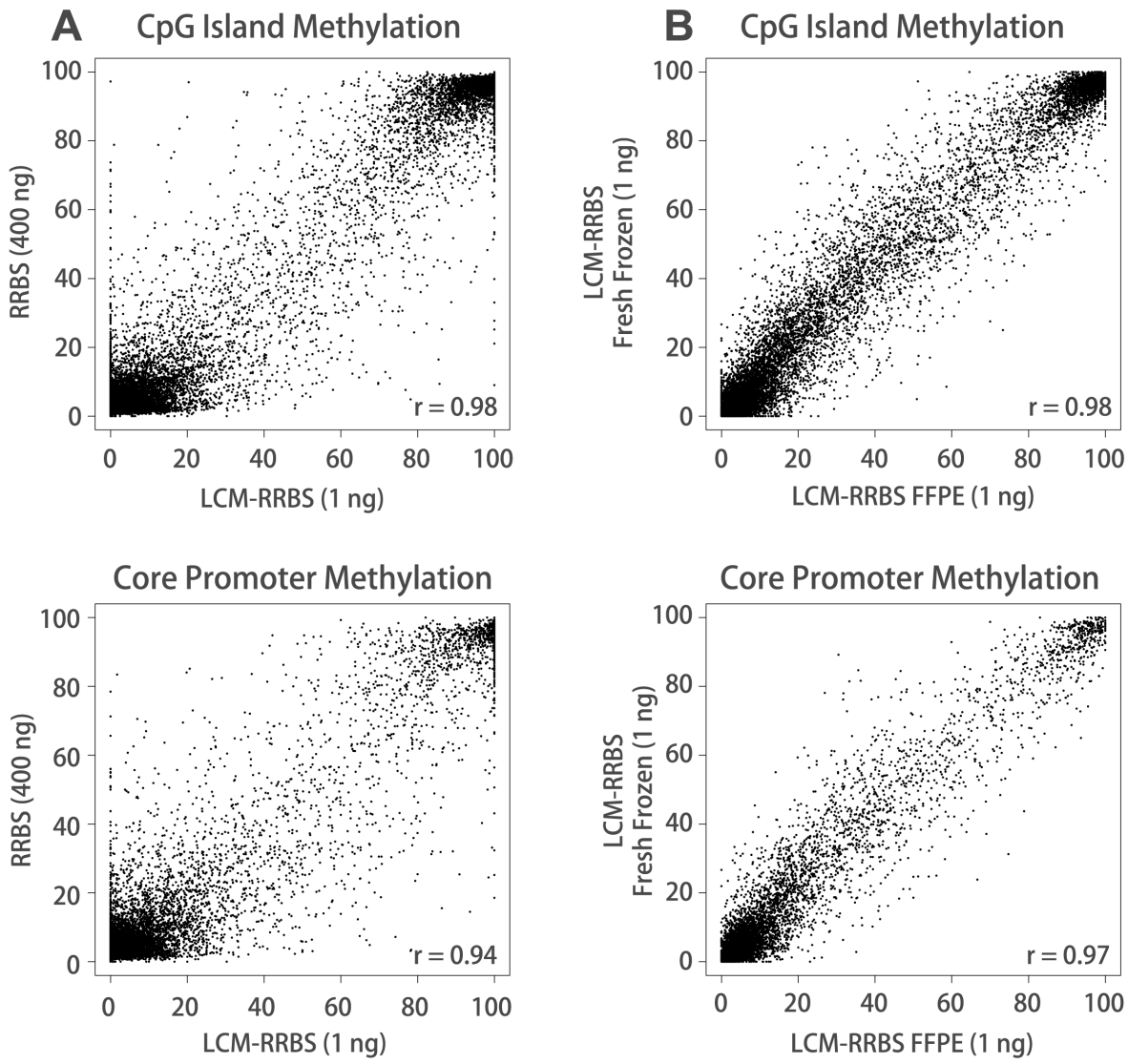

Figure 2.2 LCM-RRBS is reproducible and robust across 1 ng extracted from bulk fresh frozen and formalin-fixed paraffin embedded (FFPE) samples. CpG Island methylation (top panels) and the methylation at 2 kb regions flanking the transcription start site (bottom panels) were compared between (**A**) 1 ng (LCM-RRBS) and 400 ng of purified leukocyte genomic DNA (RRBS); and (**B**) 1 ng of FFPE DNA and 1 ng of fresh frozen genomic DNA extracted from the same endometrial tumor (LCM-RRBS).

FIGURE 2.3

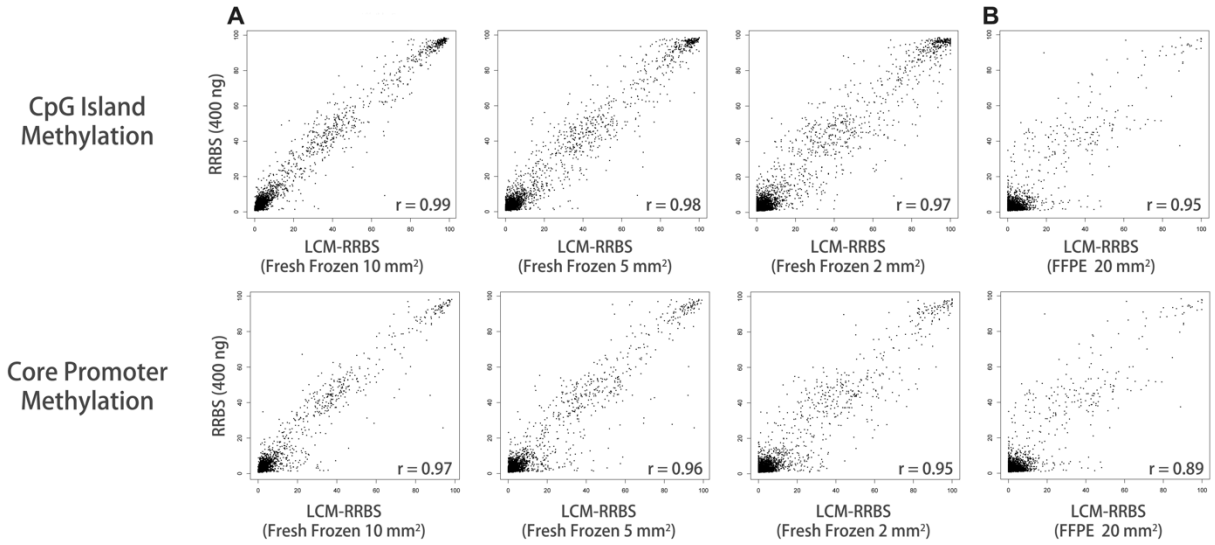

Figure 2.3 LCM-RRBS is robust across microdissected samples collected from fresh frozen and FFPE tissues. Fresh frozen and formalin-fixed paraffin-embedded (FFPE) mouse liver was collected for DNA methylation profiling. LCM was used to collect tissue from areas ranging in size from 20 to 2 mm^2. CpG Island methylation (top panels) and methylation at 2 kb regions flanking the transcription start site (bottom panels) were compared between (**A**) fresh frozen samples (LCM-RRBS) and 400 ng of purified mouse liver genomic DNA (RRBS); and (**B**) FFPE samples (LCM-RRBS) and 400 ng of purified mouse liver genomic DNA (RRBS).

FIGURE 2.4

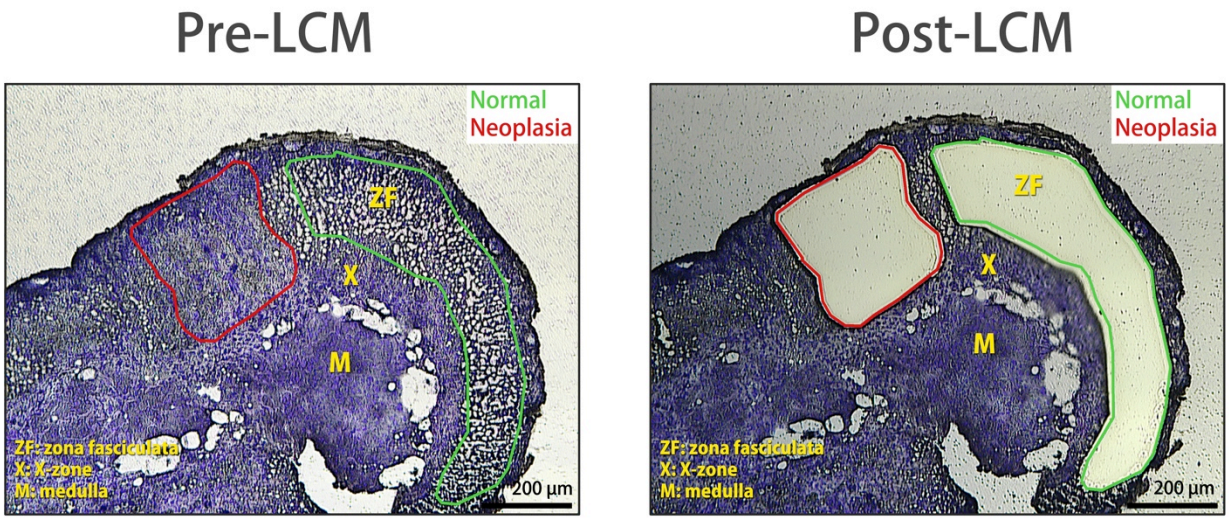

Figure 2.4 DNA methylation profiling of GDX-induced adrenocortical neoplasms and adjacent normal tissue using LCM-RRBS. The adrenal glands of 3 ovariectomized DBA/2J mice were fresh frozen in Tissue-tek O.C.T. compound, cryosectioned, and stained. Shown are representative cryosections pre- and post-LCM. Normal cells in the zona fasciculata (ZF) contain large lipid droplets that are easily recognized. In contrast, neoplastic cells distort the normal adrenal zonal architecture and contain relatively few lipid droplets. The microdissected normal tissue included zona glomerulosa and zona fasciculata cells; care was taken to avoid dissection of X-zone (X), medulla (M), or capsule cells, as these cell types have distinct developmental origins (Morohashi and Zubair 2011) and therefore may have different epigenetic fingerprints. An average of 5.5 mm^2 of neoplastic (red) and normal (green) tissue per adrenal pair was collected and analyzed using LCM-RRBS.

FIGURE 2.5

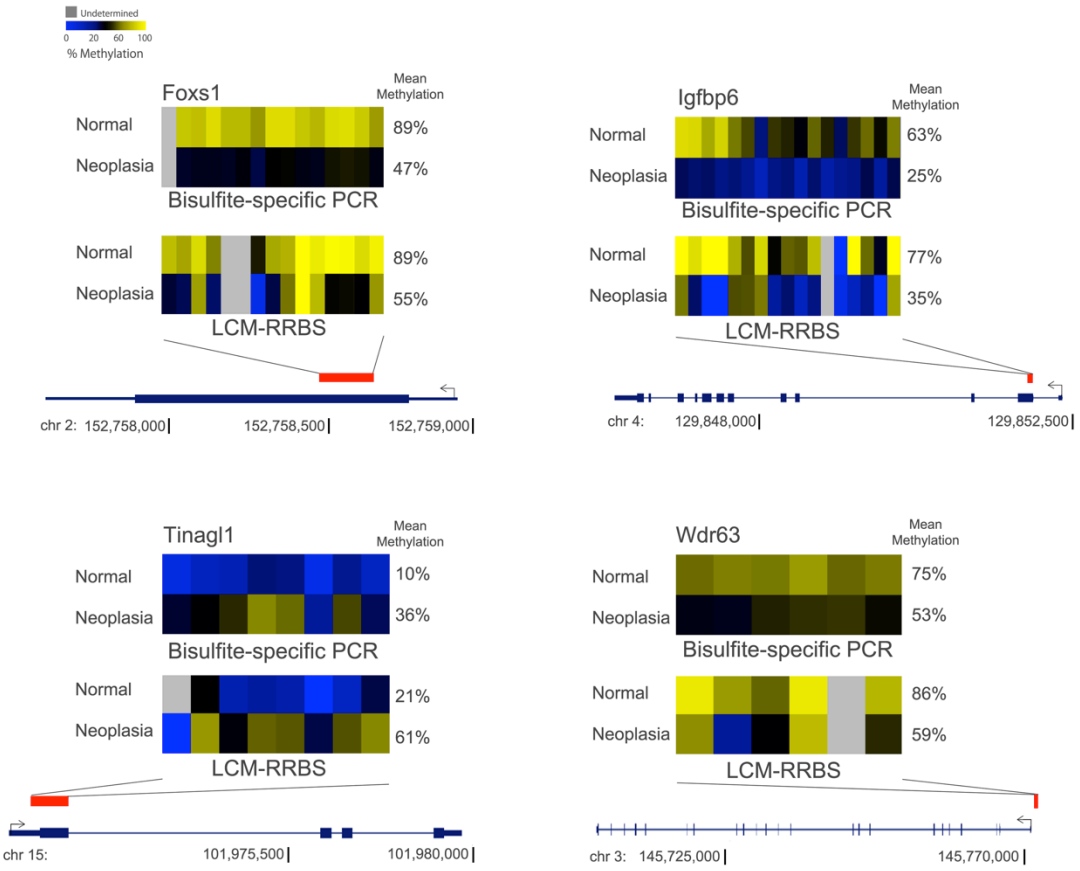

Figure 2.5 Validation of differentially methylated promoters. The DNA methylation of one hypermethylated and three hypomethylated promoters was interrogated by bisulfite-specific PCR (BSP) and sequencing across enriched neoplastic and normal samples. All genes show a statistically significant difference (Fisher's exact test, $P < 10^{-15}$) in DNA methylation using bisulfite-specific PCR. Each colored box represents an individual CpG dinucleotide within a 2 kb region centered around the transcription start site. High (yellow), moderate (black), low (blue), and undetermined methylation levels are shown for each CpG. The mean methylation of each region interrogated is shown to the right of each heatmap. The red box indicates the region of the promoter that was interrogated by LCM-RRBS and BSP.

TABLE 2.1

Gene symbol	Entrez gene name	Function[a]	Percent methylation[b]			*P* value[c]
			Normal	Neoplasia	Difference	
Tinagl1	Tubulointerstitial nephritis antigen-like 1	Zonal differentiation of adrenocortical cells	18	64	46	0.0235
Dennd4b	DENN/MADD domain containing 4B	unknown	34	62	29	0.0125
L1td1	LINE-1 type transposase domain containing 1	Embryonic stem cell renewal and identity	33	59	27	0.0241
Syne4	Spectrin repeat containing, nuclear envelope family member 4	Microtubule-dependent nuclear positioning	45	70	25	0.0189
Fignl2	Fidgetin-like 2	Unknown	40	63	23	0.0125
Slc5a5	Solute carrier family 5 (sodium iodide symporter), member 5	Regulates iodine uptake in the thyroid	33	51	18	0.0352
Nsmce1	Non-SMC element 1 homolog	Pancreas development	67	81	14	0.0320
Hoxc10	Homeobox C10	Kidney development and limb formation	42	54	12	0.0125
Myo18b	Myosin XVIIIb	Myofibrillar structure maintenance	86	34	52	0.0230
B3gnt8	UDP-GlcNAc:betaGal beta-1,3-*N*-acetylglucosaminyltransferase 8	Positive regulator of cell proliferation	72	32	40	0.0125
P2rx7	Purinergic receptor P2X, ligand-gated ion channel, 7	Neuronal differentiation and migration	79	46	33	0.0244
Dpep1	Dipeptidase 1 (renal)	Kidney and genitourinary development	80	49	31	0.0027
Foxs1	Forkhead box S1	Neuronal differentiation and testicular development	95	64	31	0.0027
Angptl2	Angiopoietin-like 2	Mediates differentiation, migration and inflammation	84	54	30	0.0125
Igfbp6	Insulin-like growth factor binding protein 6	Igf2 signaling	76	46	30	0.0406
Nt5dc2	5′-nucleotidase domain containing 2	Unknown	85	57	28	0.0224
Trim63	Tripartite motif-containing 63	Cardiomyocyte development	69	41	28	0.0352
Ica	RIKEN cDNA 1300017J02 gene	Putative hepatic iron regulator	86	59	27	0.0125
Wdr63	WD repeat domain 63	Gonad development	85	59	27	0.0128

[a]Gene function from NCBI, GeneRIF.
[b]Mean methylation of three mice.
[c]FDR corrected.

Table 2.1 Top hypermethylated and hypomethylated genes in GDX-induced adrenocortical neoplasms of the mouse.

SUPPLEMENTARY FIGURES AND TABLES

FIGURE 2.S1

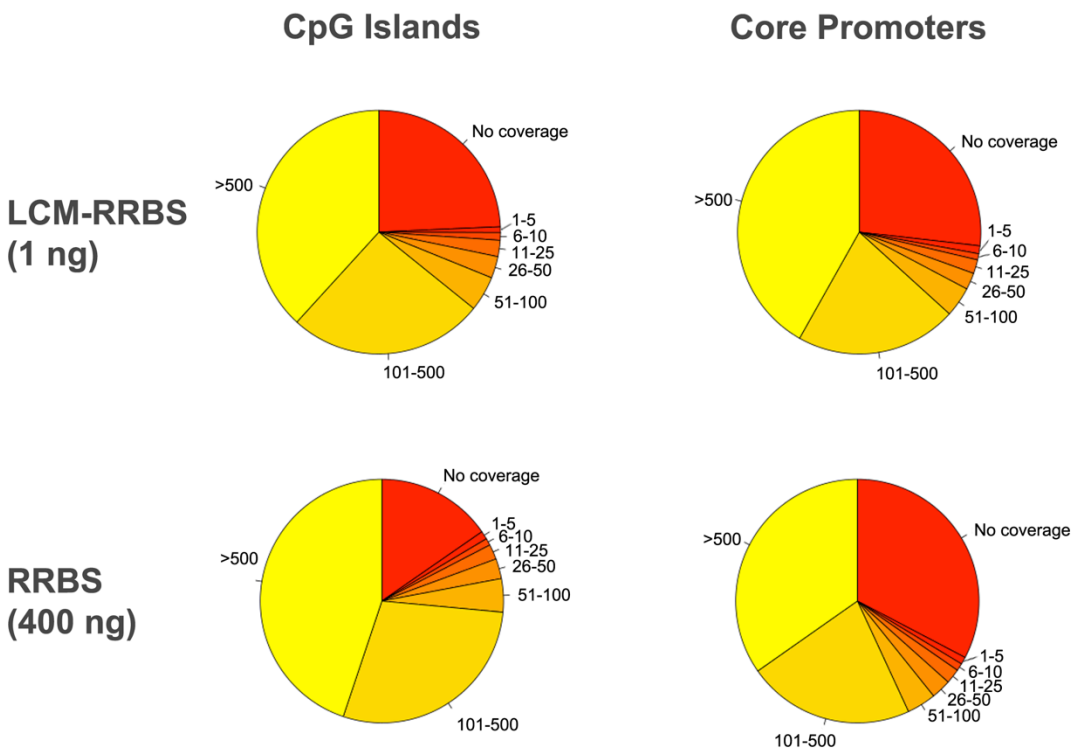

Figure 2.S1 LCM-RRBS can interrogate the DNA methylation status of most CpG Islands and core promoters. For all UCSC annotated CpG Islands and core promoters, defined as a 2 kb region centered on the transcription start site, we determined the read coverage. Coverage is defined as the total number of reads covering a CpG within a given genomic interval.

FIGURE 2.S2

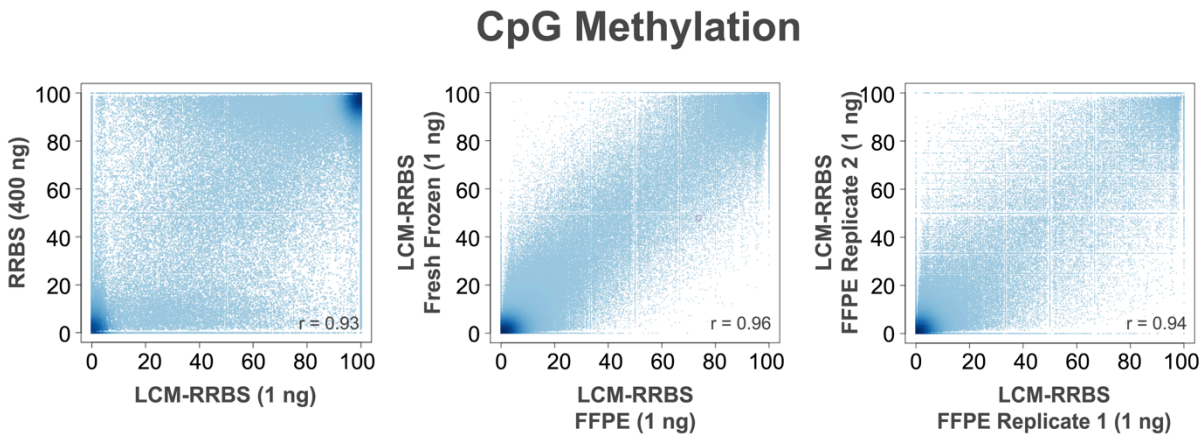

Figure 2.S2 LCM-RRBS is reproducible and robust at the CpG level across 1 ng of formalin-fixed paraffin-embedded DNA. LCM-RRBS and RRBS were applied to 1 ng and 400 ng of commercially purchased leukocyte genomic DNA, respectively, and showed a high concordance across individual CpGs (Pearson, 0.93). LCM-RRBS was applied to 1 ng of DNA extracted from fresh frozen or formalin-fixed paraffin-embedded endometrial tumor tissue. CpG methylation was highly concordant between fresh frozen and FFPE samples and FFPE technical replicates (Pearson, 0.96 and 0.94, respectively). Only CpGs covered by at least 20 reads in each sample were considered.

FIGURE 2.S3

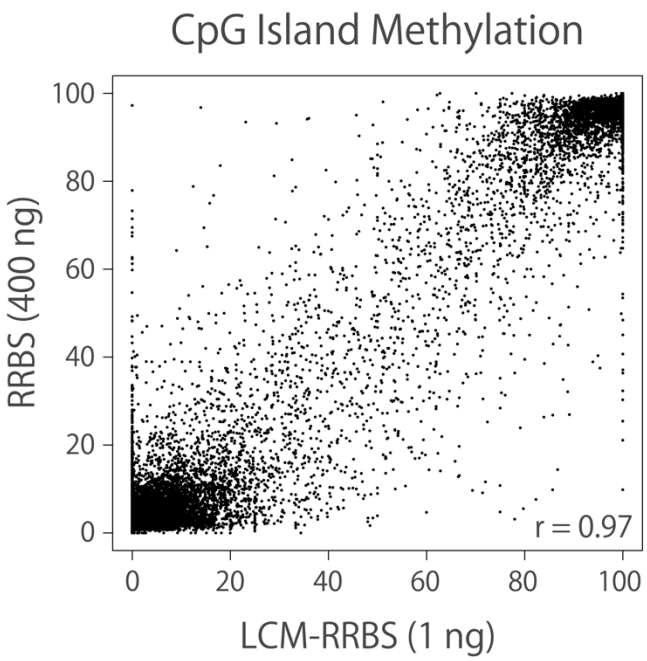

Figure 2.S3 Increasing individual CpG coverage does not significantly alter the concordance of CGI methylation between data collected by RRBS (400 ng) and LCM-RRBS (1 ng). The methylation of each CGI was calculated by determining the mean methylation of only those CpGs within the region that were covered by at least 10 reads in both datasets. Only CGIs that were covered by a total of 100 reads were considered for concordance analysis.

FIGURE 2.S4

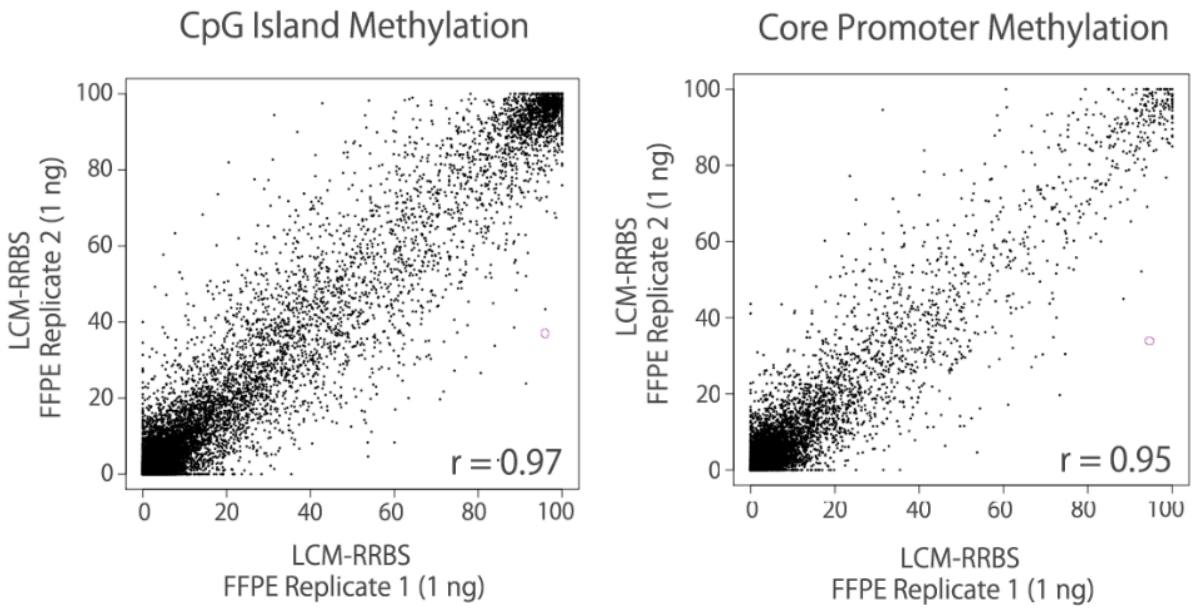

Figure 2.S4. **LCM-RRBS is reproducible and robust across 1 ng extracted from bulk FFPE samples.** CpG Island methylation (top panels) and the methylation at 2 kb regions flanking the transcription start site (bottom panels) were compared between two technical replicates of 1 ng of FFPE endometrial tumor DNA using LCM-RRBS.

FIGURE 2.S5

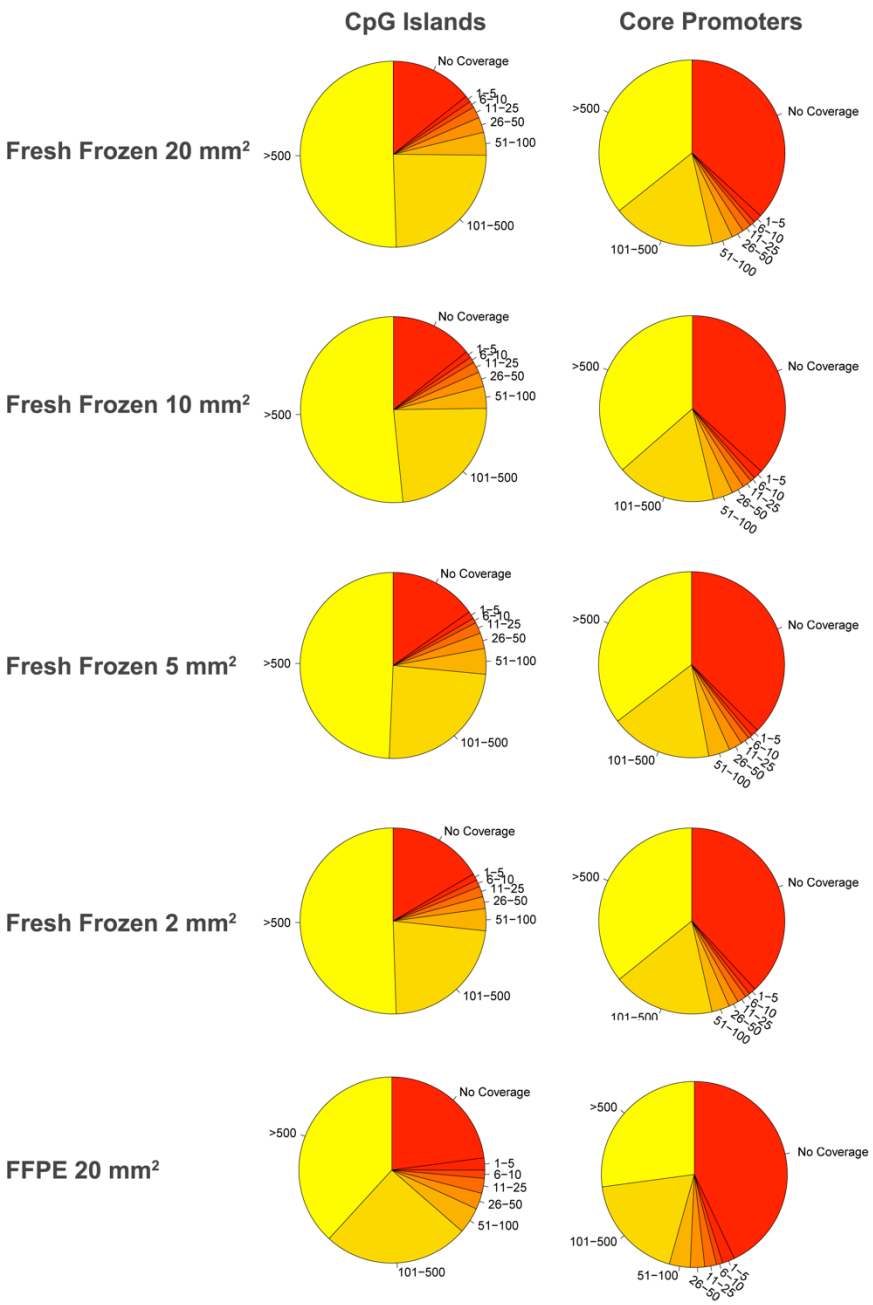

Figure 2.S5 LCM-RRBS interrogates most CpG Islands and core promoters of liver mouse samples collected by laser capture microdissection (LCM). For CpG Islands and core promoters, defined as a 2 kb region centered on the RefSeq transcription start site, we determined the mean read coverage of samples collected by LCM from fresh frozen and FFPE tissues. Coverage is defined as the total number of reads covering a CpG within a given genomic interval.

FIGURE 2.S6

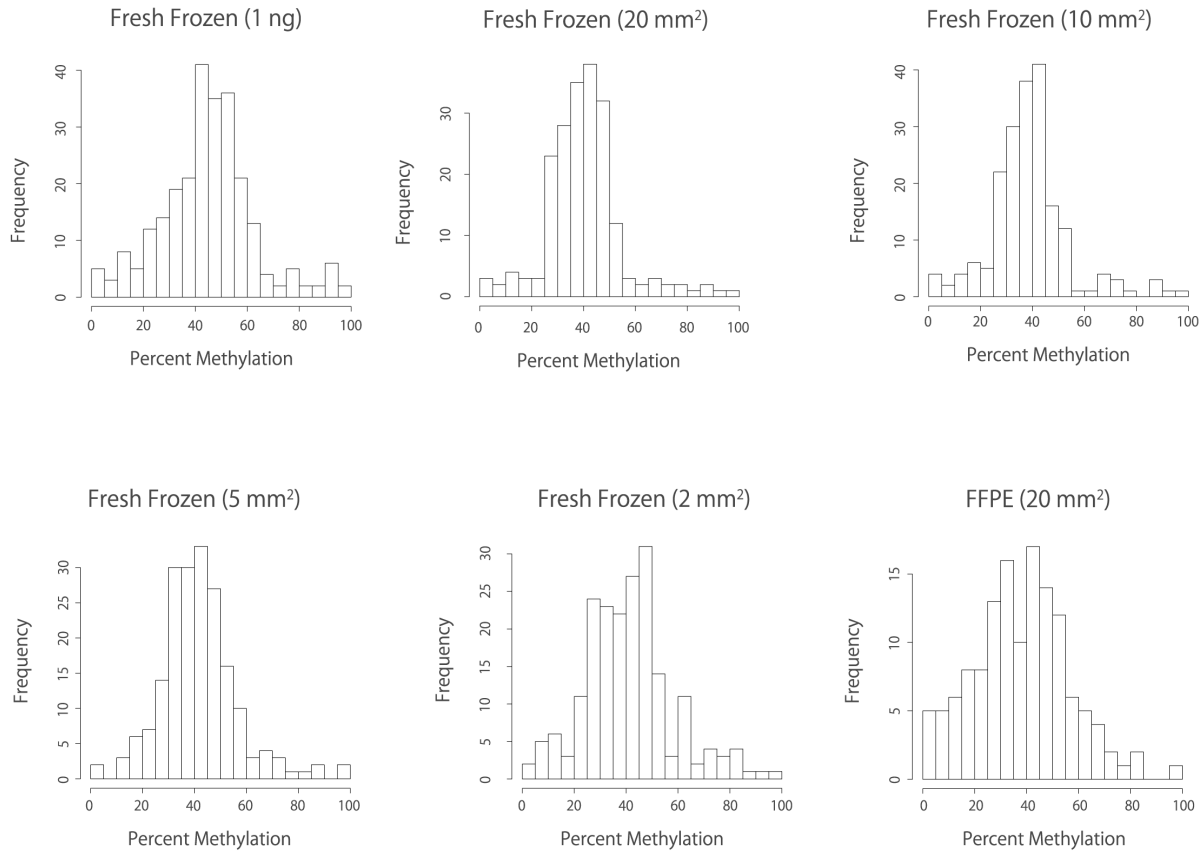

Figure 2.S6 DNA methylation distribution of CpG Islands on chromosome X. The DNA methylation status of all UCSC annotated CpG islands across the X chromosome was determined for 1 ng of fresh frozen DNA extracted from bulk liver tissue and fresh frozen and FFPE samples collected by LCM to assess the presence of PCR bias. Due to X-inactivation, we expect the majority of CpG Islands to be moderately methylated (30-70%), as observed, confirming that LCM-RRBS is largely immune to PCR bias down to 2 mm^2 of fresh frozen and 20 mm^2 of FFPE tissue.

TABLE 2.S1

http://nar.oxfordjournals.org/content/suppl/2013/03/27/gkt230.DC1/nar-01911-met-h-2012-File009.xlsx

Table 2.S1 DNA Methylation of core promoters (2kb centered around the TSS) across normal and neoplasia adrenocortical samples. Percent DNA methylation of annotated CpG Islands. ND: Not determined. The URL address above contains a link to Table 2.1S.

TABLE 2.S2

http://nar.oxfordjournals.org/content/suppl/2013/03/27/gkt230.DC1/nar-01911-met-h-2012-File010.xlsx

Table 2.S2 Gene promoter DNA methylation analysis of normal and neoplasia tissue. Promoters are defined as the regions within 1 kb upstream and downstream of the transcription start site. This table lists the genes and respective DNA methylation levels for those that pasted multiple hypothesis corrections (FDR adjusted). The URL address above contains a link to Table 2.2S.

CHAPTER 3: NOVEL MARKERS OF GONADECTOMY-INDUCED ADRENOCROTICAL NEOPLASIA IN THE MOUSE AND FERRET

In review at Molecular Cellular Endocrinology (2014).

ABSTRACT

Gonadectomy (GDX) induces sex steroid-producing adrenocortical tumors in certain mouse strains and in the domestic ferret. Complementary approaches, including DNA methylation mapping and microarray expression profiling, were used to identify novel genetic and epigenetic markers of GDX-induced adrenocortical neoplasia in female DBA/2J mice. Markers were validated using a combination of laser capture microdissection, quantitative RT-PCR, *in situ* hybridization, and immunohistochemistry. Two genes with hypomethylated promoters, *Igfbp6* and *Foxs1*, were upregulated in post-GDX adrenocortical neoplasms. The neoplastic cells also exhibited hypomethylation of the fetal adrenal enhancer of *Sf1*, an epigenetic signature that typifies descendants of fetal adrenal cells. Expression profiling demonstrated upregulation of gonadal-like genes, including *Spinlw1*, *Insl3*, and *Foxl2*, in GDX-induced adrenocortical tumors of the mouse. One of these markers, FOXL2, was detected in adrenocortical tumor specimens from gonadectomized ferrets. These new markers may prove useful for studies of steroidogenic cell development and for diagnostic testing.

INTRODUCTION

Steroidogenic cells in the adrenal cortex and gonads arise from a common pool of progenitors in the adrenogonadal primordia, but the mechanisms that determine whether a given precursor cell adopts an adrenocortical or gonadal phenotype are not fully understood (Val et al. 2006; Hu et al. 2007; Morohashi and Zubair 2011; Wood and Hammer 2011; Laufer et al. 2012; Shima et al. 2012; Simon and Hammer 2012; Bandiera et al. 2013; Pihlajoki et al. 2013b; Wood et al. 2013). One experimentally tractable model for the study of steroidogenic cell fate determination is gonadectomy (GDX)-induced adrenocortical neoplasia. In response to GDX and the ensuing changes in serum hormone levels [↑ luteinizing hormone (LH), ↓ inhibins, etc.], sex steroid-producing tumors arise in the adrenal glands of certain mouse strains and ferrets (Bielinska et al. 2005; Bielinska et al. 2006a; Johnsen et al. 2006a; Bernichtein et al. 2007; Bernichtein et al. 2008a; Bernichtein et al. 2009; Doghman and Lalli 2009; Beuschlein et al. 2012; Miller et al. 2013). This phenomenon is thought to reflect gonadotropin-induced metaplasia of stem/progenitor cells in the adrenal capsule or cortex, although the term “neoplasia” is used more often than “metaplasia” to describe the process (Bielinska et al. 2006a). The neoplastic tissue resembles luteinized ovarian stroma and expresses gonadal-like differentiation markers, including LH receptor (*Lhcgr*), anti-Müllerian hormone (*Amh*) and its receptor (*Amhr2*), inhibin-α (*Inha*), transcription factors *Gata4* and *Wt1*, and enzymes required for sex steroid biosynthesis (*Cyp17a1*, *Cyp19a1*) (Bielinska et al. 2003a; Bielinska et al. 2005; Bielinska et al. 2006a; Johnsen et al. 2006a; Krachulec et al. 2012; Bandiera et al. 2013). Prototypical markers of adrenocortical cell differentiation, such as the ACTH receptor (*Mc2r*) or corticoid biosynthetic enzymes (*Cyp21*, *Cyp11b1*, *Cyp11b2*), are not expressed in the neoplastic tissue (Bielinska et al. 2003a; Bielinska et al. 2005; Bielinska et al. 2006a; Johnsen et al. 2006a).

The genetic basis of GDX-induced adrenocortical neoplasia has been investigated in the mouse. Hypophysectomy, parabiosis, and transplantation experiments have shown that the adrenal glands of susceptible strains of mice exhibit an inherent predisposition to develop tumors in response to gonadotropin stimulation [reviewed in (Bielinska et al. 2005; Bielinska et al. 2006a)]. Linkage analysis of crosses between susceptible (DBA/2J) and non-susceptible (C57Bl/6) inbred strains has established that post-GDX adrenocortical neoplasia is a complex trait influenced by multiple genetic loci (Bernichtein et al.

2007; Bernichtein et al. 2008a; Bernichtein et al. 2009). Targeted mutagenesis of *Gata4*, a gene normally expressed in gonadal but not adrenal steroidogenic cells of the adult mouse, attenuates post-GDX adrenocortical tumor formation in susceptible strains (Krachulec et al. 2012), and transgenic expression of *Gata4* induces adrenocortical neoplasia in a non-susceptible strain (Chrusciel et al. 2013).

In addition to genetic factors, epigenetic changes such as DNA methylation may contribute to the pathogenesis of GDX-induced adrenocortical neoplasia. Altered methylation of cytosine residues in CpG dinucleotides has been shown to modulate gene expression and progenitor cell fate in various tissues, including endocrine organs (Aranda et al. 2009; Hoivik et al. 2011). For example, conditional mutagenesis of the mouse *Dnmt1* gene, which encodes the maintenance DNA methyl-transferase, causes reprogramming of pancreatic β-cells into α-cells (Akerman et al. 2011; Dhawan et al. 2011a). It has been suggested that GDX-induced adrenocortical neoplasia may be another example of DNA methylation-regulated cell fate interconversion in an endocrine tissue (Bielinska et al. 2009; Schillebeeckx et al. 2013). According to this hypothesis, epigenetic alterations affect the phenotypic plasticity of adrenocortical stem/progenitor cells, allowing them to respond to the hormonal changes associated with GDX (Feinberg et al. 2006b; Bielinska et al. 2009).

The current study was undertaken to identify novel genetic and epigenetic markers of GDX-induced adrenocortical neoplasia, so as to gain a better foothold for investigations into the mechanistic basis of tumorigenesis. Complementary approaches, including genome-wide DNA methylation mapping and microarray expression profiling, were used to screen for genes that are hypomethylated and/or overexpressed in post-GDX adrenocortical neoplasms of the mouse. Candidate genes were validated using a combination of laser capture microdissection (LCM), quantitative RT-PCR (qRT-PCR), and *in situ* hybridization or immunohistochemistry. One of the validated genes was found to be a marker of post-GDX adrenocortical neoplasia in not only mice but also ferrets.

RESULTS AND DISCUSSION

Genes identified as hypomethylated by genome-wide analysis are upregulated in GDX-induced adrenocortical neoplasms of the mouse

In a prior report, we used a highly sensitive method of global DNA methylation analysis, termed LCM-RRBS, to identify gene promoters that are differentially methylated in neoplastic *vs*. normal

adrenocortical tissue from ovariectomized DBA/2J mice (Schillebeeckx et al. 2013). We reasoned that promoters that are hypomethylated in the neoplastic tissue could be novel markers of GDX-induced adrenocortical tumorigenesis. In the current study, we subjected these hypomethylated candidate genes to a series of validation studies. To be considered a *bona fide* marker of GDX-induced adrenocortical neoplasia, we stipulated that three criteria had to be met: 1) qRT-PCR analysis of mRNA isolated from whole adrenal glands had to show a significant increase in expression of the gene in gonadectomized *vs*. intact mice**,** 2) qRT-PCR analysis of mRNA isolated by LCM had to show a significant increase in expression of the gene in neoplastic *vs*. adjacent normal adrenocortical tissue, and 3) *in situ* hybridization or immunohistochemistry of adrenal tissue from gonadectomized mice had to demonstrate expression in the neoplastic cells. Of the 37 hypomethylated genes identified in the genome-wide screen, we assessed the expression difference for 31 genes using whole adrenal qRT-PCR (criteria 1), 6 genes using LCM qRT-PCR (criteria 2), and 5 genes using *in situ* staining or immunohistochemistry (criteria 3). Two hypomethylated genes, *Igfbp6* and *Foxs1*, fulfilled all three criteria (Figure 3.1A-F).

Igfbp6 encodes a member of a family of insulin-related growth factor binding proteins (IGFBPs), which modulate the interaction of insulin-related growth factors (IGFs) with their cell surface receptors. IGFBP6 differs from the other 5 members of the IGFBP family in that it has a 30- to 100-fold preferential binding affinity for IGF2 over IGF1 (Bach et al. 1993). Both IGF2 and IGF1 impact the growth, differentiation, and function of adrenocortical cells (Weber et al. 1999), and IGFBP6 has been shown to inhibit the actions of IGF2 in experimental systems (Bach et al. 1994). Hormone-dependent *Igfbp6* expression has been documented in somatic cells of the rodent ovary (Rohan et al. 1993). It is conceivable that IGFBP6 produced in post-GDX adrenocortical neoplasms serves an insulator function by blocking the activity of IGF2, thereby favoring gonadal-like differentiation over adrenocortical differentiation.

Foxs1 (*Fkhl18*) encodes a forkhead-domain transcription factor that is expressed in Sertoli cells and periendothelial cells of the developing mouse fetal testis (Sato et al. 2008). Male and female *Foxs1* knockout mice are viable and fertile, but the mutant males accumulate blood in the fetal testis, presumably due to apoptosis of periendothelial cells (Sato et al. 2008). *Foxs1* is also expressed in gonadal-like cells that accumulate in the adrenal subcapsule of *Gata6*$^{flox/flox}$; *Sf1*-cre mice (Pihlajoki et al.

2013a). In addition to gonadal(-like) cells, *Foxs1* is expressed in neural crest derivatives (Heglind et al. 2005; Montelius et al. 2007); consistent with this notion, *in situ* hybridization showed strong *Foxs1* expression in the neoplastic adrenocortical cells and weak expression in chromaffin cells of the adrenal medulla (Figure 3.1E).

That only two of the hypomethylated genes identified in the original screen fulfilled all three validation criteria is not surprising. Hypermethylation is correlated with increased gene expression, but the relationship is not absolute. Whole adrenal qRT-PCR may not detect small, but significant, expression differences. The small amounts of RNA isolated by LCM precluded qRT-PCR validation of a large number of genes.

Locus-specific DNA methylation analysis of a fetal adrenal enhancer in GDX-induced adrenocortical neoplasms

LCM-RRBS does not interrogate all methylated loci. To supplement the genome-wide screen, we performed locus-specific DNA methylation analysis of an intronic enhancer of steroidogenic factor-1 (*Sf1*, *AdBP4*, *Nr5a1*), a transcription factor that regulates steroidogenic cell differentiation in the adrenal cortex and gonads (Buaas et al. 2012). The fetal adrenal enhancer (FAdE) of *Sf1* (Zubair et al. 2006; Zubair et al. 2008; Morohashi and Zubair 2011) is hypomethylated both in the fetal adrenal and adult adrenal cortex (Hoivik et al. 2011; Hoivik et al. 2013). In contrast, the FAdE is hypermethylated in other tissue types, including tissues that express *Sf1* (ventromedial nucleus of the hypothalamus, pituitary gonadotropes) and those that do not (pituitary nongonadotropes, liver) (Hoivik et al. 2011). The adrenal cortex of the adult mouse is derived from fetal adrenal cells in which the FAdE was transiently active during development (Zubair et al. 2008; Morohashi and Zubair 2011), implying that once demethylation has occurred in the fetal adrenal, the methylation pattern is maintained during subsequent cell divisions and conserved in the adult cortex. We hypothesized that the methylation status of the *Sf1* FAdE in post-GDX adrenocortical neoplastic cells could serve as an epigenetic marker for the neoplastic tissue. To assess the DNA methylation status of the *Sf1* FAdE, we performed bisulfite-specific PCR on tissue isolated by LCM (Figure 3.2). Consistent with published reports (Hoivik et al. 2011; Hoivik et al. 2013), the FAdE was hypermethylated in the liver and hypomethylated in both the adrenal X-zone [a remnant of the fetal

adrenal (Morohashi and Zubair 2011)] and the adult adrenal cortex (zF + zG). Importantly, the FAdE was also hypomethylated in post-GDX adrenocortical neoplasms.

To determine whether FAdE methylation status could distinguish the embryonic origin of the neoplastic cells (fetal adrenal *vs*. ectopic ovarian cells), we measured FAdE methylation in whole ovary. Ovarian tissue exhibited an intermediate level of FAdE methylation that was statistically greater than that of adult adrenal cortex ($P < 0.01$) and X-zone ($P < 0.005$) but not that of post-GDX tumors ($P = 0.085$, not significant). We conclude that GDX-induced tumor cells, like adjacent normal adrenocortical cells, carry an epigenetic mark (FAdE hypomethylation) that typifies descendants of fetal adrenal cells; however, the methylation status of this locus cannot exclude a gonadal origin for the neoplastic cells.

RNA expression profiling identifies novel markers of GDX-induced adrenocortical neoplasia in the mouse

Since differential DNA methylation is only one of several mechanisms that can influence gene expression, we performed microarray hybridization to screen for additional transcripts that are upregulated during GDX-induced adrenocortical tumorigenesis. For this analysis we used mRNA isolated from whole adrenal glands of ovariectomized *vs*. intact DBA/2J mice. We identified 89 and 38 genes that were upregulated and downregulated, respectively, in adrenal glands from gonadectomized mice (Supplemental Table 1). We hypothesized that the upregulated genes were enriched for markers of gonadal tissue. To test this hypothesis, genes that were differentially expressed in the adrenal glands of gonadectomized mice were compared to pooled GEO microarray data for different mouse tissues, using an established algorithm (Chen et al. 2013). Enrichment analysis demonstrated that genes upregulated in the adrenal glands of gonadectomized mice were more likely to be highly expressed in ovary (Figure 3.3A; $P = 2.01 \times 10^{-8}$) or testis (Figure 3.3B; $P = 2.16 \times 10^{-7}$) but not in the non-steroidogenic tissue brain (Figure 3.3D; $P = 0.90$). Genes downregulated in the adrenal glands of gonadectomized mice were more likely to be highly expressed in the normal adrenal tissue (Figure 3.3C; $P = 7.83 \times 10^{-4}$) than in brain (Figure 3.3E; $P = 1.00$). This systematic, transcriptomic comparison reinforces the longstanding tenet that GDX induces the selective accumulation of gonadal-like cells in the adrenal glands of DBA/2J mice.

Among the genes upregulated in the microarray analysis were gonadal-like markers known to be expressed in post-GDX adrenocortical neoplasms, such as *Cyp17a1* (460-fold), *Lhcgr* (43-fold), *Inha* (17-fold), and *Amhr2* (6.2- to 9.1-fold on different microarray probes) (Bielinska et al. 2006a). Two well-documented markers of GDX-induced adrenocortical neoplasia, *Gata4* and *Amh* (Bielinska et al. 2006a), were not upregulated, underscoring the inherent limitations of microarray technology (each of these genes was represented by a single probe on the microarray and showed a non-significant expression change of ~0.5-fold). Two genes identified in the microarray screen, *Spinlw1* and *Insl3*, were validated as novel markers of GDX-induced adrenocortical neoplasia in the mouse, using the same rigorous criteria described above (Figure 3.4A-E).

Spinlw1, which was upregulated 16- to 51-fold on different microarray probes, encodes EPPIN, a serine protease inhibitor secreted by Sertoli cells and epididymal epithelial cells (O'Rand et al. 2011) but not by cells of the ovary or uterus (Sivashanmugam et al. 2003; Silva et al. 2012). Thus, EPPIN is generally considered to be a "male-specific" marker. Consistent with this notion, expression of *Spinlw1* has been shown to be androgen-dependent (Denolet et al. 2006; Schauwaers et al. 2007; Willems et al. 2010; Silva et al. 2012).

Insl3, which was upregulated 50-fold on the microarray, encodes insulin-like 3, a hormone that is secreted by fetal Leydig cells. INSL3 mediates the trans-abdominal phase of testicular descent (Ivell et al. 2013), and loss-of-function mutations in *Insl3* cause cryptorchidism (Ivell and Anand-Ivell 2011). Additionally, INSL3 is constitutively secreted by adult Leydig cells and serves as a serum biomarker of this cell type (Anand-Ivell et al. 2009), although the function of INSL3 in the adult testis remains unclear (Ivell et al. 2013). *Insl3* is also expressed in the ovary, particularly in theca interna cells of the maturing follicle (Satchell et al. 2013), where it induces androgen production (Glister et al. 2013). It is conceivable that serum levels of INSL3 could be a biomarker of post-GDX adrenocortical neoplasms, and future experiments will explore this possibility.

Foxl2, a gene that was marginally upregulated (1.3-fold) in the microarray analysis, was found to be another *bona fide* marker of GDX-induced adrenocortical neoplasia in the mouse (Figure 3.5A-F). *Foxl2* encodes a forkhead transcription factor that is expressed in granulosa and interstitial cells of the ovary (Schmidt et al. 2004; Mork et al. 2012). Consequently, FOXL2 is generally considered to be a

"female-specific" marker (Georges et al. 2013). Mice harboring null mutations in *Foxl2* develop ovaries that express testicular differentiation markers (Schmidt et al. 2004; Uda et al. 2004; Ottolenghi et al. 2005; Garcia-Ortiz et al. 2009; Uhlenhaut et al. 2009). Of note, extracts of adrenal glands from non-gonadectomized mice have been shown to contain FOXL2 mRNA and protein at levels 40-50 times lower than extracts of ovary (Yang et al. 2010). The significance of this low level expression of *Foxl2* in non-neoplastic adrenal glands is unclear.

None of these validated, gonadal-like genes showed evidence of differential methylation in the LCM-RRBS analysis (Schillebeeckx et al. 2013). The degree of methylation of the *Foxl2* promoter was <1% in both neoplastic and normal tissue, and the degree of methylation of the *Insl3* promoter was comparable in neoplastic and normal tissue (13% *vs*. 16%; σ = 4.1 and σ = 5.3, respectively). Coverage of the *Spinlw1* promoter was inadequate for quantification of its methylation status.

In addition to gonadal-like markers, 3 mast cell enzyme genes, *Cma1*, *Cma2*, and *Cpa3*, were upregulated (2.5- to 8.4-fold) in the adrenal glands of gonadectomized mice (Supplemental Table 1). This finding is consistent with the well-documented phenomenon of mast cell infiltration of post-GDX adrenocortical neoplasms (Kim et al. 1997; Bielinska et al. 2005).

FOXL2 is a marker of adrenocortical neoplasia in gonadectomized ferrets

Sex steroid-producing neoplasms arise in up to 20% of gonadectomized ferrets and are a major cause of morbidity and mortality in this species (Bielinska et al. 2006a; Beuschlein et al. 2012). Immunostaining of archival veterinary pathology specimens showed that FOXL2 is a marker of post-GDX adrenocortical neoplasia in the ferret (Figure 3.6A-F; Figure 3.S1). The specimens were obtained from gonadectomized ferrets with signs of ectopic sex steroid production (e.g., alopecia, vulvar hyperplasia, or stranguria) and included examples of adrenocortical carcinoma, adenoma, and nodular hyperplasia. Several of the specimens contained residual normal cortex, which served as a negative control (Figure 3.6A, C, E). Nuclear FOXL2 immunoreactivity was detected in 50% of the specimens examined [4 of 10 cases of adrenocortical carcinoma, 4 of 6 cases of adrenocortical adenoma, and 1 of 2 cases of nodular hyperplasia]. FOXL2-positive tumors were seen in both female (Figure 3.6) and male (Figure 3.S1) ferrets. Only a minority of the cells within a given tumor reacted with FOXL2 antibody (Figure 3.6D;

Figure 3.S1). By comparison, nuclear GATA4 immunoreactivity has been observed in >90% of ferret adrenocortical tumors (Peterson et al. 2004), and GATA4 antibody typically stains a higher percentage of cells within a given tumor (Figure 3.6F).

We conclude that FOXL2 is a marker of GDX-induced adrenocortical tumors in not only the mouse but also the ferret. This observation expands the list of genes known to be expressed in sex steroidogenic adrenocortical neoplasms of the ferret (Schoemaker et al. 2002; Peterson et al. 2003; Peterson et al. 2004; Bielinska et al. 2006a; Wagner et al. 2008; de Jong et al. 2013; de Jong et al. 2014). The relatively low percentage of FOXL2 immunoreactive cells within ferret adrenocortical tumors may limit the diagnostic utility of this marker. It remains to be determined whether the other markers of adrenocortical neoplasia that emerged from our screens of gonadectomized mice are also markers of tumorigenesis in ferrets.

Summary

Using complementary approaches, including DNA methylation analysis and microarray expression profiling, we have identified novel epigenetic (*Igfbp6*, *Foxs1*) and genetic (*Spinlw1*, *Insl3*, *Foxl2*) markers of GDX-induced adrenocortical neoplasia in the mouse. That both "male-specific" (*Spinlw1*) and "female-specific" (*Foxl2*) markers were detected is noteworthy and implies that the neoplasms exhibit mixed characteristics of male and female gonadal somatic cells. Such indeterminate steroidogenic cell phenotypes have been reported in other experimental models (Heikkila et al. 2002; Val et al. 2006). One of the markers, FOXL2, was observed in adrenocortical tumor specimens from gonadectomized ferrets. These new markers may prove useful for studies of steroidogenic cell development and for tumor classification.

Highlights

- *Igfbp6* & *Foxs1* are hypomethylated and upregulated in murine post-GDX adrenocortical tumors.
- The *Sf1* FAdE is hypomethylated in murine post-GDX adrenocortical tumors.
- *Spinlw1*, *Insl3*, & *Foxl2* are upregulated in murine post-GDX adrenocortical tumors.
- FOXL2 immunoreactivity is evident in adrenocortical tumors from gonadectomized ferrets.

- Post-GDX adrenocortical tumors exhibit properties of female and male gonadal cells.

ACKNOWLEDGMENTS

We thank members of the histology, laser microscopy, and microarray cores for technical assistance. We thank Michael Brooks for providing assistance with the microarray analysis, Mikko Anttonen for helpful suggestions, and Matti Kiupel for providing ferret tissue specimens.

AUTHOR CONTRIBUTIONS

This work was done in collaboration with Marjut Pihlajoki, Elisabeth Gretzinger, Franziska Tholl, Theresa Hiller, Ann-Kathrin Löbs, Theresa Röhrig, Anja Schrade, Rebecca Cochran, Patrick Jay, Markku Heikinheimo, Rob Mitra, and Dave Wilson. Dave Wilson, Markku Heikinheimo, Rob Mitra, Patrick Jay, Marjut Pihlajoki, and I designed and conceived the experiments. Tissue immunochemistry, LCM, nucleic acid extractions, *in situ* hybridization, and qRT-PCRs were completed by Marjut Pihlajoki, Elisabeth Gretzinger, Franziska Tholl, Theresa Hiller, Ann-Kathrin Löbs, Theresa Röhrig, Anja Schrade, and Rebecca Cochran. I processed gene expression data, completed enrichment analysis, and performed bisulfite specific PCR experiments. Dave Wilson, Rob Mitra, Markku Heikinheimo, Patrick Jay, Marjut Pihlajoki, and I wrote the manuscript. This work is in review for the journal *Molecular Endocrinology* and supported by NIH grants DK52574, DK075618, and DA025744, American Heart Association grant 13GRNT16850031, the Sigrid Jusélius Foundation, the Academy of Finland, and CIMO (Centre for International Mobility Finland) fellowship TM-13-8769.

MATERIALS AND METHODS

Experimental animals

Procedures involving mice were approved by an institutional committee for laboratory animal care and were conducted in accordance with NIH guidelines for the care and use of experimental animals. DBA/2J mice were purchased from Jackson Laboratories (Bar Harbor, ME). Female mice were anesthetized and gonadectomized at 3-4 weeks of age (Bielinska et al. 2005). We limited our analysis to females because they develop post-GDX adrenocortical neoplasms more readily than their male counterparts (Bielinska et al. 2006a; Beuschlein et al. 2012). Adrenal tissue was harvested for analysis 3 months later.

Isolation of neoplastic and normal tissue using LCM

Cryosections (10 μm) of adrenal glands from ovariectomized or intact mice were collected on membrane slides (PEN-Membrane 2.0 μm; Leica), fixed in acetone (for collecting DNA) or ethanol (for collecting RNA) at -20º C, stained with hematoxylin and eosin (H&E) or crystal violet, and then dehydrated by passage through successively higher concentrations of ethanol followed by xylene (Pihlajoki et al. 2013a; Schillebeeckx et al. 2013). LCM was used to isolate samples from GDX-induced adrenocortical neoplasms, adjacent normal adrenocortical tissue [zona glomerulosa (zG) + zona fasciculata (zF) cells], and the adrenal X-zone. Dissectates were collected in SDS/proteinase K for genomic DNA isolation or in RNA extraction buffer (RNeasy Mini Kit, Qiagen, Valencia, CA).

Global and locus-specific DNA methylation analyses

DNA from neoplastic or normal adrenocortical tissue was subjected to genome-wide methylation analysis using LCM-reduced representation bisulfite sequencing (LCM-RRBS) (Schillebeeckx et al. 2013). Locus-specific DNA methylation of the *Sf1* enhancer locus was measured via bisulfite-specific PCR and pyrosequencing as described previously (Schillebeeckx et al. 2013).

Microarray expression profiling

RNA was isolated from whole adrenal glands of intact, virgin (n = 3) or ovariectomized (n = 3) female DBA/2J mice using RNeasy® Mini Kit (Qiagen, Valencia, CA), amplified using the TotalPrep RNA amplification kit (Illumina, San Diego, CA), and hybridized on an Illumina Mouse6v2 oligonucleotide array. Array hybridization was performed by the GTAC Microarray Core facility at Washington University according to standard protocols. Probes with a hybridization signal less than signal background were excluded from the analysis, leaving 15,066 probes. Probes that did not show at least a 2-fold change in either direction were then excluded, leaving 628 probes (Table 3.S1). Finally, we determined statistical significance using the Student's *t* test and correcting for multiple hypothesis testing using the Benjamini-Hochberg method, leaving a total of 127 probes (q-value < 0.05). Statistically significant upregulated (n = 89) and downregulated (n = 38) probes correspond to 85 and 36 genes, respectively (Table 3.S1).

Genes that were differentially expressed in adrenal glands from gonadectomized *vs*. intact mice were compared to pooled Gene Expression Omnibus (GEO) data for different mouse tissues, using non-parametric statistical testing (Chen et al. 2013). Genes with a statistically significant expression specificity (q-value < 0.01; Benjamini-Hochberg corrected) for adrenal, brain, ovary, or testis tissues were considered to be tissue-specific (Table 3.S2). Enrichment for upregulated or downregulated genes within the tissue-specific gene dataset was determined using the Fisher's Exact Test.

qRT-PCR

Total RNA was isolated and subjected to qRT-PCR analysis as described (Slott et al. 1993). Expression was normalized to the housekeeping genes *Actb* and *Gapdh*. Primer pairs are listed in Table 3.3S.

***In situ* hybridization**

Nonradioactive *in situ* hybridization was performed (Val et al. 2006) using paraformaldehyde-fixed, paraffin-embedded adrenal sections (5 µm). To prepare riboprobes, cDNA fragments of *Igfbp6*, *Foxs1*, *Spinlw1*, and *Foxl2* were amplified by RT-PCR (annealing temperature = 52º C) and cloned into the vector pCRII-TOPO (Invitrogen, Carlsbad, CA) using the manufacturer's guidelines. RT-PCR primers

and cDNA fragment sizes are specified in Table 3.S4. Digoxygenin-labeled antisense riboprobes were synthesized from *EcoR*V-linearized plasmids using Sp6 RNA polymerase (Heikinheimo et al. 1994). Bound riboprobe was detected using an alkaline phosphatase conjugated anti-digoxigenin antibody, as described (Val et al. 2006). Sections were subsequently counterstained with nuclear fast red.

Immunohistochemistry

Mouse tissues were fixed overnight in 4% paraformaldehyde in PBS, embedded in paraffin, sectioned (5 μm), and subjected to immunoperoxidase staining (Anttonen et al. 2003). Formalin-fixed, paraffin-embedded, adrenocortical neoplasms from gonadectomized ferrets were obtained from the archives of a veterinary diagnostic laboratory, as reported previously (Peterson et al. 2003; Peterson et al. 2004; Wagner et al. 2008); included were cases of adrenocortical carcinoma, adenoma, and nodular hyperplasia. Criteria for classification of these tumors are listed elsewhere (Peterson et al. 2003). The primary antibodies were: a) rabbit anti-INSL3 (sc-134587; Santa Cruz Biotechnology, Santa Cruz, CA; 1:200 dilution), b) goat anti-GATA4 (sc-1237, Santa Cruz Biotechnology; 1:200 dilution), and c) goat anti-FOXL2 (IMG-3228; Imgenex, San Diego, CA; 1:400 dilution). Secondary antibodies were: a) goat anti-rabbit biotinylated IgG (NEF-813, NEN Life Science, Boston, MA; 1:2000 dilution) and b) donkey anti-goat biotinylated IgG (Jackson Immunoresearch, West Grove, PA; 1:1000 dilution). The avidin-biotin immunoperoxidase system (Vectastain Elite ABC Kit, Vector Laboratories, Inc., Burlingame, CA) and diaminobenzidine were used to visualize the bound antibody. The analysis included negative control studies in which the primary antibodies were omitted.

Data Release

The DNA methylation data generated for the study can be found under GEO accession number GSE45361. DNA methylation and raw sequence data are also publically available at the Center for Genome Sciences (www.cgs.wustl.edu/~maxim/). The microarray hybridization data has been deposited under GEO accession number GSE54393.

FIGURE 3.1

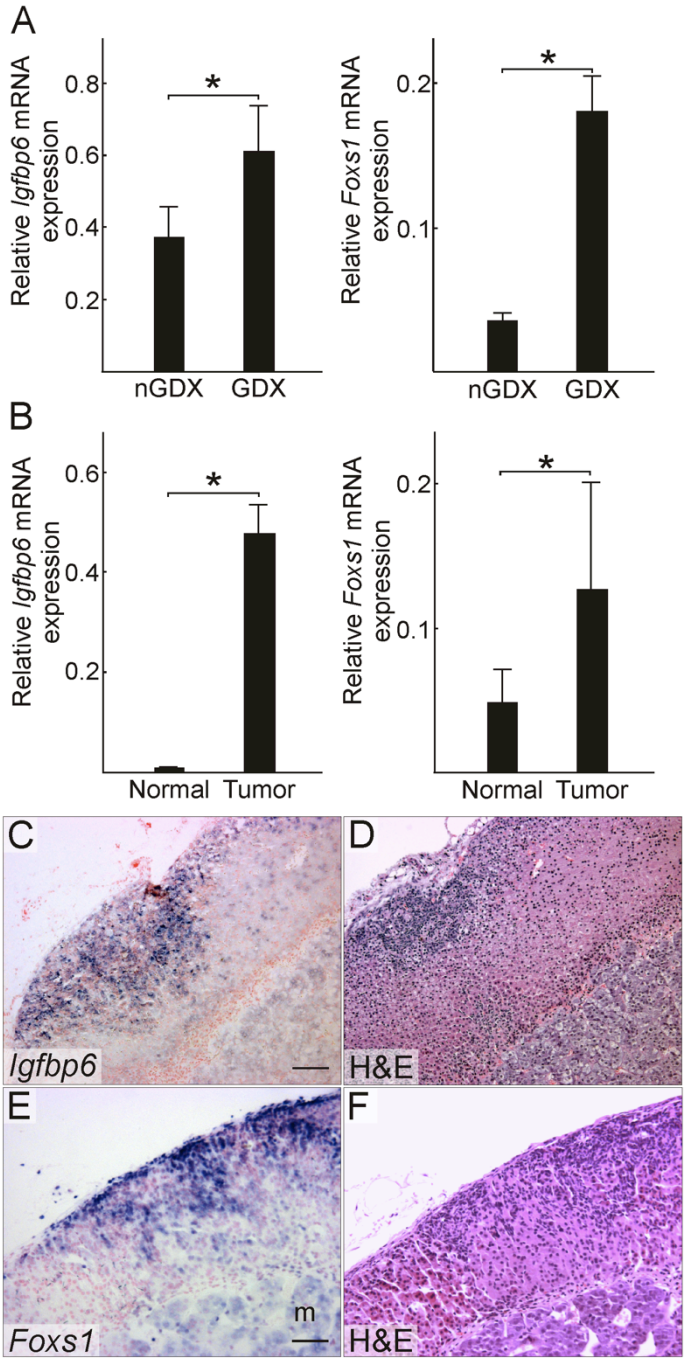

Figure 3.1 Two genes identified as hypomethylated by genome-wide mapping, *Igfbp6* and *Foxs1*, are upregulated in GDX-induced adrenocortical neoplasms of the mouse. (**A**) qRT-PCR analysis of *Igfbp6* and *Foxs1* in whole adrenal mRNA from non-gonadectomized (nGDX) or gonadectomized (GDX) female DBA/2J mice (n = 4 per group). (**B**) qRT-PCR analysis of *Igfbp6* and *Foxs1* in normal or tumor tissue isolated by LCM from the adrenal cortex of ovariectomized DBA/2J mice (n = 4 per group). qRT-

PCR results were normalized to expression of *Actb* (*, $P < 0.05$); normalization to *Gapdh* yielded similar results. Statistical significance was determined using the Student's *t* test. (**C**) *In situ* hybridization of adrenal tissue from an ovariectomized DBA/2J mouse using an *Igfbp6* antisense riboprobe. (**D**) H&E staining of an adjacent tissue section. *Igfbp6* mRNA localized to neoplastic cells in the subcapsular region. (**E**) *In situ* hybridization of adrenal tissue from an ovariectomized DBA/2J mouse using a *Foxs1* antisense riboprobe. (**F**) H&E staining of an adjacent tissue section. *Foxs1* mRNA localized to neoplastic tissue in the subcapsular region and to cells in the medulla (m). Bars = 50 μm.

FIGURE 3.2

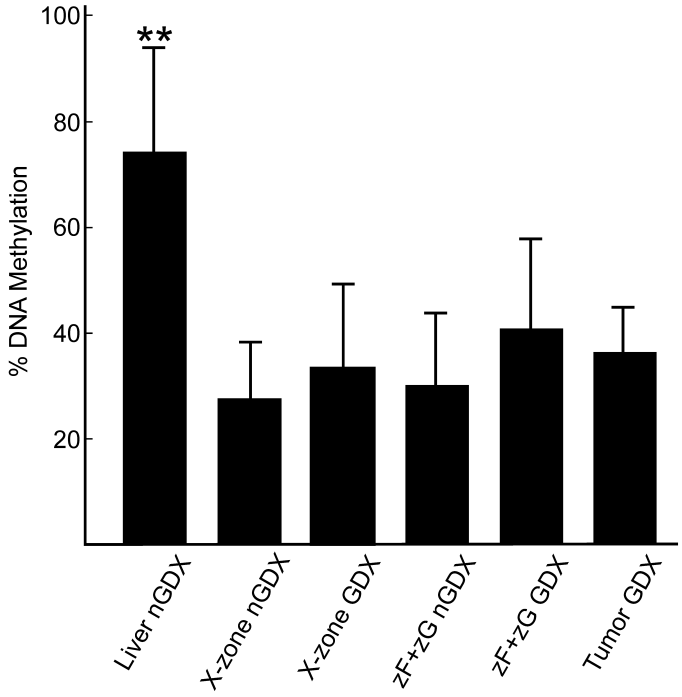

Figure 3.2 The *Sf1* FAdE is hypomethylated in normal and neoplastic adrenocortical cells. Tissue from female DBA/2J mice was subjected to DNA methylation analysis using bisulfite-specific PCR. (**A**) The graph compares the level of methylation across 7 of 8 CpG dinucleotides in the *Sf1* FAdE for various tissues. The 8th CpG (CpG-2) was omitted from the analysis because the data did not pass quality control measures for pyrosequencing. (**B**) The inset shows *P* values for paired *t*-tests among the different tissue types. Synthetic DNA is included as a negative control. Note that the FAdE is hypomethylated in the X-zone, normal cells of the zona fasciculata + zona glomerulosa (zF + zG), and post-GDX adrenocortical tumor tissue.

FIGURE 3.3

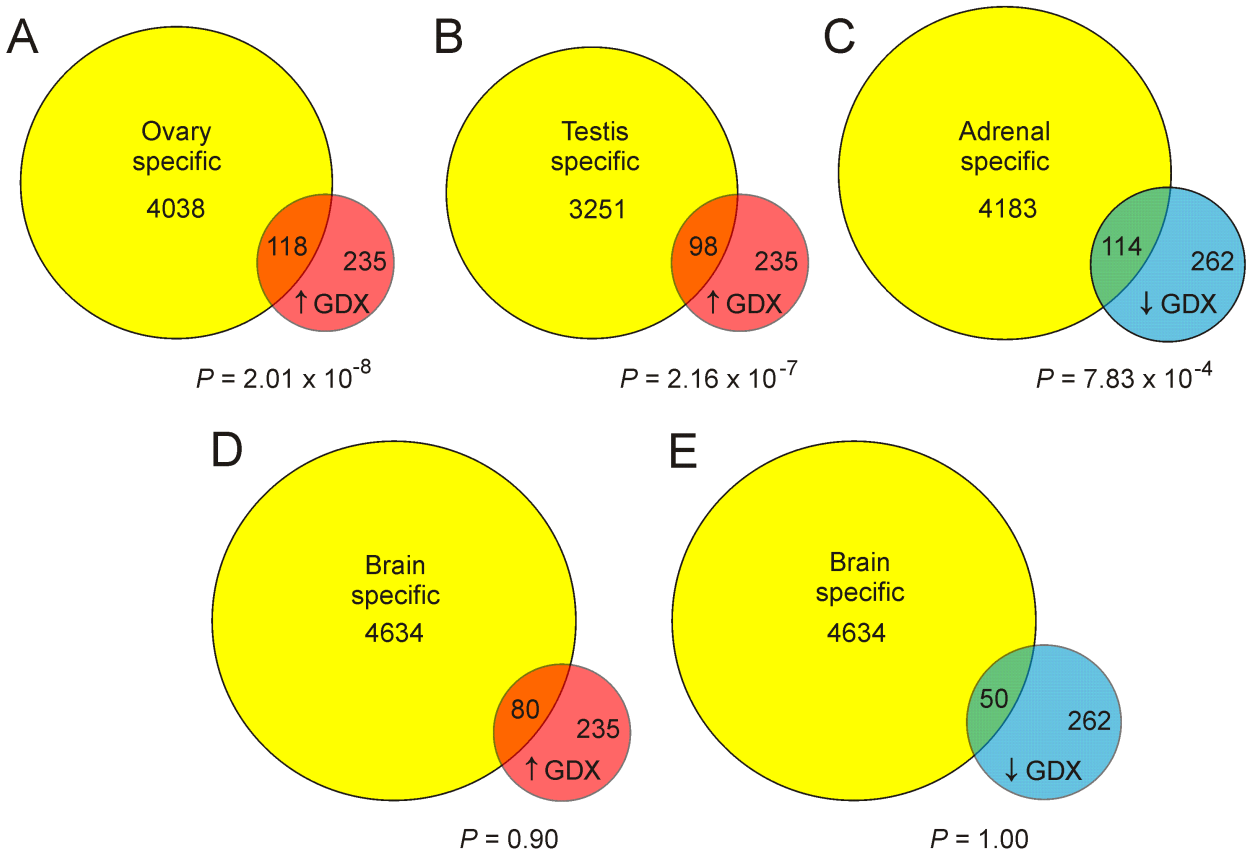

Figure 3.3 Systematic transcriptome analysis shows that genes upregulated in the adrenal glands of gonadectomized mice are more likely to be highly expressed in ovary or testis than in other tissues. Probes that were differentially expressed in microarrays of whole adrenal glands from gonadectomized *vs*. non-gonadectomized female DBA/2J mice were compared to pooled GEO microarray data for different mouse tissues using an established rank sum algorithm (Chen et al. 2013). Probes specific for ovary, testis, adrenal or brain are shown as yellow circles. Probes that were upregulated in adrenal glands from gonadectomzied mice are shown as red circles (↑ GDX). Probes that were downregulated in adrenal glands from gonadectomzied mice are shown as blue circles (↓ GDX). Enrichment analysis demonstrated that genes upregulated in the adrenal glands of gonadectomized mice were more likely to be highly expressed in ovary (**A**) or testis (**B**) than in brain (**D**). Genes downregulated in the adrenal glands of gonadectomized mice were more likely to be highly expressed in the normal adrenal tissue (**C**) than in brain (**E**).

FIGURE 3.4

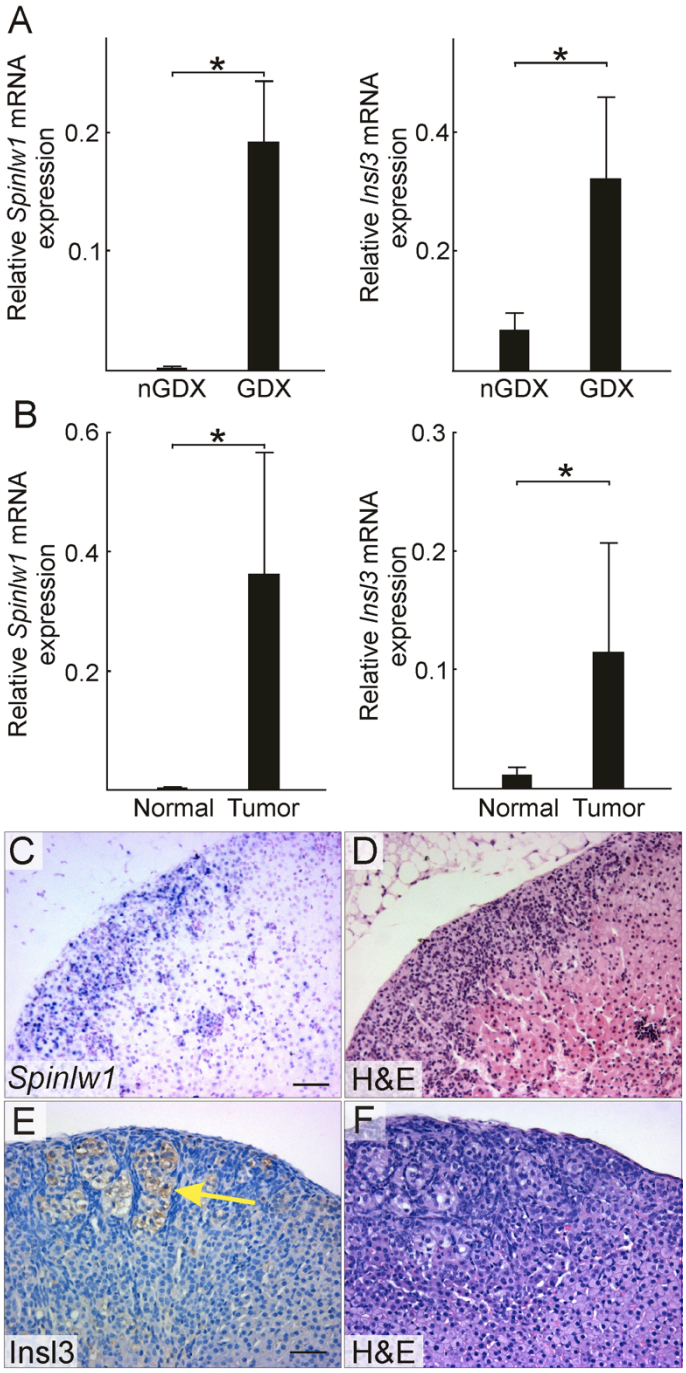

Figure 3.4 Two genes identified by microarray expression profiling, *Spinlw1* and *Insl3*, are upregulated in GDX-induced adrenocortical neoplasms of the mouse. (**A**) qRT-PCR analysis of Spinlw1 and Insl3 in whole adrenal mRNA from non-gonadectomized (nGDX) or gonadectomized (GDX) female DBA/2J mice (n = 4 per group). (**B**) qRT-PCR analysis of Spinlw1 and Insl3 in normal or tumor

tissue isolated by LCM from the adrenal cortex of ovariectomized DBA/2J mice (n = 4 per group). qRT-PCR results were normalized to expression of the housekeeping gene Actb (*, $P < 0.05$); normalization to Gapdh yielded similar results. Statistical significance was determined using the Student's t test. (**C**) In situ hybridization of adrenal tissue from an ovariectomized DBA/2J mouse using a Spinlw1 antisense riboprobe. (**D**) H&E staining of an adjacent tissue section. Spinlw1 mRNA localized to neoplastic cells in the subcapsular region. (C) Immunoperoxidase staining of adrenal tissue from an ovariectomized DBA/2J mouse using anti-INSL3. (D) H&E staining of an adjacent tissue section. INSL3 immunoreactivity was evident in lipid-laden type B neoplastic cells [see (Bielinska et al. 2006a) for a description of this cell type]. Control experiments demonstrated INSL3 immunoreactivity in Leydig cells of the adult testis (data not shown). Bars = 50 μm.

FIGURE 3.5

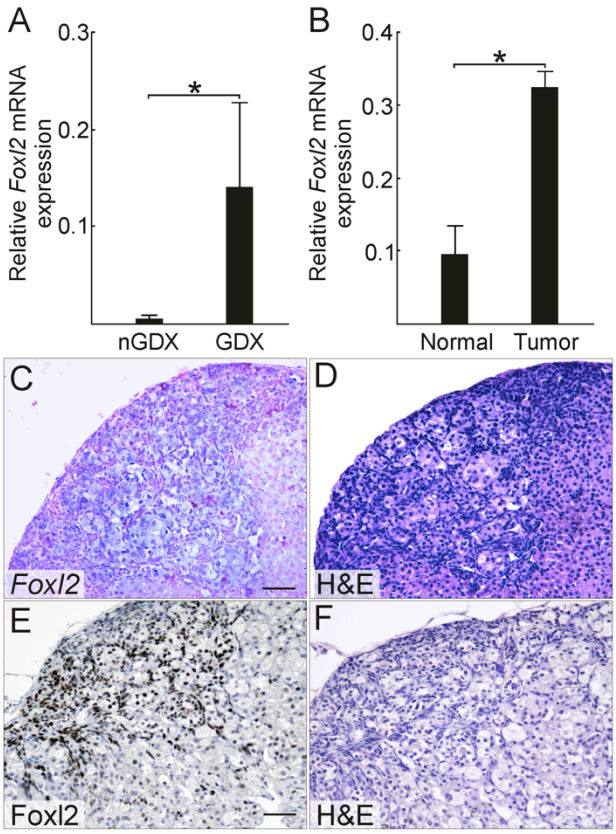

Figure 3.5 *Foxl2* is upregulated in GDX-induced adrenocortical neoplasms of the mouse. (**A**) qRT-PCR analysis of *Foxl2* in whole adrenal mRNA from non-gonadectomized (nGDX) or gonadectomized (GDX) female DBA/2J mice (n = 4 per group). (**B**) qRT-PCR analysis of *Foxl2* in normal or tumor tissue isolated by LCM from the adrenal cortex of ovariectomized DBA/2J mice (n = 4 per group). qRT-PCR results were normalized to expression of the housekeeping gene *Actb* (*, $P < 0.05$); normalization to *Gapdh* yielded similar results. Statistical significance was determined using the Student's *t* test. (**C**) *In situ* hybridization of adrenal tissue from an ovariectomized DBA/2J mouse using a *Foxl2* antisense riboprobe. (**D**) H&E staining of an adjacent tissue section. Bar = 50 µm. (**E**) Immunoperoxidase staining of adrenal tissue from an ovariectomized DBA/2J mouse using anti-FOXL2. (**F**) H&E staining of an adjacent tissue section. Nuclear FOXL2 immunoreactivity was evident in both small, basophilic type A neoplastic cells and large, lipid-laden type B neoplastic cells [see (Bielinska et al. 2006a) for a description of these cell types]. Control stainings demonstrated nuclear FOXL2 immunoreactivity in granulosa cells of the adult ovary (data not shown). Bar = 50 µm.

FIGURE 3.6

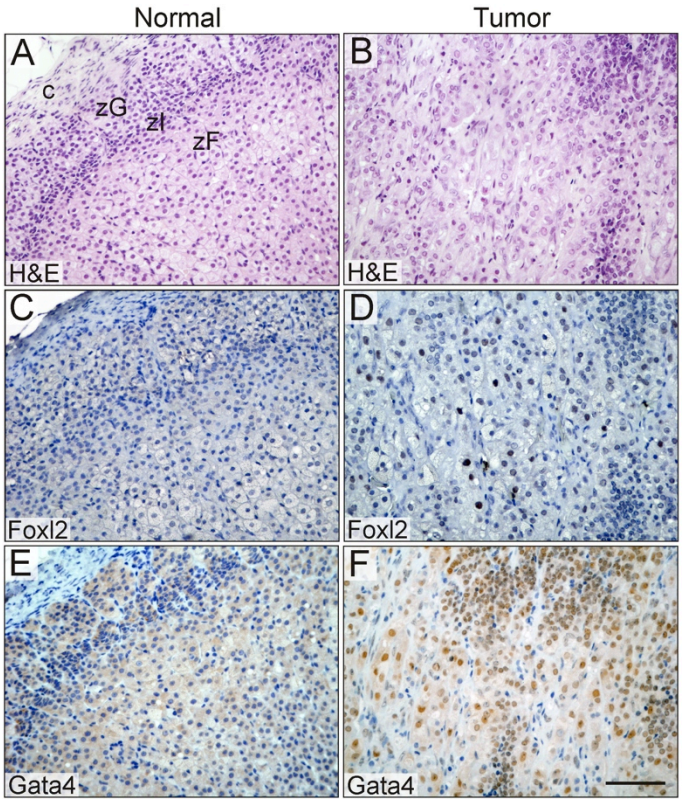

Figure 3.6 FOXL2 immunoreactivity in neoplastic adrenocortical cells from a gonadectomized female ferret with clinical evidence of ectopic sex steroid production. Sections of normal (A,C,E) or neoplastic (B,D,F) adrenal cortex were subjected to H&E staining (A,B) or to immunoperoxidase staining with anti-FOXL2 (C,D) or anti-GATA4 (E,F). The neoplastic tissue contained a mixture of small, basophilic cells and large, polyhedral cells. Nuclear GATA4 immunoreactivity was evident in both the small and large neoplastic cells, whereas nuclear FOXL2 immunoreactivity was limited to the large neoplastic cells in this specimen. Bar = 50 μm. Abbreviations: c, capsule; zF, zona fasciculata; zG, zona glomerulosa; zI, zona intermedia [a zone characteristic of ferrets and other carnivores (Holmes 1961)].

SUPPLEMENTARY FIGURES AND TABLES

FIGURE 3.S1

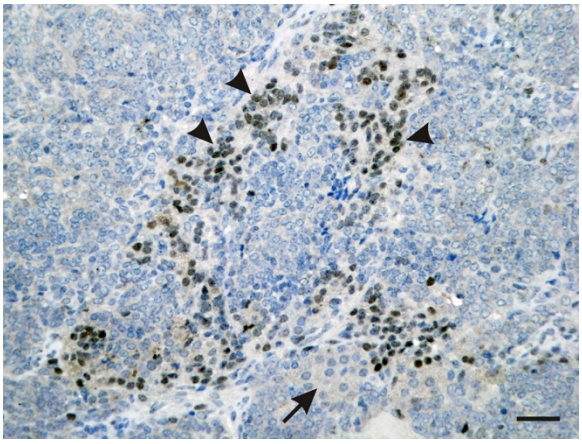

Figure 3.S1 FOXL2 immunoreactivity in an adrenocortical tumor from a gonadectomized male ferret with clinical signs of ectopic sex steroid production. Nuclear FOXL2 staining is evident a subset of the neoplastic cells (arrowheads) but not in normal adrenocortical cells (arrow). Bar = 50 µm.

TABLE 3.S1

https://cgs.wustl.edu/~maxim/Thesis/Chapter3/Table3.S1.xlsx

Table 3.S1 Microarray expression profiling of adrenal glands from gonadectomized (GDX) *vs*. non-gonadectomized (nGDX) female DBA/2J mice. Genes that were significantly upregulated or downregulated (> 2-fold) are listed in the first tab. Genes that survived the Benjamini-Hochberg correction are listed in the second tab. See *Materials and Methods* for details.

TABLE 3.S2

https://cgs.wustl.edu/~maxim/Thesis/Chapter3/Table3.S2.xlsx

Table 3.S2 Genes deemed tissue-specific for adrenal, brain, ovary, and testis based on analysis of pooled GEO microarray data. See *Materials and Methods* for details.

TABLE 3.S3

Gene	Primer sequence (5' → 3')	cDNA size (bp)	Reference
Actb	F: GCGTGACATCAAAGAGAAGC R: AGGATTCCATACCCAAGAAGG	187	NM_007393.2
Foxl2	F: GCAAGGGAGGCGGGACAACAC R: GAACGGGAACTTGGCTATGATGT	154	NM_02020.2
Foxs1	F: TACCTCGCCCTCACCGTGCC R: CAAGGCCTGGGTCAGTCCCCA	167	NM_010226
Gapdh	F: GCTCACTGGCATGGCCTTCCGTG R: TGGAAGAGTGGGAGTTGCTGTTGA	200	NM_008084.2
Igfbp6	F: CAGAGACCGGCAGAAGAATC R: GCTTCCTTGACCATCTGGAG	289	NM_008344
Insl3	F: CACGCAGCCTGTGGAGACCC R: CGCTGGCGCTGAGAAGCCT	134	NM_013564.7
Spinlw1	F: TGACTTGCTGTTTCCCAGGAG R: AAGCCATACAGTAGCCGGAG	191	NM_029325.2

Table 3.S3 Primers used for qRT-PCR.

TABLE 3.S4

Gene	Primer sequence (5' → 3')	Size (bp)
Spinlw1	F: GTGCTATTTGGCCTGCTTGC R: TAGGACCCCACAACTGGGAA	513
Foxl2	F: CAAGTACCTGCAATCGGGGT R: TGCGTCTCAGACACTTCGAC	490
Foxs1	F: TCACCGTGCCCAGCATTCGG R: CCCCAAGGGACCTGCCTGACT	513
Igfpb6	F: TGTTGGTTCGTTGCGGGCTCA R: CCTGCGAGGAACGACCTGCTG	561

Table 3.S4 Primers used to generate riboprobes for *in situ* hybridization.

CHAPTER 4: ACTIVE METHYLATION AND DEMETHYLATION IN MOTOR NEURON MATURATION

ABSTRACT

Cytosine methylation (5mC) and hydroxymethylation (5hmC) play important roles in transcriptional regulation in post-mitotic neurons. Neuronal activity-induced gene activation in the adult brain is mediated through the Tet oxidation of 5mC to 5hmC, which is replaced with unmodified cytosine by the base excision repair pathway to complete the demethylation cycle. Little is known about the dynamics of how 5mC and 5hmC are established in adult neurons partially owing to the difficulty of studying neuron maturation *ex vivo*. Here, we leverage a recently developed 5mC genome-wide mapping technique and a motor neuron puromycin-selectable transgenic embryonic stem cell line to characterize the DNA methylome of motor neurons (MNs) as they differentiate and mature. We show that motor neuron differentiation and maturation is characterized by a loss of 5mC at genes critical for neuron function and by a global gain in 5hmC. These changes in cytosine modification result in attendant gene activation and show enrichment for known sequence motifs of motor neuron-specific transcription factors. Furthermore, single, double, and triple knockout experiments show that individual Tet proteins are not required for MN maturation and that *Tet3* is sufficient for proper MN differentiation and maturation, and that MN can differentiate and form without the presence of any TET enzymes. Our study suggests that transcription factor-mediated active demethylation of 5mC through the 5hmC intermediate occurs in a temporal manner and that the TET proteins play redundant and compensatory roles in MN differentiation and maturation.

INTRODUCTION

Motor neurons (MNs) connect the central nervous system and skeletal muscle to direct muscle contractions and movement. They are primarily located in the ventral horn of the spinal cord where axons extend to muscle cells at neuromuscular junctions. MN diseases, like amyotrophic lateral sclerosis (ALS), result in axon degeneration and eventual paralysis, organ failure, and death. In normal development, MNs are completely formed by embryonic day 13 (E13) in the mouse and gestation week 5.5 in humans (Altman J. 2001). *In vivo* studies have shown that MN maturation is characterized by changes in morphology and electrophysiology. As MNs mature, they increase in cell body size, extend peripheral axons, and form dendrites (Cullheim et al. 1987; Nunez-Abades and Cameron 1995; Burke and Glenn 1996; Altman J. 2001; Carrascal et al. 2005; Li et al. 2005), and are able to fire multiple action potentials to propagate synaptic signaling (Ziskind-Conhaim 1988; Martin-Caraballo and Greer 1999; Carrascal et al. 2005). Previously, McCreedy et al. developed a transgenic mouse embryonic stem cell line that allows for the selection of motor neurons in culture. This *in vitro* differentiation method, and others, have demonstrated that normal cell morphology and electrophysiology can be recapitulated in cultured MNs (Wichterle et al. 2002; Miles et al. 2004; Wichterle and Peljto 2008; Takazawa et al. 2012).

The covalent modification of cytosine in the genome is associated with gene silencing and is vital for normal development (Reik 2007) and for maintaining cellular identity (Dhawan et al. 2011b). The methylation of cytosine (5mC) has been extensively studied in neurons of the brain (Iwamoto et al. 2011; Lister et al. 2013; Kozlenkov et al. 2014) and has been shown to influence synaptic formation (Levenson et al. 2006), learning and memory (Day and Sweatt 2010; Miller et al. 2010; Zovkic et al. 2013), emotional behavior (LaPlant et al. 2010), adult neurogenesis (Ma et al. 2009), and age-related cognitive decline (Oliveira et al. 2012). Although originally thought to be a permanent covalent modification in neurons, 5mC levels change in response to external stimuli (Guo et al. 2011a). Recent studies have identified CpGs that are modified in postnatal neural tissues upon neuronal activity induced by voluntary exercise (Guo et al. 2011a), social stress conditions (McGowan et al. 2009; Elliott et al. 2010), and electroconvulsive stimulation (Guo et al. 2011a). The necessity for functional DNA methyltransferase and CpG-binding proteins in ensuring long-term neural plasticity and cognition further highlights the importance of 5mC in normal neuron function (Moretti et al. 2006; Feng et al. 2010). Although the

importance of 5mC in neurons of the brain is well established, little is know about the role 5mC plays in motor neuron maturation.

The discovery of the hydroxymethylation of cytosine (5hmC) has introduced an additional role of cytosine modification in the brain (Penn et al. 1972; Tahiliani et al. 2009). 5hmC occurs primarily at gene bodies and is associated with gene activation (Branco et al. 2012). The presence of 5hmC varies across tissue types but is highest in the hypothalamus and cerebral cortex (Globisch et al. 2010) suggesting it plays an important role in neuron function. 5hmC is an intermediate for the demethylation of 5mC (Tahiliani et al. 2009; Guo et al. 2011b) and knockout experiments have shown that the oxidation of 5mC to 5hmC by TET1 is necessary for neuronal activity-induced active DNA demethylation in the adult mouse brain (Guo et al. 2011b). Although significant progress has been made in understanding the role of 5mC and 5hmC in neurons of the developing brain, it is unknown whether 5hmC is present in motor neurons and whether the Tet demethylating machinery is necessary for normal motor neuron function.

In this study, we used Laser Capture Microdissection-Reduced Representation Bisulfite Sequencing (LCM-RRBS) to profile the genome-wide 5mC and used immunocytochemistry to profile global 5hmC of a pure population of differentiating and maturing MNs *in vitro*. We show that post-mitotic MNs undergo significant methylome remodeling upon differentiation and functional and morphological maturation. The observed methylation changes are associated with transcriptional changes of genes important for motor neuron function. Furthermore, we show that the presence of 5hmC in MNs occurs post-differentiation upon cellular maturation and that the Tet demethylation machinery is necessary for proper motor neuron differentiation and maturation. We provide the first global view of the changing DNA methylation landscape at base pair resolution in developing motor neurons.

RESULTS

In vitro selected motor neurons model morphological and electrophysiological motor neuron maturation in culture

Current *in vitro* studies of MN maturation are limited by the presence of proliferating glial cells that prevent the isolation of individual neurons. Previously, McCreedy et al. characterized a transgenic Hb9-driving puromycin (Hb9-puro) transgenic embryonic stem cell line that, upon retinoic acid and sonic

hedgehog induction, differentiates into physiologically normal MNs. Induced neural progenitor cells (NPCs) express the spinal motor neuron-specific transcription factor, *Hb9*, and therefore also express Puromycin N-acetyltransferase, which can be used to select for MNs (Figure 4.1). To better understand the utility of the Hb9-puro line as a model for studying MN maturation, we characterized the morphology and electrophysiology of maturing MNs at 24 hours (D0), two days (D2), and four days (D4) after differentiation and selection. We first assessed the formation of neurites, branch points, neuronal networks, cell area, and neurite length as a function of time. β-tubulin class III staining, imaging, and quantification showed a statistically significant increase (ANOVA, $P<0.05$) in neurite number, neurite length, neurite branch points, and cell body area from D0 to D4 (Figure 4.2 and 4.3). Furthermore, as MNs developed neurites, they formed an increasing number of contacts with neighboring cells to establish closed networks as represented by a decrease in independent networks across time (Figure 4.3). Most neurite growth and branching occurred from D0 to D2 after which neurite growth slowed (Figure 4.2 and 4.3). We conclude that cultured MNs are morphologically mature at four days after differentiation.

We next investigated the electrophysiological properties of D0, D2, and D4 MNs. Previous work by McCreedy et al. showed that D4 MNs have the ability to fire multiple action potentials during periods of prolonged depolarization and exhibit adaptation in both spike frequency and amplitude. To evaluate the functional properties of selected MNs over time, we performed whole-cell current-clamp recordings between 0 and 6 days after differentiation (puromycin selection; Figure 4.4). To test for expression of functional neurotransmitter receptors, we exposed cells to selective agonists for AMPA/kainate, NMDA, glycine, and GABA receptors and recorded the resulting inward current. Figure 4.4 (top panel) shows that the amplitude of agonist-evoked total whole-cell current density increases with time (Pearson, $P<10^{-4}$) after puromycin selection, suggesting the number of channels expressed on the surface of each cell increases over time. Interestingly, currents evoked by AMPA/kainite, NMDA, and glycine agonists showed the greatest increase from D2 to D4 while the largest increase in current density evoked by the GABA agonists was observed between D0 and D2, suggesting GABA receptor expression occurs earlier than AMPA/kainite, NMDA, and glycine receptors. Furthermore, whole-cell capacitance, which is proportional to surface area, significantly increased from 22.6 ± 1.5 pF (n=11) at D0 to 61.2 ± 4.0 pF (n=16) at D6

(ANOVA, $P<10^{-4}$) while input resistance decreased significantly between D0 and D2 but not between D2 and D4 or D6 (ANOVA, $P<10^{-4}$), which confirms our morphology experiments that a significant increase in cell size occurs between D0 and D4 (Figure 4.4, bottom panel). Our results and those from McCreedy et al. together demonstrate that selected MNs mature into functional neurons by four days after puromycin selection.

Maturing motor neurons show active gains and losses in DNA methylation at genes implicated in neurogenesis

Little is known about the role DNA methylation plays in MN maturation. To determine whether DNA methylation changes occur in maturing MNs, we applied LCM-RRBS (Schillebeeckx et al. 2013) to three biological replicates of D0, D2, and D4 MNs. Using LCM-RRBS, we were able to interrogate the DNA methylation status of 1,074,439 CpGs at all three time points. Although >99% of CpGs did not show a significant change in DNA methylation from D0 to D4, 2,894 CpGs and 1,147 CpGs showed a statistically significant loss and gain of at least 25 percent methylation, respectively (Figure 4.5, left panel; Table 4.S1). Although LCM-RRBS enriches for CpGs in promoter regions (Figure 4.S1; top left panel), we observed that most CpG changes occurred >10 kilobases from the nearest gene (Figure 4.S1; bottom panel). We, therefore, assessed whether there was an enrichment at introns, exons, promoters, or intergenic regions. We see a significant depletion of differentially methylated CpGs in promoters (p-value$<10^{-16}$; OR<0.32) and a significant enrichment for differentially methylated CpGs at introns (p-value$<10^{-9}$; OR>1.49) and intergenic regions (p-value$<10^{-9}$; OR>1.46) (Figure 4.S1). Furthermore, hypomethylated CpGs were highly enriched for CpGs distal to the nearest gene (Welch's t test; p-value$<10^{-5}$) while hypermethylated CpGs had a slight enrichment for distal CpGs (Welch's t test; p-value<0.039) with a mean distance of 17,709 bp and 14,264 bp, respectively, to the nearest transcription start site (TSS)(Figure 4.S1). Taken together, we conclude that significant changes in CpG methylation occur at putative regulatory regions distal to TSSs.

To determine whether the differentially methylated CpGs regulate genes important for motor neuron function, we identified the nearest gene to each differentially methylated CpG and performed Gene Ontology enrichment analysis. Hypomethylated CpGs were associated with genes enriched for

neuronal function, including neuron projection guidance, axon guidance, and neuron differentiation while hypermethylated CpGs were associated with genes enriched for cell differentiation, including embryonic morphogenesis, regulation of cell development, and multicellular organismal development (q-value<10^{-6}; FDR-adjusted; Table 4.1). Interestingly, changes in DNA methylation occurred at various stages in MN maturation. Roughly 65% of hypomethylated CpGs lost more methylation between D0 and D2 than between D2 and D4 while roughly 65% of hypermethylated CpGs gained more methylation between D2 and D4 than between D0 and D2 (Figure 4.5, right panel; Table 4.S1), suggesting gene activation occurs earlier than gene silencing. We conclude that hypomethylated CpGs are associated with genes important for motor neuron maturation and that the loss of CpG methylation occurs within two days after differentiation.

To assess whether the observed loss of methylation is specific to MNs, we determined the methylation states of hypomethylated CpGs for astrocytes, a glial cell type. Astrocytes and MNs share a common progenitor and, therefore, have a close epigenetic origin in which they stem from one DNA methylation profile. We find that CpGs hypomethylated in MNs are enriched for CpGs that are highly methylated in astrocytes (Figure 4.6; Fisher's Exact Test, p-value<10^{-16}). Indeed, 82% of hypomethylated CpGs are highly methylated in astrocytes with a percent methylation of 75% or greater, suggesting the observed loss of DNA methylation is specific to MNs. Together, we conclude that maturing MNs show significant gains and losses of DNA methylation at CpGs associated with genes implicated in neuronal functions and cellular differentiation and that the loss of DNA methylation is specific to motor neurons.

Motor neurons show a significant increase in expression of genes associated with active demethylation

DNA methylation at gene promoters acts to silence genes by recruiting chromatin modifying proteins or by preventing transcription factor binding and is therefore negatively correlated with gene expression (Watt and Molloy 1988; Robertson 2005). We hypothesized that the observed gains and losses of DNA methylation would be associated with gene silencing and gene activation, respectively. To test this hypothesis, we performed gene expression profiling of neural progenitor cells (NPC) and selected MNs at D0, D2, and D4 (Table 4.S2). We found 142 and 50 genes that showed a statistically

significant (q-value<0.05; FDR-adjusted) increase or decrease in expression between D0 and D4 MNs, respectively (Figure 4.7; left panel). Upregulated genes were enriched (q-value<0.05; FDR-adjusted) for neuronal function, including cell signaling, synaptic transmission, and neuromuscular process while downregulated genes were enriched (q-value<0.05; FDR-adjusted) for various biological processes including apoptosis and biosynthesis (Table 4.2). Furthermore, upregulated genes were enriched (q-value<0.05; FDR-adjusted) for ion channels, transporters, and neurotransmitter receptors (Table 4.3). Unsupervised hierarchical clustering of differentially expressed genes suggests that the majority of expression changes occur with two days of MN differentiation with fewer changes occurring between D2 and D4 than between D0 and D2 (Figure 4.7; right panel). Therefore, within two days of differentiation, motor neurons undergo significant gene activation and gene silencing as they mature into functioning neurons.

To assess whether differentially methylated CpGs were associated with an expression change, we determined the expression level of genes associated with hypomethylated and hypermethylated CpGs across NPCs, D0 MNs, D2 MNs, and D4 MNs. We found that 82 and 21 genes upregulated and downregulated in D4 MNs, respectively, were associated with differentially methylated CpGs (Fisher's Exact; p-value<10^{-16}; Figure 4.8). Of the 82 upregulated genes, 64 were associated with a distal CpG that significantly lost DNA methylation between D0 and D4 MNs (Figure 4.S2) while 13 downregulated genes were associated with a distal CpG that gained DNA methylation (data not shown). Taken together, we find that as they mature, MNs undergo significant expression changes that correlate with concomitant changes in DNA methylation.

Motif scanning identifies factors associated with active gains and losses of DNA methylation in maturing motor neurons

DNA methylation can prevent the binding of transcription factors (TFs) to DNA by directly occluding binding sites or by recruiting methyl-binding proteins (Watt and Molloy 1988). We hypothesize that differentially methylated CpGs may lie within gene regulatory regions important for MN maturation and differentiation and that the methylation status of these CpGs may regulate TF binding and gene regulation. To elucidate whether differentially methylated regions are associated with TF binding, we

evaluated 100 bp regions flanking hypomethylated and hypermethylated CpGs for the enrichment of known TFs using the Hypergeometric Optimization of Motif EnRichment (HOMER) suite of tools (Heinz et al. 2010). We find 96 and 9 predicted motifs corresponding to known cognate TFs that were enriched in hypomethylated and hypermethylated regions, respectively (q-value<0.01; FDR-adjusted). To elucidate which TFs may contribute to MN function, we performed gene ontology enrichment analysis on the TFs that were found predicted by HOMER to bind hypomethylated and hypermethylated regions. We found that 14 TFs associated with hypomethylation are enriched for neuronal function (Table 4.4). Indeed, the TFs *Isl1*, *Phox2a*, *Lhx3*, *Tbx20* are known to play critical roles in spinal and cranial MN differentiation (Mazzoni et al. 2013). Of these 14 TFs, 10 show a statistically significant (p-value<0.05; FDR-adjusted) expression difference between NPCs and D0 MNs as assessed by microarray analysis (Figure 4.S3). *Isl1* and *Lhx3* had the highest increase and decrease in expression, respectively, from NPCs to D0 MNs. We also find that hypermethylated regions were enriched for *Oct4*, a pluripotency factor, and *Olig2*, a neuro-oligodendrocyte cell-type specification factor. These findings suggest *Isl1* may direct demethylation of CpGs during motor neuron maturation while *Lhx3* may play a role in maintaining CpGs in a methylated state prior to differentiation.

Maturing motor neurons gain 5-hydroxymethylation over time

We previously demonstrated that MNs lose DNA methylation at 2,894 CpGs as they mature. Because MNs no longer undergo cell division after selection (D0), the loss of DNA methylation we observe between D0 and D4 is indicative of an active demethylation pathway, independent of cell division. Several mechanisms for active demethylation have been proposed (Branco et al. 2012; Kohli and Zhang 2013) but all require the active oxidation of 5-methylcytosine (5mC) to a 5-hydroxymethylcytosine (5hmC) intermediate by the Ten-eleven translocation (Tet) family of proteins. We hypothesized that the loss of 5mC in maturing MNs occurs with a concomitant gain of 5hmC. To test this hypothesis, we performed immunocytochemistry on fixed D0, D2, and D4 MNs with an antibody directed at 5hmC. Immunocytochemistry showed very little positive staining for 5hmC at D0 and strong 5hmC signal at D2 and D4 (Figure 4.9). We conclude that MNs gain significant 5hmC as they mature to functional neurons.

TET proteins are necessary for proper motor neuron differentiation and maturation

Our microarray expression analysis shows a statistically significant increase in *Tet2* and *Tet3* expression in D0 MNs as compared to NPCs (Student's t test; p-value<0.05; Figure 4.S4). Furthermore, qRT-PCR analysis of ES cells shows an increase in *Tet3* expression as they differentiate into NPCs (Figure 4.S5). The observed changes in gene expression of the TET proteins and the observed increase in 5hmC in maturing MNs (Figure 4.9) prompted us to hypothesize that the Tet family plays a role in MN differentiation and maturation. To determine whether the Tet family members are necessary for MN cell-type specification and maturation, we created *Tet1*, *Tet2*, and *Tet3* single knockout, *Tet1*/*Tet2* double knockout (DKO), and *Tet1*/*Tet2*/*Tet3* triple knockout (TKO) Hb9-puro ES cells using the CRISPR/Cas9 genome engineering system (Figure 4.S6). Similar to wildtype cells, all knockout cell lines formed EBs and NPCs when ES cells were induced to differentiate with retinoic acid and sonic hedgehog agonist (data not shown). *Tet1*, *Tet2*, and *Tet3* single knockout NPCs developed into normal MNs and showed normal gains of 5hmC by D4 (Figure 4.10). Interestingly, single knockout MNs, unlike wildtype MNs showed the presence of 5hmC at D0 as well (Figure 4.10). Immunocytochemistry staining showed a dramatic reduction in MN differentiation in DKO cells compared to wildtype cells but showed normal gains of 5hmC in the viable MNs that formed (Figure 4.11). TKO cells showed an almost complete inability to differentiate into MNs; curiously, the few MNs that did differentiate and mature showed strong 5hmC staining (Figure 4.11).

Differentiation experiments with single knockout ES cells suggest all three TET proteins can convert 5mC to 5hmC and play redundant roles in MN maturation. The presence of viable *Tet1*/*Tet2* DKO MNs suggests *Tet3* is sufficient for MN differentiation and can compensate for the lack of TET1 and TET2 protein, which is consistent with experiments showing that *Tet1*/*Tet2* DKO ES cells form viable and grossly normal mice (Dawlaty et al. 2013). The significant reduction in differentiation potential of *Tet1*/*Tet2*/*Tet3* knockout NPCs suggests the TET proteins are necessary for normal differentiation. A few TKO MNs, however, did successfully differentiate and mature. These TKO MNs seem morphologically normal and show strong 5hmC staining. The presence of 5hmC in MNs that completely lack any Tet

machinery suggests the TET proteins are not necessary for MN maturation and that an as of unknown protein can mediate the hydroxymethylation of cytosine.

DISCUSSION

In this study, we show that MNs undergo significant epigenetic and gene expression changes as they mature. We leverage an Hb9-puro cell line to culture pure populations of spinal MNs and leverage the LCM-RRBS method to interrogate genome-wide DNA methylation of the limited genomic material. We find that MN maturation is characterized by a loss of cytosine methylation (5mC) and a gain of cytosine hydroxymethylation (5hmC) although gains of cytosine methylation are also observed. We find that the majority of methylation loss results in an increase in expression of the nearest gene. Interestingly, methylation and expression changes did not always co-occur. Indeed, we find CpGs that require four days to become fully demethylated but the CpG's nearest gene is significantly activated within one or two days after differentiation. For example, methylation loss of a CpG near the *App* gene occurs 2-4 days after differentiation yet *App* expression significantly increases within one day of differentiation (Figure 4.S7). In most cases, however, CpG methylation correlates well with expression change (Table 4.S3). For example, CpGs associated with *Ret*, on the other hand, show a loss of DNA methylation at Day 2 with a concomitant increase in *Ret* expression (Figure 4.S7). The difference in transcriptional timing and loss of DNA methylation suggests other epigenetics factors could be regulating the activation of MN-specific transcriptional profiles.

Combinatorial TF expression defines a cell's identify. *Isl1*, *Lhx3*, and *Hb9* expression characterize spinal MNs while *Isl1*, *Phox2a*, and *Tbx20* expression characterize branchiomotor and visceromotor (cranial) neurons (Mazzoni et al. 2013). Interestingly, motif scanning of hypomethylated regions revealed the enrichment of putative binding motifs for both spinal (*Isl1* and *Lhx3*) and cranial (*Phox2a* and *Tbx20*) MN TFs (Table 4.4). Mazzoni *et al.* showed that *Isl1* directly interacts with *Lhx3* and *Phox2a* in a cell-type dependent manner, which may explain the enrichment for *Phox2*a binding motifs. We predict that known Isl1-Phox2a binding sites of cranial MNs show an enrichment for hypomethylated regions relative to NPCs in cranial MNs. Furthermore, the hypomethylated regions we have identified may also be important for proper differentiation and maturation of cranial motor neurons.

We also evaluated the presence of genome-wide cytosine hydroxymethylation (5hmC) in maturing MN cultures by immunocytochemistry. We find that MNs have very little 5hmC 24 hours after differentiation (Day 0) but gain genome-wide 5hmC within two days of maturation (Figure 4.9). Whole-genome 5hmC mapping techniques will elucidate where in the genome the 5hmC increases as MNs mature and whether the gain of 5hmC correlates with gene expression as is suggested by recent studies (Jin et al. 2011; Tan et al. 2013). *Hb9* expression peaks during NPC differentiation 4 days after EBs are induced with retinoic acid/Shh agonist and marks post-mitotic cells (Li et al. 2008b). Hb9-expressing NPCs, therefore, do not undergo any cell division during puromycin selection or maturation. Surprisingly, despite seeing no 5hmC in NPCs (Tan et al. 2013) or D0 MNs (Figure 4.9), we observe active demethylation at over 23,000 CpGs between NPCs and D0 MNs (data not shown), suggesting the active loss of methylation from NPCs to D0 MNs does not occur through the Tet pathway. Because NPCs consist of heterogeneous cells expressing different neuronal markers, DNA methylation analysis of Hb9+ FACS-enriched NPCs along with genome mapping of 5hmC in Hb9+ NPCs and D0 MNs would validate whether the aforementioned hypomethylated events are due to active demethylation mediated by the Tet machinery or due to other active demethylating pathways.

Finally, we show that *Tet1*, *Tet2*, and *Tet3* play redundant, compensatory roles in MN differentiation and maturation, which is consistent with single knockout mice studies (Dawlaty et al. 2011). Expression analysis shows an increase in *Tet2* and *Tet3* expression between NPCs and D0 MNs (Figure 4.S4) as well as high expression of *Tet3* in whole mouse spinal cords (Figure 4.S5) suggesting *Tet2* and *Tet3* are important for MN function. We find that single knockout ES cells can be induced to differentiate into normal MNs that show normal 5hmC gains demonstrating that no single Tet is required for MN differentiation and maturation. Furthermore, we find that *Tet1*/*Tet2* double knockout cells also differentiate into MNs, albeit less efficiently than wildtype cells, and show normal 5hmC gains, demonstrating that *Tet3* is sufficient for MN differentiation and maturation. *Tet1*/*Tet2*/*Tet3* triple knockout cells show a striking reduction in ability to differentiate but are able to form mature MNs, demonstrating that the presence of Tet is necessary for normal differentiation of MNs. Significant validation of the lack of Tet protein as well as quantification of morphology, differentiation, and 5hmC levels is necessary to support the experiments herein. *Tet1*/*Tet3* and *Tet2*/*Tet3* double knockout studies along with DNA methylation analysis of all

knockouts will further elucidate the overlapping role and genome specificity of each of the Tet proteins in MN maturation.

Model for active DNA methylation and gene activation in motor neuron differentiation

Our findings support a model where neuron-specific transcription factors recruit the Tet machinery to actively demethylate enhancer regions and activate genes critical to normal MN function (Figure 4.12). In this model, *Isl1* binding sites at distal regulatory regions are methylated in NPCs (1). Upon NPC differentiation into MNs, *Isl1* expression increases, binds at these regions, and recruits the Tet demethylases and other factors to remove DNA methylation (2). Upon demethylation by Tet, MN-specific target genes (e.g. *Ret*) are activated (3). Co-immunoprecipitation experiments of *Isl1* and *Tet* will directly test whether they interact and form a DNA-binding complex.

In conclusion, we present the first known evidence of active demethylation in maturing spinal MNs and report the derivation of *Tet1*, *Tet2*, and *Tet2* single, *Tet1*/*Tet2* double, and *Tet1*/*Tet2*/*Tet3* triple ES knockouts that can be differentiated and selected to create pure MN cultures. We show that all knockouts are able to form morphologically mature MNs and that the double and triple knockout lines have a significantly reduced ability to differentiate into MNs. The DNA methylation (5mC and 5hmC) and gene expression data from this study will inform future experiments to help elucidate the role of cytosine methylation in MN differentiation, maturation, and viability. Furthermore, the knockout embryonic stem cell lines derived in this study will facilitate research on dissecting the role of Tet family members in the differentiation of many other cell types as well as their roles in MN maturation.

ACKNOWLEDGEMENTS

We would like to thank Shelly for her willingness to collaborate and share unpublished reagents. I thank members of the Sakiyama-Elbert lab, especially Dylan McCreedy, for insightful discussion and help troubleshooting cell culture experiments; Andrew Yoo and Matthew McCoy for a willingness to collaborate and for discussions on DNA methylation and its role in MN cellular identity; James Huettner for assistance and guidance in assessing electrophysiological maturity of the MNs; and members of the Mitra, Cohen,

and Gordon lab for helpful discussion; and Jamie Kwasnieski and Vanessa Ridaura for comments and suggestions with writing this chapter.

AUTHOR CONTRIBUTIONS

This work was done in collaboration with Dylan McCreedy, Sumithra Sankararaman, Michael Brooks, James Huettner, Matthew McCoy, Shelly Sakiyama-Elbert, and Rob Mitra. Dylan McCreedy, James Huettner, Shelly Sakiyama-Elbert, Rob Mitra, and I conceived the project and designed experiments. Dylan McCreedy developed the Hb9-puro RW4 cell line and collected motor neurons for DNA methylation interrogation. James Huettner performed electrophysiology experiments. Sumithra Sankararaman performed DNA methylation, gene enrichment, and motif finding analysis. Michael Brooks performed motor neuron maturation image processing and analysis. Matthew McCoy completed 5-hydroxymethylation staining experiments. Vanessa Ridaura performed RNA extractions and qRT-PCRs. I performed primary DNA methylation and gene expression analysis, cell culture experiments, derived the knockout cell lines, and wrote this chapter.

MATERIALS AND METHODS

Embryonic stem cell cultures

Mouse ES cells were cultured on gelatin-coated T75 flasks in complete media consisting of Dulbecco's Modified Eagle Medium (DMEM; Life Technologies, Carlsbad, CA) containing 10% newborn calf serum (Life Technologies), 10% fetal bovine serum (Life Technologies), and 30 μM of each of the following nucleosides (Sigma): adenosine, cytosine, guanosine, and uridine. For routine culture, ES cells were passaged every two days. Briefly, ES cells were washed with DMEM containing 25 mM HEPES (Life Technologies) and dissociated with 0.25% trypsin-EDTA (Life Technologies) for 5 min. Trypsin was quenched with fresh complete media and cells were transferred to a new gelatin-coated flask at a 1:5 ratio in fresh complete media containing 1000 U/mL leukemia inhibitory factor (LIF; Millipore, Bellerica, MA) and 100 μM β-mercaptoethanol (BME; Life Technologies). All cells were cultured at 37°C in the presence of 5% CO_2.

Derivation of *Tet1*, *Tet2*, *Tet3* knockout Hb9-puro embryonic stem cells

Plasmids containing the Cas9 open reading frame or a guide RNA (gRNA) targeting *Tet1*, *Tet2*, or *Tet3* were obtained from the Genome Editing Core at Washington University in St. Louis. Hb9-puro mouse ES cells were transfected with 0.5 μg of pPB_donor plasmid (which contains a neomycin and dsRed construct), 1 μg Cas9 plasmid, and 1 μg of gRNA plasmid directed at *Tet1*, *Tet2*, and/or *Tet3* depending on the intended line using lipofectamine 2000 (Life Technologies) one day after passaging to a 6-well gelatinized plate. One day after transfection, cells were trypsinized and 10% of cells transferred to a 10 cm gelatinized dish. After incubating for one day, the neomycin analog, G-418, was added at a concentration of 1-0.5 mg/ml. Two days later, media was replaced with fresh complete media containing 0.5 mg/ml and cultures left to expand for an additional 6 days. Following a media change, individual colonies were expanded in individual wells of a gelatinized, 96-well, flat-bottom plate and screening by PCR and Illumina sequencing for insertions or deletions in the target position. Putative knockout lines were expanded and validated using Sanger sequencing.

Motor neuron differentiation

Transgenic ES cells were differentiated into MNs using a 2-/4+ RA and smoothened agonist (SAG, Millipore) induction protocol as previously described (McCreedy et al. 2012). Approximately 1 x 10^6 ES cells were cultured in suspension on 150 mm petri dishes in modified DFK5 media consisting of DMEM/F12 base media (Life Technologies) containing 5% knockout serum replacement (Life Technologies), 1x insulin transferrin selenium (ITS; Life Technologies), 50 μM nonessential amino acids (Life Technologies), 100 μM β-mercaptoethanol, 5 μM thymidine, and 15 μM of the following nucleosides: adenosine, cytosine, guanosine, and uridine. During this process, ES cells aggregate into multi-cellular EBs. After the first two days (2-), the EBs were moved to a 15 mL conical and allowed to settle for 5 min. The media was aspirated and replaced with 10 mL fresh DFK5 containing 2 μM RA and 600 nM Smoothen agonist (SAG). EBs were then cultured for an additional 4 days (4+) with media replaced every 2 days.

Motor neuron maturation

After 2-/4+ RA and SAG induction, cell were dissociated in 0.25% Trypsin-EDTA for 15 minutes and quenched with complete media. Dissociated cells were counted and centrifuged at 240 x g for 5 minutes. Cells were resuspended in DFK5 media containing 5 ng/mL glial-derived neurotrophic factor (GDNF; Peprotech, Rocky Hill, NJ), 5 ng/mL brain derived neurotrophic factor (BDNF; Peprotech, Rocky Hill, NJ), 5 ng/mL neurotrophin-3 (NT-3; Peprotech, Rocky Hill, NJ), 4 μg/mL puromycin in water (Sigma) and plated at ~6 x 10^4 cells/cm^2 in individual wells of a poly-ornithine-coated (Sigma) 6-well plate for 24 hours. In parallel, for control cultures not receiving puromycin, cells were resuspended in DFK5 media with the growth factor cocktail and plated at ~6 x 10^4 cells/cm^2 in individual wells of a poly-ornithine-coated coated 6-well plate for 24 hours. Following selection, cells were fixed for immunofluorescence or media was replaced with modified DFKNB media consisting of DFK5 and Neurobasal (NB) media (Life Technologies) mixed at a 1:1 ratio and supplemented with 1x B27, 5 ng/mL GDNF, 5 ng/mL BDNF, and 5 ng/mL NT-3. Cells were cultured in DFKNB media for up to 4 additional days.

Immunohistochemistry

Cells were fixed in 4% (w/v) paraformaldehyde (Sigma) in phosphate buffered saline (PBS) for 15 minutes at room temperature. Fixed cultures were washed once with PBS and permeabilized with 0.2% triton-X in PBS for 10 minutes. Cultures were then blocked for 1 hour at room temperature with 5% normal goat serum (NGS; Sigma) in PBS. Primary antibodies were added to blocked cultures overnight at 4°C at the following dilutions in 2% NGS in PBS: class III β-tubulin (Tuj1, 1:5000, Covance) or 5-hydroxymethylation (5hmC, 1:250, Active Motif). Primary antibodies were labeled with the appropriate AlexaFluor conjugated goat antibodies (Life Technologies) at a 1:1000 dilution in PBS for 1 hours. All wells were counterstained with DAPI (1:10000; Sigma) to label cell nuclei. Following immunofluorescence, phase contrast and fluorescent images were captured.

Quantification of motor neuron neurite formation

Five biological replicates of D0, D2, and D4 wildtype MNs were fixed and stained with class III β-tubulin as described above (see "Immunohistochemistry). For each time point, 60 images with a ~1/3 overlap at 10x magnification were taken and combined into one image using the photomerge tool in Adobe Photoshop. Each sample image was then cropped to form images of equal size. Cropped images were uploaded to the WimNuerite (Wimasis) software for neurite morphology quantification. Student's t test was applied to determine statistical significance.

Electrophysiology

D0, D2, D4, and D6 MN cultures were bath perfused with Tyrode's solution (in mM): 150 NaCl, 4 KCl, 2 $MgCl_2$, 2 $CaCl_2$, 10 glucose, 10 HEPES, pH adjusted to 7.4 with NaOH. Whole-cell electrodes pulled from borosilicate glass capillaries had an open tip resistance of 2 to 5 MOhm when filled with one of the following internal solutions (in mM): (1) 140 K-glucuronate, 10 NaCl, 5 $MgCl_2$, 0.2 EGTA, 5 Na-ATP, 1 Na-GTP, 10 HEPES, pH adjusted to 7.4 with KOH; or (2) 140 Cs-glucuronate, 5 CsCl, 5 $MgCl_2$, 10 EGTA, 5 Na-ATP, 1 Na-GTP, 10 HEPES, pH adjusted to 7.4 with CsOH. Current and voltage were recorded with an Axopatch 200A amplifier, filtered at 1 kHz, digitized at 10 kHz and analyzed off-line with

Clampfit software (pClamp 9.2). Ligand-gated channel agonists were dissolved in 160 NaCl, 10 HEPES, 2 $CaCl_2$ and applied by local perfusion (Kim et al. 2009).

DNA extraction

Selected cultures contain viable MNs as well as dead cells adhering to the well which can confound DNA methylation analysis. To degrade non-MN DNA, cells were treated with cell-impermeable DNase as follows: cells were washed three times with PBS to remove dead cells, 160 µl of DNase (10mg/ml; Sigma) in 1 ml of DFK5 and 10mM MgCl2 were added to each well and incubated for 2 hours at 37°C, then stained with DAPI or Hoechst to verify complete degradation of non-MN DNA. Cells were then washed three times with PBS and scraped off with a rubber scraper and transferred to a 1.5 ml Eppendorf tube. Collected cells were spun at 16000 rpm for 10 minutes and the supernatant removed with a pipet. DNA samples were purified using NucleoSpin Tissue XS columns (Clontech) following the protocol for cultured cells and eluted in 20 µl of nuclease-free water. Genomic DNA was quantified using the Quant-it dsDNA High Sensitivity kit (Invitrogen) and the Qubit fluorometer (Invitrogen).

LCM-RRBS

LCM-RRBS was performed on three biological replicates of D0, D2, and D4 MNs as described previously (Schillebeeckx et al. 2013). Briefly, 10 ng of DNA from each of the nine samples was digested with *MspI* (NEB), end-repaired, and ligated with pre-annealed methylated paired-end Illumina adapters. Adapter-ligated fragments were purified using MinElute columns (Qiagen) and bisulfite treated using the EZ DNA Methylation Gold Kit (Zymo). Each sample was PCR amplified in triplicate using sample-specific indexed primers using Platinum Taq Polymerase (Invitrogen) and 10 melting-annealing-extension cycles. All PCR products and replicates were pooled and analyzed by electrophoresis on a 3% 1X Tris-acetate-EDTA (TAE) NuSieve agarose gel (Lonza). Fragments between 150 bp and 350 bp were gel extracted and purified using MinElute columns (Qiagen). To minimize PCR bias, the final PCR library was amplified in quadruplicate using 11 PCR cycles. The four replicates are pooled and gel extracted to remove remaining adapter dimers and primers, then purified, and sequenced on Illumina HiSeq 2000 machines.

Sequence alignment and methylation calling

All analysis was performed using the July 2007 (NCBI37/mm9) build of the mouse genome. On average 25 million single-end 42-bp raw high quality reads per sample were aligned to the reduced reference using RRBSMAP (Xi et al. 2012) filtering against reads that contain adapter sequence. Reads that showed less than 90% bisulfite conversion (~1 unconverted non-CpG cytosine per read) were filtered to remove those that resulted from incomplete bisulfite converted molecules. Aligned reads with a mapping quality of zero were also discarded. The resulting high quality uniquely mapped reads were used for methylation calling. We identified the genomic coordinates of all CpGs in the reference sequence and assessed percent DNA methylation by calculating the fraction of reads that had an unconverted cytosine at the CpG position relative to the total reads. We required that each read have either a “TG” or “CG” dinucleotide at the expected CpG coordinate to be considered for analysis. Differentially methylated CpGs were determined using a Fisher’s Exact Test comparing the ratio of methylated versus unmethylated read for each sample. CpGs were considered hypomethylated or hypermethylated if they showed a statistically significant decrease or increase of at least 25% between D0 and D4 MNs, respectively.

Microarray expression profiling

RNA was extracted from 2-/4+ RA and SAG induced embryonic stem cells (neural progenitor cells), D0 MNs, D2 MNs, and D4 MNs using the QIAGEN RNeasy kit. Briefly, cells were resuspended in lysis buffer RLT (cat no. 79216, QIAGEN) and β-mercaptoethanol, and then disrupted and lysed using a QIAshredder column (cat no. 79656, QIAGEN). RNA was extracted using a QIAGEN RNeasy Mini kit (cat no. 74104, QIAGEN) with on-column DNase treatment (cat no. 79254, QIAGEN). Approximately 2 mg of total RNA per sample was used to prepare cDNA using SuperScript® III, First-strand Synthesis system (cat no. 18080-051, Invitrogen). Array hybridization was performed by the GTAC Microarray Core facility at Washington University. Briefly, 50ng of total RNA was used to generate biotinylated cDNA, according to the standard NuGen WT-Ovation Pico RNA Amplification kit. Following fragmentation with NuGen Encore Biotin Module, 2.5 ug of cDNA were hybridized onto Mouse Gene 1.0 ST Arrays (Affymetrix) in the GeneChip Hybridization Oven 640 for 18 hr at 45°C. GeneChips were then washed and stained in the

Affymetrix Fluidics Station 450 and were scanned using the Affymetrix GeneChip 7G 3000 Scanner. The scanned raw .CEL files were analyzed to export signal intensity values, using Affymetrix Expression Console software with Affymetrix default RMA Gene analysis settings. Probe summarization (Robust Multichip Analysis, RMA), quality control analysis, and probe annotation were performed according to recommended guidelines (Expression Console Software, Affymetrix). We determined statistical significance using the Student's *t* test and corrected for multiple hypothesis testing using the Benjamini & Hochberg's method, leaving a total of 252 probes (q-value < 0.05). Statistically significant upregulated and downregulated probes correspond to 142 and 50 genes, respectively (Table 4.S2).

Motif scanning analysis

Differentially methylated regions were defined as 200 bp genomic windows flanking a hypomethylated or hypermethylated CpG. If the CpGs were in a cluster separated by less than 200 bp, then the window was defined to be 100 bp upstream of the first CpG and 100 bp downstream of the second CpG. Motif scanning was performed on hypomethylated and hypermethylated regions independently by using the Hypergeometric Optimization of Motif EnRichment (HOMER) suite of tools HOMER to identify enrichment of known TFs (Heinz et al. 2010). Only motifs with q-values < 0.01 (Benjamini-Hochberg corrected) were considered enriched. We found 92 and 9 sequence motifs enriched in hypomethylated and hypermethylated regions, respectively, which were used for gene ontology (GO) enrichment analysis using HOMER tools. GO terms associated with biological processes were Bonferonni corrected for multiple hypotheses, filtered to remove those with a corrected p-value > 0.01, and ranked according to fold enrichment leaving 149 GO terms. Biological processes associated with neuronal function are listed in Table 4.4.

Quantitative Reverse Transcription Polymerase Chain Reaction

Tet1, *Tet2*, and *Tet3* expression of Hb9-puro embryonic stem cells, embryoid bodies, 2-/2+ embryoid bodies, and neural progenitor cells was determined by qRT-PCR using pre-validated TaqMan probes (Life Technologies) (Figure 4.S5). RNA was extracted following the protocol mentioned above and 1.5 ml of cDNA was used to run Taq-Man qPCR reactions, using TaqMan® Universal Master Mix II,

without UNG (cat no. 4440043, Life Technologies), plus commercially available TaqMan primers to Tet1 (Mm01169087_m1), Tet2 (Mm00524395_m1) and Tet3 (Mm00805756_m1) genes. All data were normalized to the endogenous control 18S ribosomal RNA (Rn18S, Mm03928990_g1) (Life Technologies) and quantitative measurements determined using the ΔΔCT approach.

DNA Methylation Datasets

Embryonic stem cell, neuronal progenitor cell, and astrocyte processed DNA methylation calls were downloaded from NCBI Gene Expression Omnibus repository under accession numbers GSE30206 and GSE11034 (Meissner et al. 2008; Stadler et al. 2011).

FIGURE 4.1

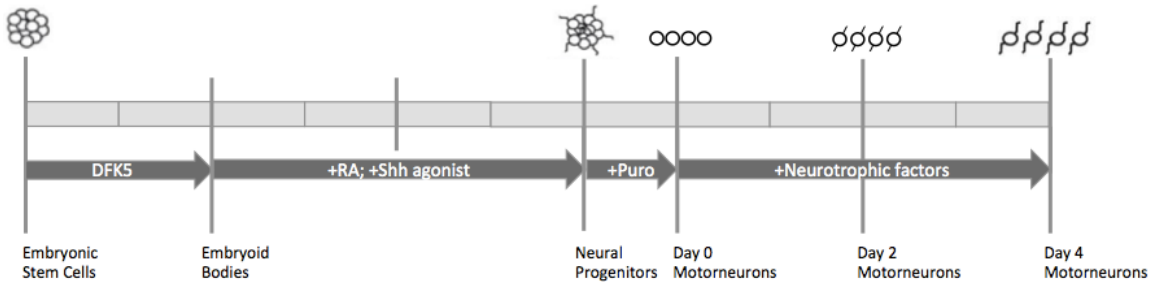

Figure 4.1 Schematic of *in vitro* differentiation protocol for creating pure motor neuron populations. Embryonic stem cells are differentiated into embryoid bodies (EBs) in DFK5 media for two days and further induced down the neural lineage with retinoic acid and sonic hedgehog agonist for four days to form cell aggregates that express neural progenitor markers Olig2 and Hb9/MNR2. Neural progenitor cells (NPCs) are dissociated and plated on poly-ornithine plates in DFK5 media containing puromycin and neurotrophic factors. Twenty-four hours after selection (Day 0), media is replaced and cells cultured for four days (Day 4). Gray boxes represents one day. See Materials and Methods for additional culturing details.

FIGURE 4.2

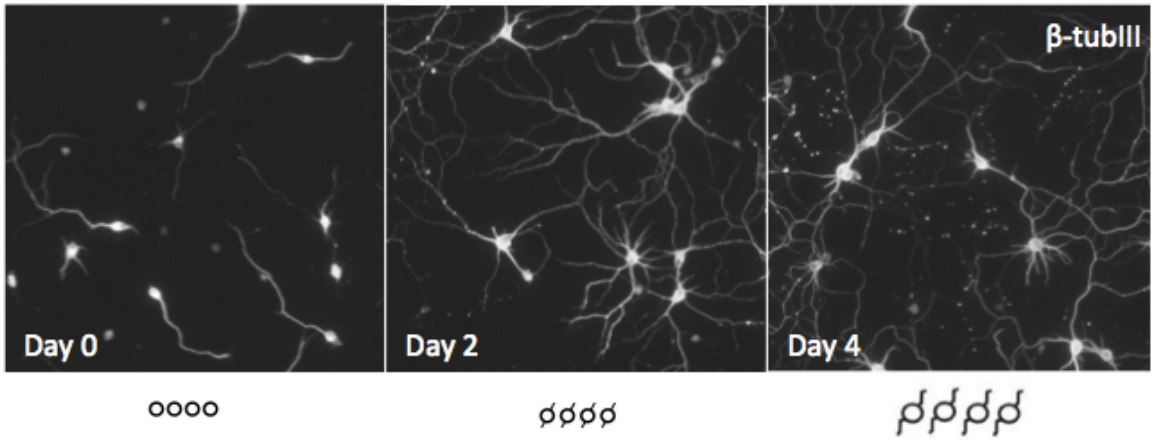

Figure 4.2 Immunocytochemistry of motor neurons after puromycin selection. Cultures were fixed with paraformaldehyde and stained with class III β-tubulin (β-tubIII, recognized by Tuj1 antibody) at 24 hours (Day 0), two days (Day 2), or four days (Day 4) after puromycin selection to visualize neurite growth.

FIGURE 4.3

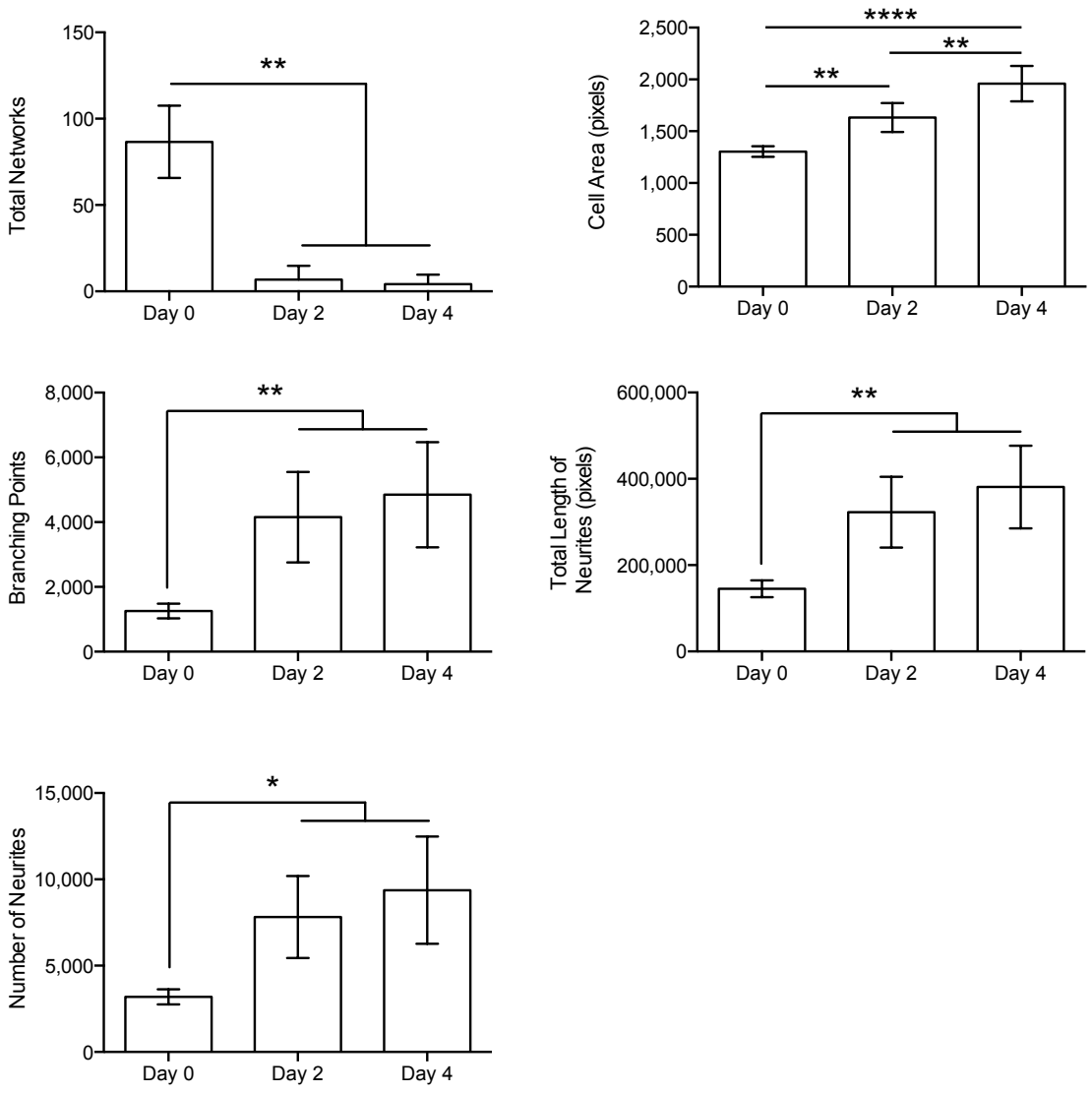

Figure 4.3 Quantification of the morphological features of motor neurons over time. Motor neurons exhibit the growth of processes from Day 0 to Day 4 as quantified by total networks (top left), cell area (top right), number of neurite branching points (middle left), total length of neurites (middle right), and number of total neurites (bottom). Each time point represents the average of five biological replicates. Error bars show the standard error of the mean. * p-value < 0.05; ** p-value <0.01; ****p-value <0.001.

FIGURE 4.4

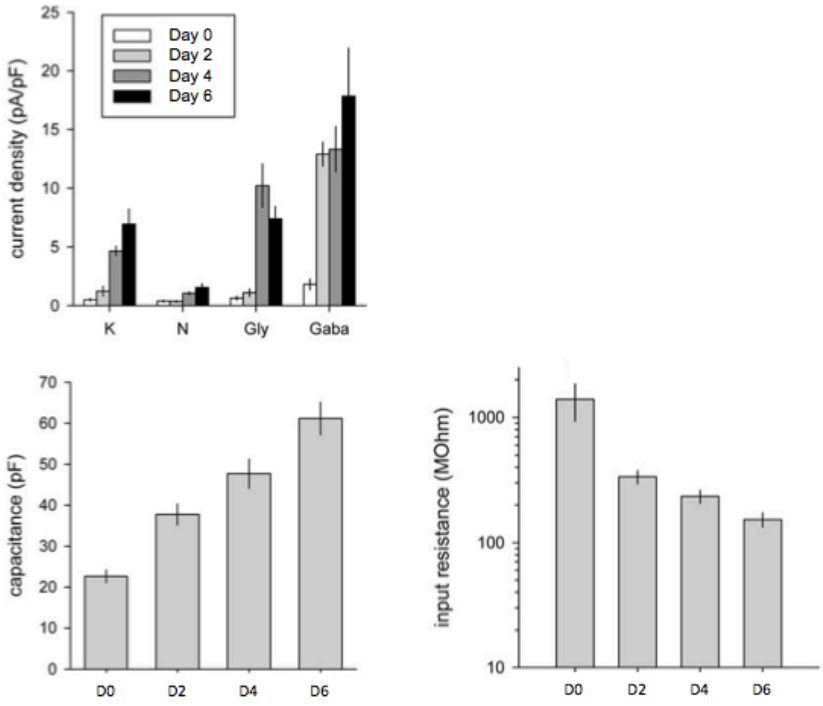

Figure 4.4 Quantification of electrophysiological features of motor neurons over time. (Top row) Whole-cell current density (pA/pF) was recorded of D0, D2, D4, and D6 motorneurons in response to agonists for specific ligand-gate ion channels; K: AMPA/kainate, N: NMDA, Gly: glycine, and Gaba: GABA receptors. (Bottom row) Left panel shows whole-cell capacitance (pF) of motor neurons at D0, D2, D4, and D6 post-selection and the right panel shows whole-cell input resistance across D0, D2, D4, and D6 time points.

FIGURE 4.5

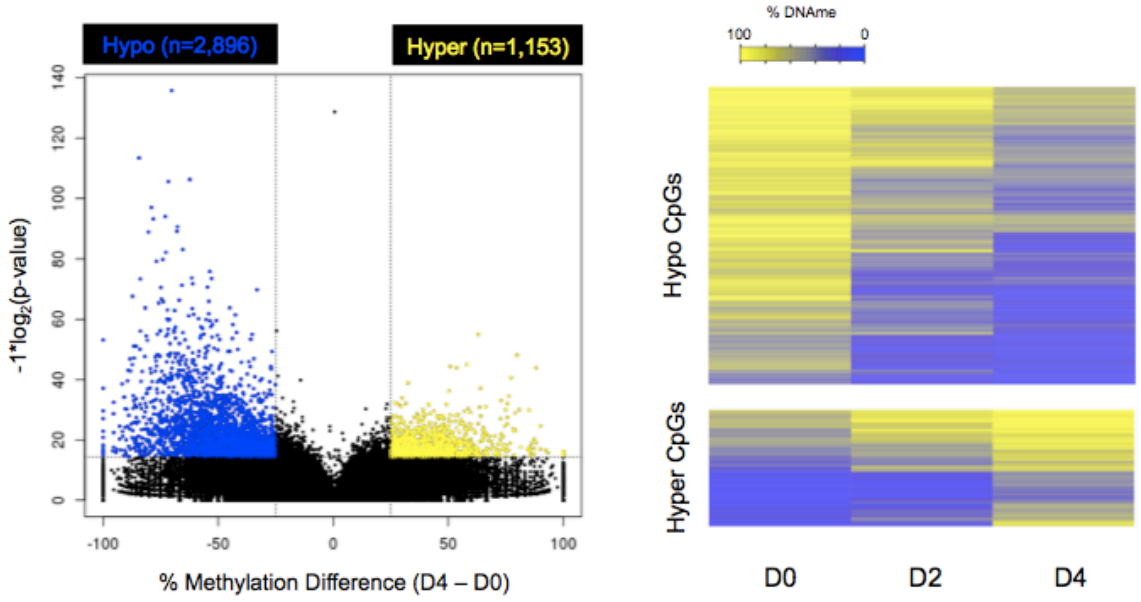

Figure 4.5 Motor neurons gain and lose DNA methylation with time. DNA methylation changes between D0 and D4 motorneurons was assessed using a Fisher's Exact Test. (Left panel) Dot plot of the inverse $\log_2$ p-value as a function of percent methylation difference between D0 and D4 motor neurons. CpGs passing Bonferonni multiple-hypothesis corrections with a percent methylation change of less than -25 or greater than 25 were considered as hypomethylated (blue) and hypermethylated (yellow), respectively. Horizontal dotted line represents the Bonferonni corrected p-value cutoff and the vertical dotted lines show a -25 and 25 percent methylation difference cutoff. (Right panel) Unsupervised hierarchical clustering heatmap of percent methylation for the 2,896 hypomethylated and the 1,153 hypermethylated CpGs in D0, D2, and D4 motor neurons. Each row represents an individual CpG. Yellow denotes high (e.g. 100) percent methylation levels and blue denotes low (e.g. 0) percent methylation levels at a particular time point.

TABLE 4.1

GO Term	Description	P-value	FDR q-value	Enrichment
GO:0097485	neuron projection guidance	5.77E-10	2.58E-07	3.05
GO:0007411	axon guidance	5.77E-10	2.49E-07	3.05
GO:0001525	angiogenesis	1.06E-08	2.73E-06	2.54
GO:0010975	regulation of neuron projection development	3.18E-08	7.88E-06	2.27
GO:0045664	regulation of neuron differentiation	1.04E-10	5.78E-08	2.2
GO:0051960	regulation of nervous system development	7.45E-13	9.64E-10	2.13
GO:0016477	cell migration	2.62E-11	1.91E-08	2.13
GO:0031344	regulation of cell projection organization	3.70E-08	8.61E-06	2.13
GO:0050767	regulation of neurogenesis	1.77E-11	1.47E-08	2.12
GO:0051270	regulation of cellular component movement	8.85E-11	5.15E-08	2.09
GO:0048598	embryonic morphogenesis	5.28E-11	2.56E-08	2.75
GO:0007389	pattern specification process	2.16E-08	4.57E-06	2.5
GO:0048731	system development	4.49E-11	2.28E-08	2.37
GO:0045597	positive regulation of cell differentiation	5.12E-09	1.32E-06	2.27
GO:0051960	regulation of nervous system development	7.15E-09	1.66E-06	2.25
GO:0007275	multicellular organismal development	2.56E-14	9.94E-11	2.24
GO:0051094	positive regulation of developmental process	1.04E-10	4.64E-08	2.17
GO:0000122	negative regulation of transcription from RNA polymerase II promoter	5.45E-08	9.91E-06	2.17
GO:0060284	regulation of cell development	5.27E-08	9.74E-06	2.09
GO:0045892	negative regulation of transcription, DNA-dependent	1.78E-09	5.30E-07	2.04

Table 4.1 Gene Ontology (GO) enrichment analysis of the genes nearest to differentially methylated CpGs. GO term p-values were adjusted for multiple hypothesis (FDR<5%) and ranked according to enrichment. The top ten GO terms of genes associated with hypomethylated (upper) and hypermethylated CpGs (lower) are shown.

FIGURE 4.6

Figure 4.6 CpGs hypomethylated in motorneurons are hypermethylated in NPCs and astrocytes. Unsupervised hierarchical clustering of the methylation status of hypomethylated CpGs in NPCs, D0 motor neurons, D2 motor neurons, D4 motor neurons, and astrocytes. The heatmap shows that 87% and 82% of hypomethylated CpGs are highly methylated (>=75%) in NPCs and astrocytes, respectively. The colored bar in the lower portion represents the expected result for motor neuron-specific methylation loss. Each row of the heatmap represents an individual CpG. Yellow denotes high (e.g. 100) percent methylation levels and blue denotes low (e.g. 0) percent methylation levels.

FIGURE 4.7

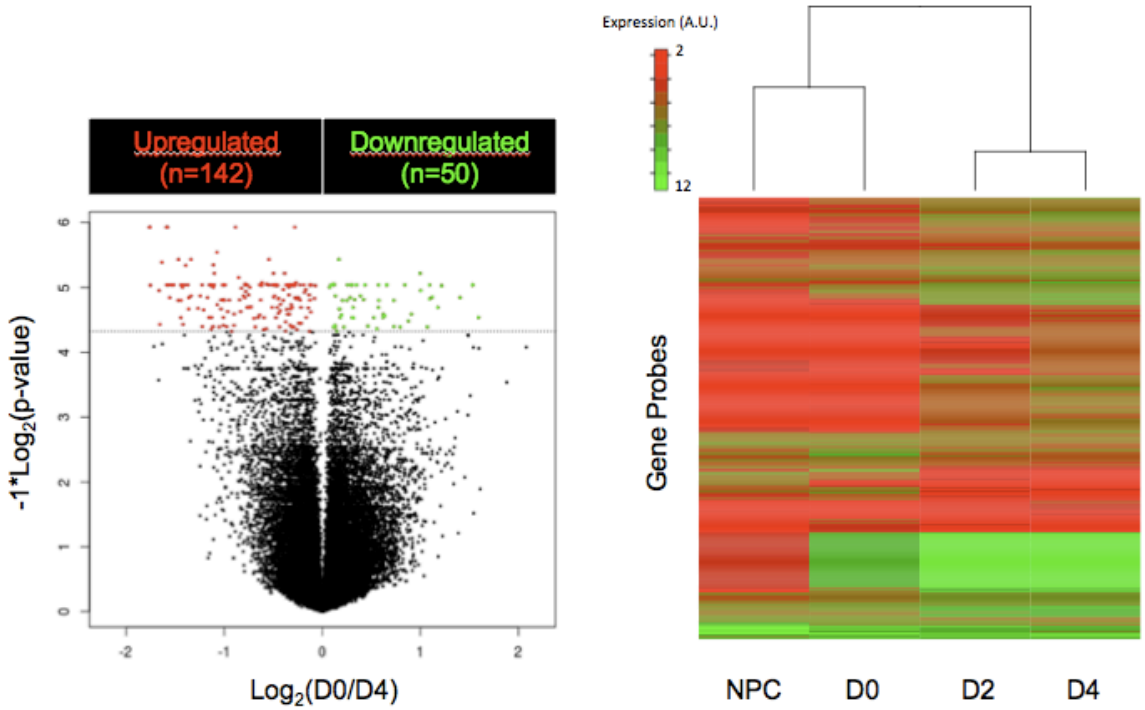

FIGURE 4.7 Gene expression profiling of neural progenitors and motor neurons. Microarray hybridization was performed on RNA isolated from neural progenitor cells (NPC), day 0 motor neurons (D0), day 2 motor neurons (D2), and day 4 motor neurons (D4). Genes differentially expressed between D0 and D4 motor neurons were identified using Student's t test and corrected for multiple hypothesis using Benjamini & Hochberg's (FDR) method. (Left panel) Dot plot of the inverse $\log_2$ p-value as a function of the $\log_2$ ratio of D0 and D4 expression. Red dots represent statistically significant upregulated probes and green dots represent statistically significant downregulated probes. Horizontal dotted line represents the multiple hypothesis corrected significant cutoff (q-value<0.05). (Right panel) Unsupervised hierarchical clustering heatmap of all probes differentially expressed between D0 and D4 motor neurons. Red and green indicate low and high expression levels, respectively. (A.U.) Arbitrary Units. Each sample represents the average of three biological replicates.

TABLE 4.2

GO Term	Description	P-value	FDR q-value	Enrichment
GO:0035235	ionotropic glutamate receptor signaling pathway	5.07E-06	4.91E-03	33.82
GO:0035249	synaptic transmission, glutamatergic	1.71E-05	1.04E-02	25.37
GO:0007215	glutamate receptor signaling pathway	5.55E-05	2.80E-02	19.02
GO:0051966	regulation of synaptic transmission, glutamatergic	7.09E-05	3.05E-02	17.90
GO:0007270	neuron-neuron synaptic transmission	1.00E-05	7.28E-03	17.29
GO:0050905	neuromuscular process	1.35E-05	8.69E-03	11.56
GO:0007268	synaptic transmission	8.73E-10	1.01E-05	11.14
GO:0007267	cell-cell signaling	1.65E-08	6.41E-05	7.64
GO:0044700	single organism signaling	7.35E-08	2.14E-04	6.73
GO:0023052	signaling	7.35E-08	1.71E-04	6.73
GO:0034645	negative regulation of muscle cell apoptotic process	2.24E-05	4.34E-02	54.17
GO:0009059	cellular macromolecule biosynthetic process	3.02E-06	3.51E-02	3.43
GO:0044249	macromolecule biosynthetic process	3.93E-06	2.28E-02	3.36
GO:1901576	cellular biosynthetic process	7.35E-06	2.85E-02	2.91
GO:0009058	organic substance biosynthetic process	1.09E-05	3.17E-02	2.83
GO:0010656	biosynthetic process	1.51E-05	3.52E-02	2.77

Table 4.2 Gene Ontology enrichment of biological processes for genes differentially expressed between D0 and D4 motorneurons. GO term p-values were adjusted for multiple hypotheses (FDR<5%) and ranked according to enrichment. All GO categories that pass multiple hypothesis corrections are shown for upregulated (top) and downregulated (bottom) genes.

TABLE 4.3

GO Term	Description	P-value	FDR q-value	Enrichment
GO:0005219	ryanodine-sensitive calcium-release channel activity	4.28E-05	7.62E-03	152.19
GO:0048763	calcium-induced calcium release activity	1.28E-04	2.09E-02	101.46
GO:0004971	alpha-amino-3-hydroxy-5-methyl-4-isoxazole propionate glutamate receptor activity	2.55E-04	3.69E-02	76.1
GO:0005234	extracellular-glutamate-gated ion channel activity	1.69E-06	4.41E-04	43.48
GO:0004970	ionotropic glutamate receptor activity	2.30E-06	5.61E-04	40.58
GO:0008066	glutamate receptor activity	9.77E-06	2.25E-03	28.99
GO:0004890	GABA-A receptor activity	2.10E-04	3.29E-02	25.37
GO:0016917	GABA receptor activity	3.38E-04	4.72E-02	21.74
GO:0005230	extracellular ligand-gated ion channel activity	2.05E-07	6.16E-05	16.65
GO:0005231	excitatory extracellular ligand-gated ion channel activity	2.15E-04	3.24E-02	13.53

Table 4.3 Gene Ontology enrichment analysis of functional categories for genes upregulated between D0 and D4 motorneurons. GO term p-values were adjusted for multiple hypotheses (FDR<5%) and ranked according to enrichment. The top ten GO terms of upregulated genes are shown.

FIGURE 4.8

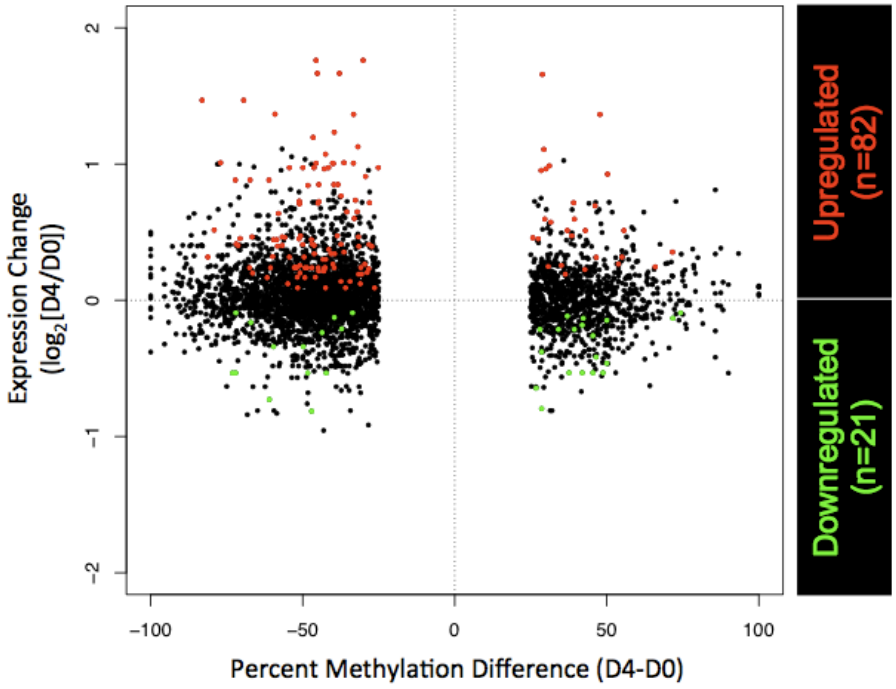

Figure 4.8 Differentially methylated CpGs are associated with differentially expressed genes. For all differentially methylated CpGs, the expression level of the nearest gene was compared between D0 and D4 motorneurons. The plot shows the expression change relative to the difference in DNA methylation between D0 and D4 motorneurons. Red dots represent upregulated genes and green dots represent downregulated genes (p-value<0.05, Student's t test).

TABLE 4.4

GO Term	Description	Genes	P-value	FDR q-value	Enrichment	Rank
GO:48665	neuron fate specification	Isl1, Lhx3, Atoh1, Foxa2	2.08E-07	2.52E-03	76.37	6
GO:21522	spinal cord motor neuron differentiation	Isl1, Phox2a, Lhx3, Tbx20	4.61E-07	5.58E-03	63.01	7
GO:21517	ventral spinal cord development	Isl1, Phox2a, Lhx3, Tbx20	7.48E-07	9.04E-03	56.01	10
GO:48663	neuron fate commitment	Isl1, Smad4, Lhx3, Atoh1, Pax7, Foxa2	1.01E-09	1.22E-05	55.59	11
GO:21515	cell differentiation in spinal cord	Isl1, Phox2a, Lhx3, Pax7, Tbx20	2.90E-08	3.50E-04	55.27	12
GO:21510	spinal cord development	Isl1, Phox2a, Lhx3, Pbx3, Pax7, Tbx20	2.53E-09	3.06E-05	47.85	13
GO:30182	neuron differentiation	Isl1, Phox2a, Smad4, Pbx3, Stat3, Pax7, Mef2c, Tbx20, Foxa2, Lhx2, Lhx3, Atoh1	5.96E-10	7.21E-06	10.33	62
GO:48699	generation of neurons	Isl1, Phox2a, Smad4, Pbx3, Atf1, Stat3, Pax7, Mef2c, Tbx20, Foxa2, Lhx2, Tcf3, Lhx3, Atoh1	3.81E-10	4.60E-06	8.04	88
GO:22008	neurogenesis	Isl1, Phox2a, Smad4, Pbx3, Atf1, Stat3, Pax7, Mef2c, Tbx20, Foxa2, Lhx2, Tcf3, Lhx3, Atoh1	8.26E-10	9.99E-06	7.58	95

Table 4.4 Gene Ontology (GO) enrichment analysis on transcription factors enriched for a sequence motif in hypomethylated regions. We identified 149 statistically enriched GO terms (q-value<0.01; FDR-adjusted) associated with hypomethylated regions and ranked them according to enrichment fold. Table 4.4 presents those GO terms and TFs associated with neuronal function. No GO terms were found to be enriched for hypermethylated regions.

FIGURE 4.9

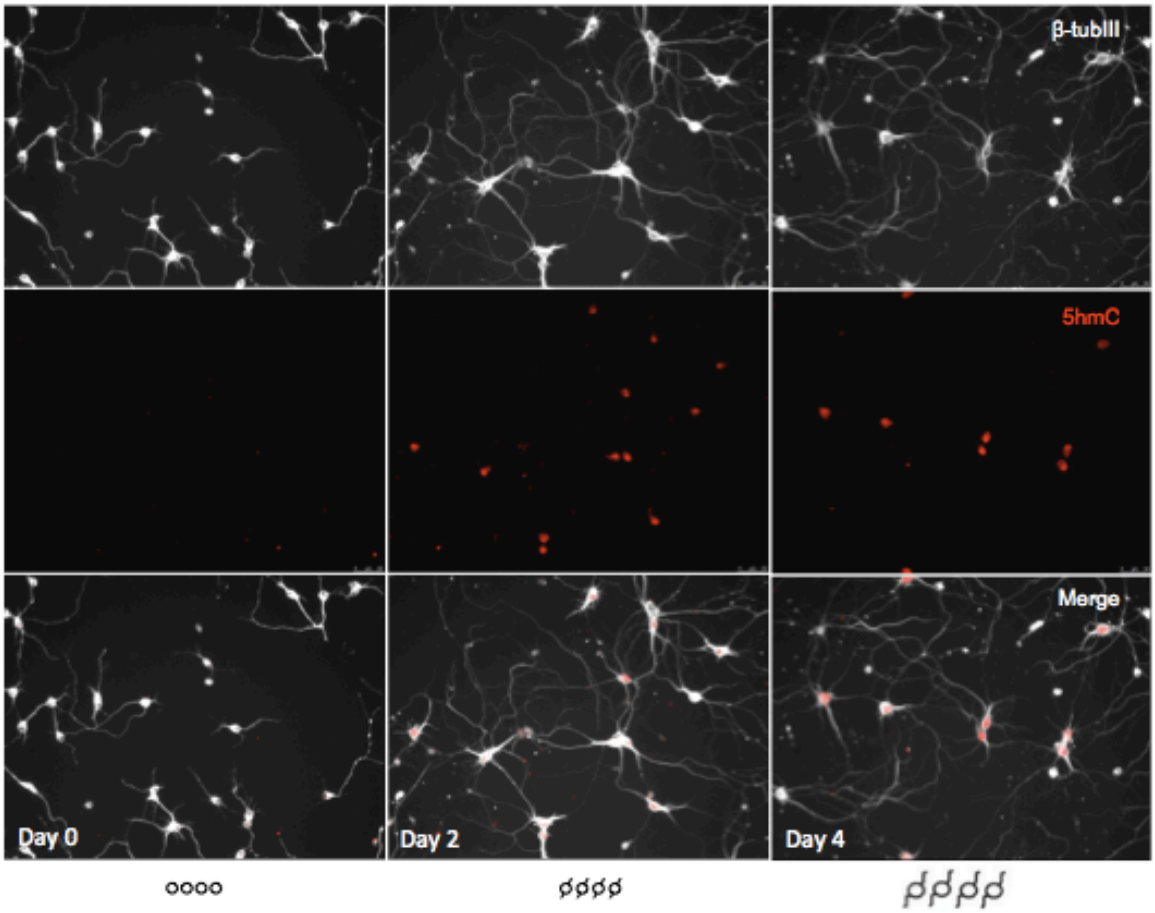

Figure 4.9 Motor neurons gain 5-hydroxymethylation with time. Puromycin-selected cultures were fixed and stained with β-tubulin class III (β-tubIII) and an antibody targeting 5-hydroxymethylation (5hmC) at 24 hours (Day 0), two days (Day 2), or four days (Day 4) after selection to visualize neurite growth and the presence of 5hmC.

FIGURE 4.10

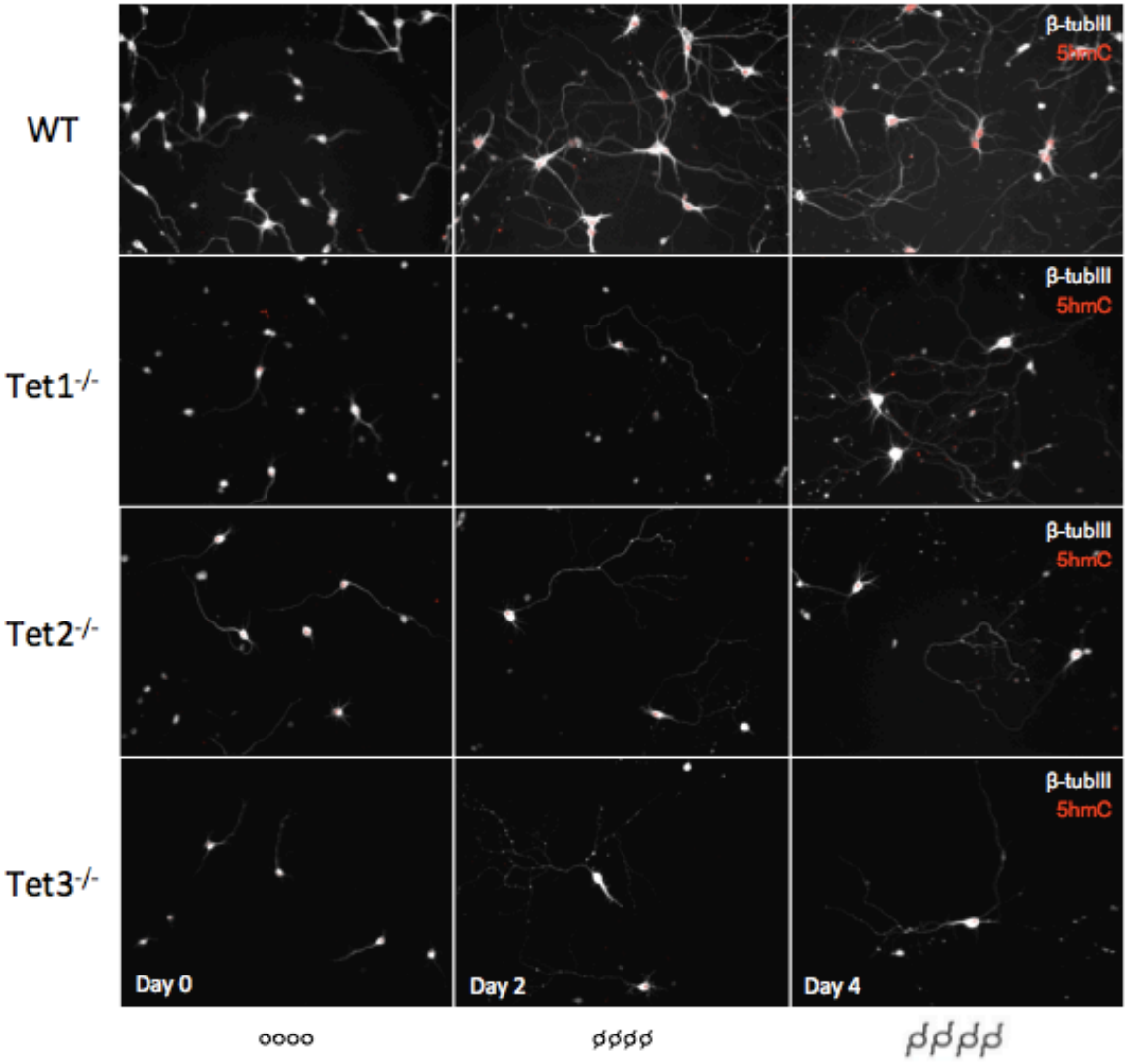

Figure 4.10 Tet single knockout motor neurons show 5-hydroxymethylation. Puromycin-selected cultures were fixed and stained with β-tubulin class III (β-tubIII) and an antibody targeting 5-hydroxymethylation (5hmC) at 24 hours (Day 0), two days (Day 2), or four days (Day 4) after selection to visualize neurite growth and the presence of 5hmC. Single knockout NPCs differentiate into normal motor neurons that look morphologically normal through Day 4. Single knockout lines also show normal gains in 5hmC by Day 4.

FIGURE 4.11

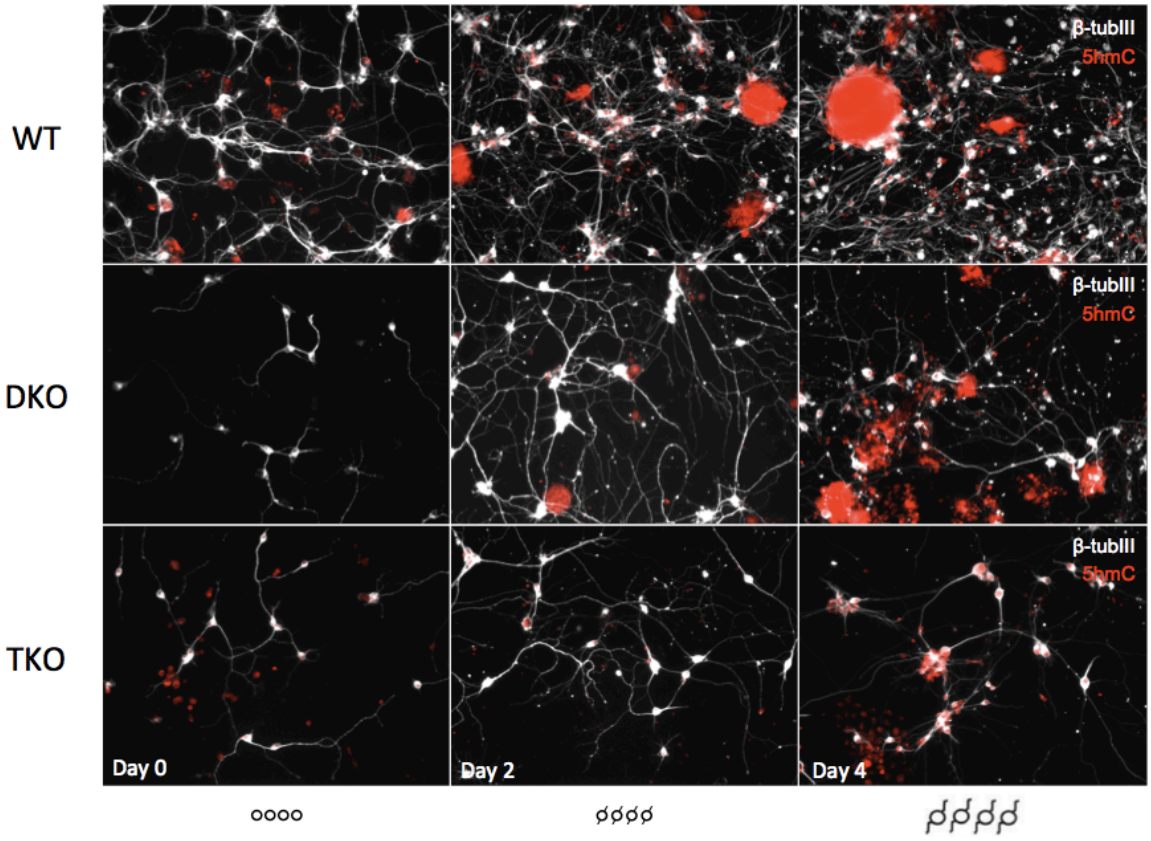

Figure 4.11 Tet double and triple knockout motor neurons show 5-hydroxymethylation. Unselected cultures were fixed and stained with β-tubulin class III (β-tubIII) and an antibody targeting 5-hydroxymethylation (5hmC) at 24 hours (Day 0), two days (Day 2), or four days (Day 4) after puromycin selection to visualize neurite growth and the presence of 5hmC. Double knockout (DKO) and triple knockout (TKO) show a significantly reduced ability to differentiate into motor neurons (MNs) as demonstrated by the lower number of normal MNs at Day 0, Day 2, and Day 4. DKO and TKO MNs also show 5hmC staining by Day 4 as demonstrated by the co-staining of β-tubulin class III (white) and 5hmC (red). Cells not positive for β-tubulin class III, presumably glial precursors, showed significant 5hmC staining in wildtype, DKO, and TKO cultures as shown by the large 5hmC positive areas (red). Note that 5hmC staining for DKO Day 0 failed.

FIGURE 4.12

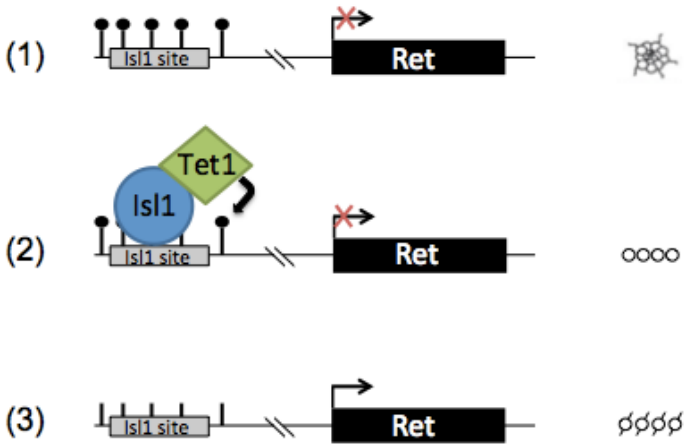

Figure 4.12 Model for active demethylation and gene activation. Our study finds motor neuron-specific transcription factor (TF) binding sites are actively demethylated during motor neuron differentiation and maturation. These findings support a model in which regions that regulate motor neuron-specific genes (e.g. Ret) contain TFs binding sites (e.g. Isl1) that are methylated in neuron progenitor cells (1). During differentiation, Isl1 binds these regulatory regions and recruits the Tet demethylation machinery, which removes the DNA methylation (2). Upon demethylation, the target gene is activated (3).

SUPPLEMENTARY FIGURES AND TABLES

FIGURE 4.S1

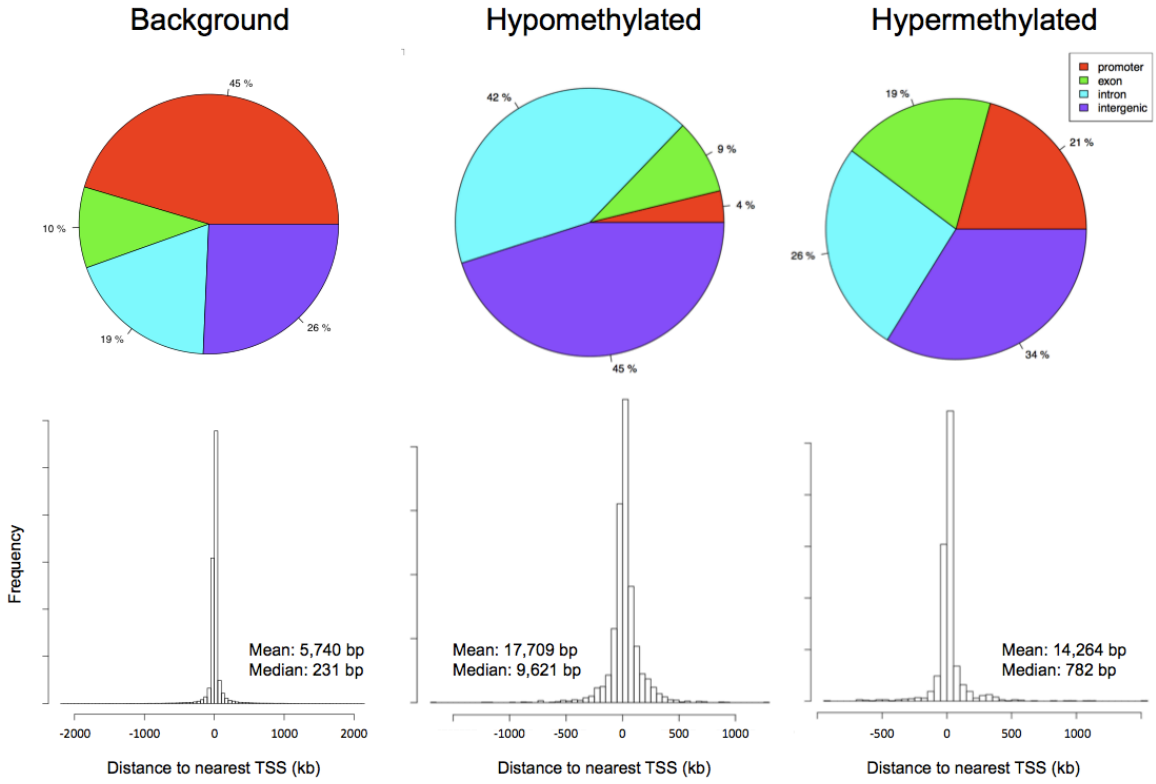

Figure 4.S1 Genomic distribution of differentially methylated CpGs. (Top panel) Background (left), hypomethylated (middle), and hypermethylated (right) CpGs were classified as located in promoters, exons, introns, or intergenic regions. (Bottom panel) Histogram of the distance to the nearest UCSC defined transcription start site for background (left), hypomethylated (middle), and hypermethylated (right) CpGs. Background, hypomethylated, and hypermethylated CpGs are located at a mean distance of 5,740 bp, 17,709 bp, and 14,262 bp from the nearest transcription start site (TSS), respectively.

FIGURE 4.S2

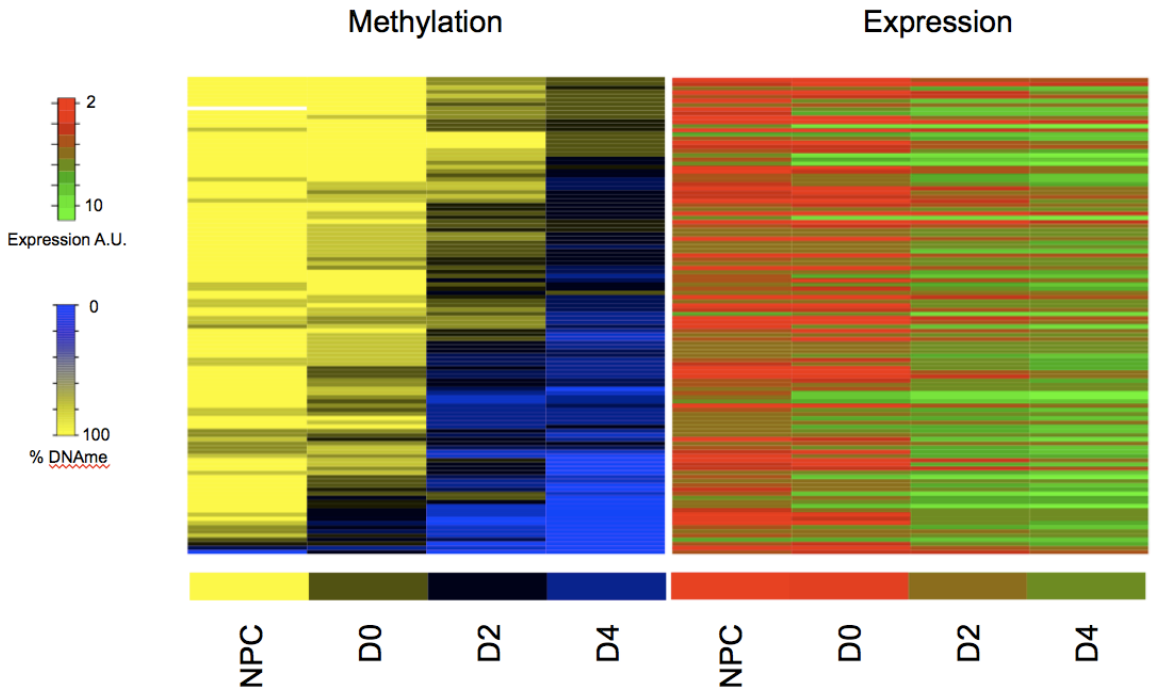

Figure 4.S2 Hypomethylated CpGs and attendant gene expression. Unsupervised hierarchical clustering was performed on the methylation status of hypomethylated CpGs across NPCs, D0, D2, and D4 motorneurons (left). The absolute expression level of the gene nearest to the hypomethylated CpG is plotted across NPCs, D0, D2, and D4 motorneurons (right). Yellow denotes high (e.g. 100) percent methylation levels and blue denotes low (e.g. 0) percent methylation levels. Red and green indicate low and high expression levels, respectively. (A.U.) Arbitrary Units. Each sample represents the average of three biological replicates.

FIGURE 4.S3

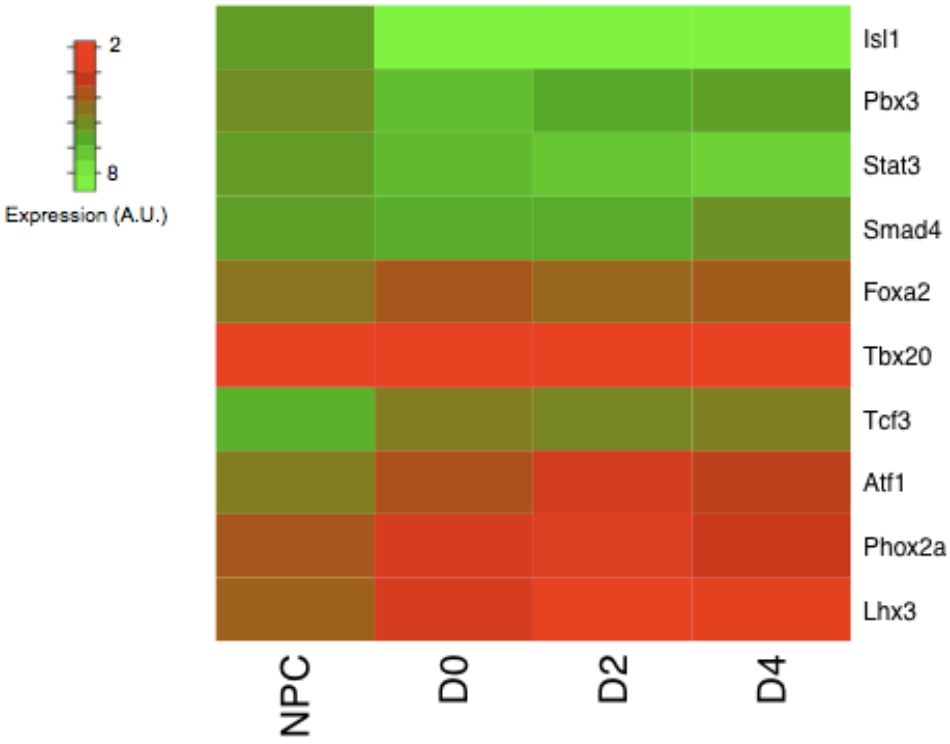

Figure 4.S3 Expression of transcription factors with enriched sequence motifs at hypomethylated regions. Heatmap of the absolute expression level of each transcription factor across NPC, D0, D2, and D4 timepoints. Rows are sorted on the ratio of NPC to D0 expression. *Isl1*, *Pbx3*, *Stat3*, and *Smad4* show a statistically significant increase in expression from NPC to D0 while *Foxa2*, *Tbx20*, *Tcf3*, *Atf1*, *Phox2a*, and *Lhx3* show a statistically significant decrease in expression. A.U.: arbitrary units.

FIGURE 4.S4

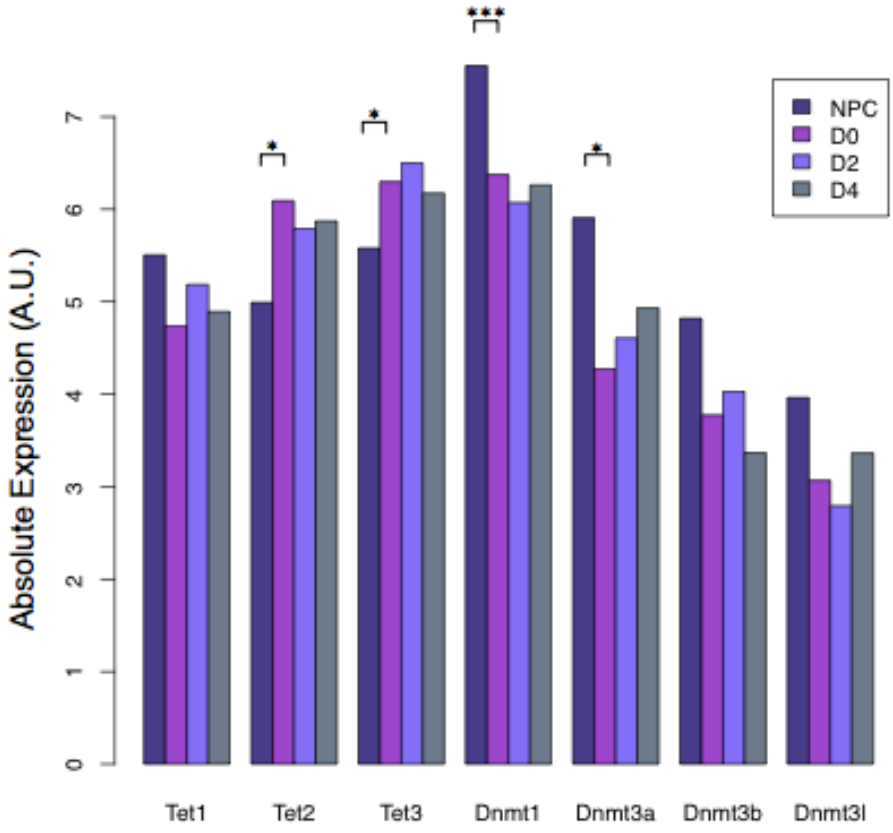

Figure 4.S4 Absolute expression of *Tet* and *Dnmt* families across NPC, D0, D2, and D4. Expression levels were determined by Affymetrix Mouse Gene 1.0 ST microarray. (A.U.) Arbitrary Units. * p-value<0.05; *** p-value<0.005 (Student's t test).

FIGURE 4.S5

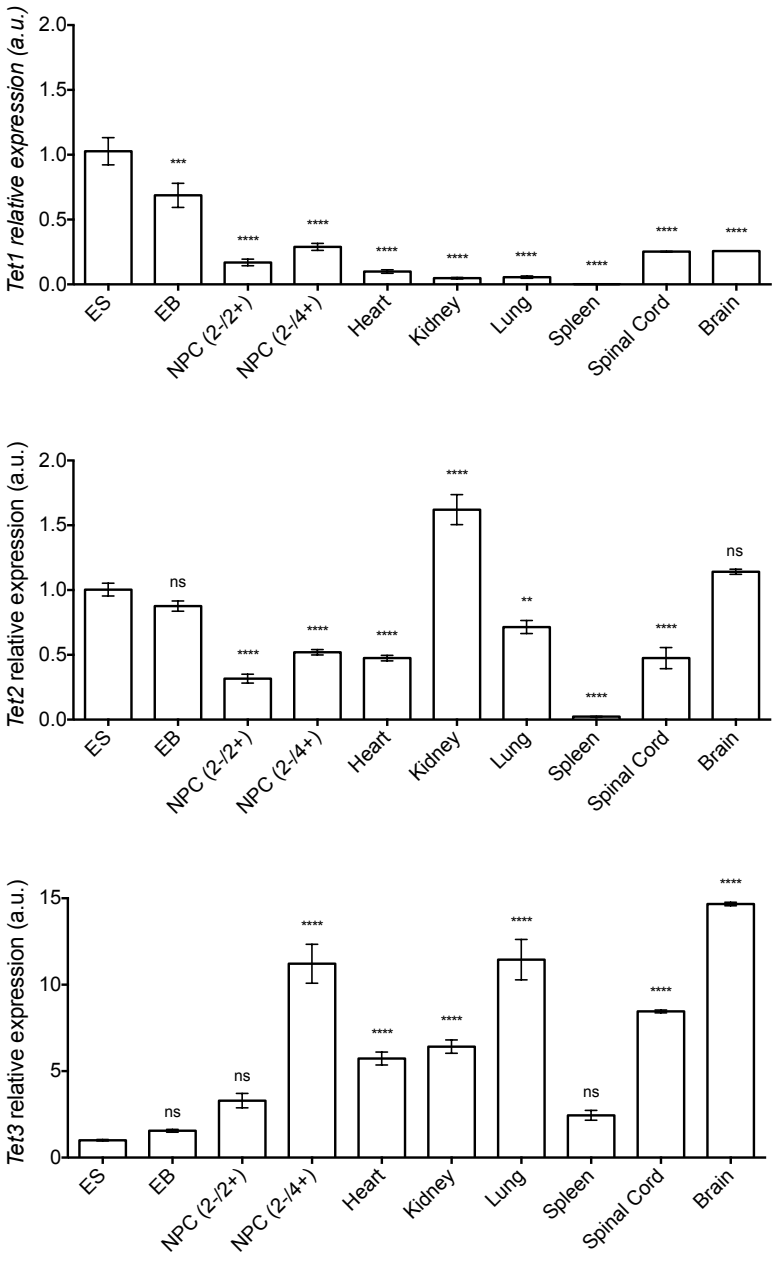

Figure 4.S5 Wildtype expression of *Tet1*, *Tet2*, and *Tet3* across various tissues. qRT-PCR analysis of Tet expression in wildtype embryonic stem cells (ES), embryoid bodies (EB), EBs induced with retinoic acid/shh agonist for two days (NPC (2-/2+)), EBs induced with retinoic acid/ssh agonist for four days (NPC (2-/4+)), and various tissues. Values are normalized to ES expression. (a.u.); arbitrary units. ns, not significant; ** p-value <0.01; *** p-value < 0.05; ****p-value <0.001.

FIGURE 4.S6

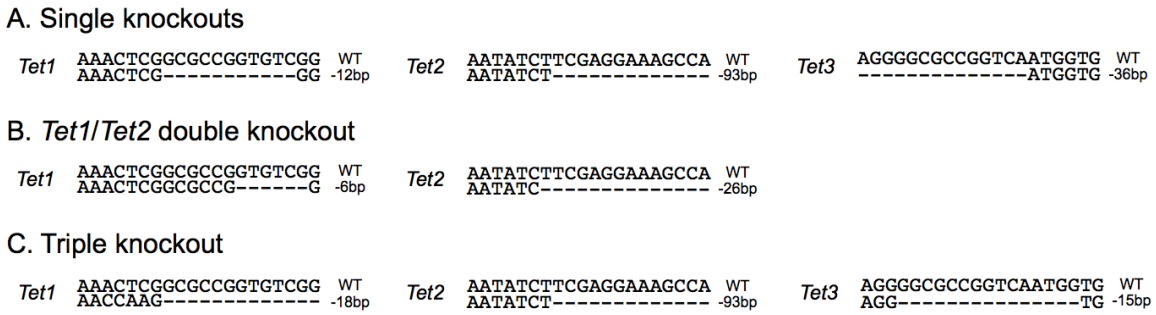

Figure 4.S6 Sanger sequence validation of knockout alleles. A portion of the wildtype and knockout *Tet1*, *Tet2*, or *Tet3* allele is represented for either (A) *Tet1*, *Tet2*, or *Tet3* single knockouts, (B) *Tet1*/*Tet2* double knockout, or (C) *Tet1*/*Tet2*/*Tet3* triple knockout.

FIGURE 4.S7

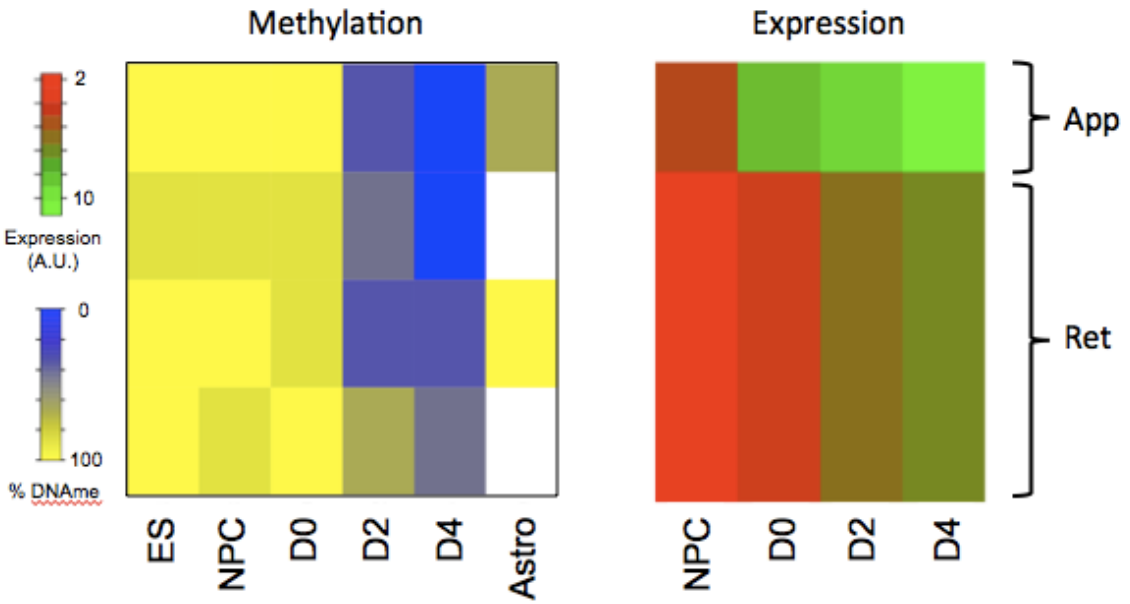

Figure 4.S7 DNA methylation and expression of putative Isl1-regulated genes *App* and *Ret*. (Left panel) DNA methylation heatmap of hypomethylated CpGs associated with *App* (1 CpG) and *Ret* (3 CpGs) were found to be enriched for the *Isl1* sequence motif. All CpGs show significant loss of DNA methylation by D4 after differentiation and are specific to motor neurons. (Right Panel) Heatmap of *App* and *Ret* absolute expression across time shows *App* is activated upon differentiation (D0) while Ret is gradually activated through D2 and D4 after differentiation. Yellow denotes high (e.g. 100) percent methylation levels and blue denotes low (e.g. 0) percent methylation. White denotes missing data. Red and green indicate low and high expression levels, respectively. (A.U.) Arbitrary Units. Each sample represents the average of three biological replicates.

TABLE 4.S1

https://cgs.wustl.edu/~maxim/Thesis/Chapter4/Table4.S1.xlsx

Table 4.S1 List of CpGs that show a statistically significant gain or loss of DNA methylation between D0 and D4 motor neurons. The table contains the DNA methylation status and read coverage for each CpG across embryonic stem cells, neural progenitor cells, D0 motor neurons, D2 motor neurons, D4 motor neurons, and astrocytes as well as categorizes each CpG as changing early (from D0 to D2) or late (from D2 to D4). The URL address above contains a link to Table 4.S1.

TABLE 4.S2

https://cgs.wustl.edu/~maxim/Thesis/Chapter4/Table4.S2.xlsx

Table 4.S2 List of differentially expressed probes. The table contains a tab with normalized expression levels of all probes ("Expression_Master"), a tab listing the probes that are differentially expressed between NPCs and D0 motor neurons ("NPCvD0"), a tab listing the probes that are differentially expressed between D0 and D4 motor neurons ("D0vD4"), a tab listing probes downregulated from D0 to D4 motor neurons, and a tab listing probes upregulated from D0 to D4 motor neurons. The URL address above contains a link to Table 4.S2.

TABLE 4.S3

https://cgs.wustl.edu/~maxim/Thesis/Chapter4/Table4.S3.xlsx

Table 4.S3 List of CpGs that show a statistically significant gain of DNA methylation and are associated with genes that increase in expression and CpGs that show a statistically significant loss of DNA methylation and are associated with genes that decrease in expression between D0 and D4 motorneurons. The table contains the DNA methylation status and expression of its associated gene of ES cells, NPCs, D0 motor neurons, D2 motor neurons, and D4 motor neurons. The URL address above contains a link to Table 4.S3.

CHAPTER 5: CONCLUDING REMARKS AND DISCUSSION

The onset of next-generation massively paralleled sequencing as a commoditized tool has significantly advanced our understanding of the role of DNA methylation in normal development, health, and disease. Before the start of this thesis, whole genome or genome-wide methylation maps were non-existent. Researchers dissected the methylation patterns of single loci under various conditions or in various cell states. Currently, whole genome patterns of 5-methylcytosine (5mC) is being catalogued across human tissues and disease states. What's more, within the past five years the field has established the importance of 5-hydroxymethylcytosine (5hmC), a novel genomic mark, in early development and has elucidated several mechanisms of active DNA demethylation. We now have a more complete understanding of the factors involved in establishing, maintaining, reading, and erasing DNA methylation. Several questions (among many) are still outstanding in the field.

What is the informative unit of DNA methylation?

A question that has always been at the forefront of the DNA methylation field is whether individual CpGs or multiple CpGs together account for functional regulation. In other words, can each individual CpG inform biology or does a region have to reach a methylation threshold to effect its role in a cell? Until recently, Differentially Methylated Regions (DMRs) and the average methylation status of all CpGs within a CpG Island (CGI), were thought to be the only informative unit for transcriptional regulation. The high density of CpGs within 200-500 base pair CGIs alluded to a mechanism in which the absolute level of CpG methylation across the entire region dictated gene regulation. Furthermore, because CGIs are enriched at gene promoters, the natural inclination has been to ascribe the CGI unit with a regulatory function controlled by CpG methylation. The focus on CGIs is an unintended consequence of the introduction of genome sequencing and the desire to interrogate as many CpG dinucleotides as possible relative to experimental costs. Naturally, focusing on has CGIs allowed for a fairly comprehensive, genome-wide evaluation but biased the field toward interrogating promoter methylation. Covariance analysis of the methylation status of CpGs and their nearest neighborhoods has strengthened the notion that the CpGs within large genome regions have a similar methylation status. These studies have shown

that if a CpG is highly methylated, the probability is high that its nearest neighbor is also methylated. Furthermore, many studies have shown that entire CGIs can show differential methylation between normal cell types and that cancer, in particular, loses methylation boundaries which results in large, megabase blocks of hypomethylated regions. The methylation profiling of adrenocortical neoplasias featured in this thesis, also demonstrates that larger regions, namely CGIs, can exhibit large methylation changes in disease models.

On the other hand, many examples abound from the 1980s and 1990s showing that the methylation status of an individual CpG is sufficient to disrupt the binding of transcription factors or to recruit methyl-binding proteins. Although the entire region could often not be analyzed, it is now clear that individual CpGs can affect transcription. The DNA-binding factor CTCF insulates the effects of distal enhancers from promoters to regulate expression and contains multiple potential CpG sites within its binding motif. The methylation of one of these CpG sites prevents DNA binding and gene repression. The methylation of a single CpG can also recruit proteins with methyl-binding domains. Their binding to methylated DNA recruits histone modifying proteins to establish repressive chromatin states. My study on the changing methylation landscape of maturing motor neurons provides evidence for both models suggesting individual CpGs can be informative and that the DNA methylation state of a cluster of CpGs confers biological information. Many of the CpGs I observed to gain or lose methylation between D0 and D4 motor neurons were the only differentially methylated CpG associated with a gene. The dramatic change in methylation observed (>25%) as well as the importance of many of the nearest genes suggests these CpGs are biologically important and play a regulatory role. Many other differentially methylated CpGs, however, were located within 200 base pairs of each other (clustered). Most often, these CpGs all showed a change in methylation status in the same direction though some instances were observed in which one CpG gained methylation and its nearest neighbor lost methylation. The former case supports the idea that large genomic regions are the informative unit while the later case suggests each CpG plays a unique role. Why nearby CpGs would both gain and lose methylation, however, is yet to be determined. From my studies, it is evident that single CpG and multi-CpG methylation changes occur and both are most likely biologically relevant.

Why does CpG methylation not correlate with expression on the global level?

The first studies of the functional role of DNA methylation definitively showed that methylation is associated with a transcriptionally repressive state. We have since come to better understand the dynamic interplay between chromatin state and DNA methylation state, but still are unable to predict transcriptional expression based on DNA methylation. What's more, global methylome and transcriptome studies show very poor correlation between DNA methylation and expression. The poor overall correlation could be due to several reason, all of which reflect a gap in our understanding and represent unique opportunities for future discovery.

One explanation for the poor correlation between methylation and expression could simply be due to our limited knowledge of the information content of each CpG. In other words, we still are unsure which CpGs regulate which genes. 3D chromatin mapping experiments have shown that the genome forms complex physical interactions between distal genomic regions. For example, enhancers have been shown to interact with gene prompters tens of thousands of kilobases away. Similarly, proximal CpGs may not provide the most predictive power for determining expression. The genomic description of the differentially methylated CpGs identified in my study of maturing motor neurons (Figure 4.S1) shows that most methylation changes occur 10-15 kilobases from transcription start sites (TSSs). Most putative enhancers are also found at roughly the same distance from TSSs as these differentially methylated CpGs. As more whole genome datasets across multiple cell types become available, correlations can be made between the methylation status of every CpG and the expression level of every gene to identify regulatory CpGs.

Moreover, the biologically significance of methylation levels further confounds our interpretation of DNA methylation and precludes correlative studies. In theory, because of the binary nature of DNA methylation, the methylation of a CpG within one cell can exist in only one of three states: 100% methylated (both alleles methylated), 50% methylated (one allele methylated, one allele unmethylated), or 0% methylated (both alleles unmethylated). Empirically, CpG methylation is bimodally distributed between fully methylated (100%) and fully unmethylated (0%) states. Imprinted loci that are 50% methylated are the exception. Furthermore, intermediate methylation states, percent methylation that is not 0, 50, or 100%, are also observed. The observation of intermediate methylation states (e.g. 30% or 70%) could

mean the presence of technical artifacts. Since all studies to date are not single cell and, therefore percent methylation of a CpG is the average methylation of all alleles in all cells, the observation of intermediate methylation most likely suggests the CpG methylation state varies among cells. The lack of correlation between methylation and expression could, therefore, be due to the fact that methylation is a binary variable while gene expression is a discrete variable. Single-cell transcription and methylation mapping experiments are necessary to more accurately describe the relationship between individual CpG methylation and gene expression.

What is the role of CpH methylation?

The methylation of cytosine outside of the CpG context was first demonstrated to occur in embryonic stem (ES) cells and subsequently lost upon differentiation. Recently, the presence of CpH (H=A, C, or T) methylation was discovered in the mammalian brain. CpH methylation occurs primarily within the gene body, whereas CpG methylation is distributed across the genome. Like methylation at CpGs, CpH methylation in the brain is negatively correlated to gene transcript abundance. In ES cells however, CpH methylation is positively correlated to gene transcript abundance.

The key to understanding the role of CpH methylation may lie in its asymmetry. Because the CpA, CpC, and CpT nucleotide pairs do not contain a cytosine on its opposite strand, methylation can only occur on one strand. When a cell with CpH methylation divides, one cell could inherit the hemi-methylated chromosome pairs while another cell could inherit the completely unmethylated chromosomes, effectively erasing the CpH methylation. Stem cells are known to undergo obligatory asymmetrical replication in which one pluripotent or multipotent cell divides into one mother cell that is identical to the original stem cell and into one daughter cell that is differentiated. The asymmetric division of CpH methylation may be a marker for stemness and instruct the dividing cell to remain in the pluripotent or multipotent state. The daughter cell that receives the unmethylated DNA strands would be signaled to differentiate due to the lack of CpH methylation. Experiments that can specifically disrupt CpH methylation in ES cells or synthetically introduce CpH methyl marks are necessary to test this hypothesis and elucidate CpH function.

How is DNA methylation directed?

Although we have a thorough understanding of the methyltransferase machinery, we have yet to determine how DNA methylation is directed to specific genomic sequences. *Dnmt1* is primarily responsible for maintaining DNA methylation states through cell divisions by faithfully adding a methyl group to the cytosines of hemi-methylated, recently replicated, DNA. *Dnmt3b* is active early in development where it adds methyl groups to completely unmethylated pericentric regions and some repetitive elements. *Dnmt3a* more broadly methylates DNA across various genomic features throughout development and in adult tissues. *Dnmt3l* confers some specificity to direct *Dnmt3a* to imprinted loci and repetitive elements. The general specificity of each of the *de novo* methyltransferases can be explained by differences in their DNA binding domains or interaction with *Dnmt3l*, but factors that determine cell type-specific deposition of methyl groups is still unclear.

An important question underlying how DNA methylation status is determined is to understand what is the default methylation state of DNA: is the natural tendency of the methyltransferase machinery to methylate all cytosines with some sequences excluding the Dnmts or is the default state of the cell to be unmethylated with Dnmts directed to specific sequences by chaperones or cofactors? This question is further complicated when considering the process of cell division and the process of passive versus active demethylation. Studies have shown that *Sp1* is associated with unmethylated sequences, especially CGIs. The additional of *Sp1* binding sites to DNA that is methylated results in the prevention (or loss) of methylation while the mutation of *Sp1* binding sites in unmethylated DNA results in gains of methylation within that region. Experiments with the Epstein-Barr virus suggest that binding of transcription factors alone is sufficient to prevent methylation of DNA in the case of replication-dependent mechanisms. These studies suggest a model where DNA-binding factors exclude the methyltransferase machinery from binding and methylating DNA.

Studies of the *Oct4* and *p15* promoters, however, suggest a different model in which specific factors direct the addition of methylation at unmethylated regions. In this case, the *Oct4* promoter acquires methylation upon differentiation. The methylation state in differentiated cells results from the binding of G9A and subsequent recruitment of DNMT3A and DNMT3B to the *Oct4* promoter.

Aberrant methylation gains in cancer can also teach us about the default state of DNA. *p15*, a tumor-suppressor genes, is highly methylated in the breast cancer cell line, MCF7. This methylation seems to be due to DNMT3A and mediated by ZNF217. Knockdown of *ZNF217* by siRNA results in complete loss of methylation at the *p15* promoter suggesting that ZNF217 directs the methylation at the *p15* promoter by recruiting DNMT3A.

Most likely, depending on the cellular context, both models described occur. My work in mapping methylation in maturing motor neuron suggests that specific factors can direct the removal and additional of DNA methylation at specific loci in an active, non-replication-dependent manner.

REFERENCES

Aapola U, Kawasaki K, Scott HS, Ollila J, Vihinen M, Heino M, Shintani A, Kawasaki K, Minoshima S, Krohn K et al. 2000. Isolation and initial characterization of a novel zinc finger gene, DNMT3L, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics* **65**(3): 293-298.

Adey A, Shendure J. 2012. Ultra-low-input, tagmentation-based whole-genome bisulfite sequencing. *Genome Res* **22**(6): 1139-1143.

Akerman I, van Arensbergen J, Ferrer J. 2011. Removing the brakes on cell identity. *Dev Cell* **20**(4): 411-412.

Altman J. BSA. 2001. *Development of the human spinal cord: an interpretation based on experimental studies in animals*. Oxford University Press, New York.

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. 1999. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* **23**(2): 185-188.

Anand-Ivell R, Heng K, Hafen B, Setchell B, Ivell R. 2009. Dynamics of INSL3 peptide expression in the rodent testis. *Biol Reprod* **81**(3): 480-487.

Anttonen M, Ketola I, Parviainen H, Pusa AK, Heikinheimo M. 2003. FOG-2 and GATA-4 are coexpressed in the mouse ovary and can modulate Mullerian-inhibiting substance expression. *BiolReprod* **68**(4): 1333-1340.

Aranda P, Agirre X, Ballestar E, Andreu EJ, Roman-Gomez J, Prieto I, Martin-Subero JI, Cigudosa JC, Siebert R, Esteller M et al. 2009. Epigenetic signatures associated with different levels of differentiation potential in human stem cells. *PLoS One* **4**(11): e7809.

Bach LA, Hsieh S, Brown AL, Rechler MM. 1994. Recombinant human insulin-like growth factor (IGF)-binding protein-6 inhibits IGF-II-induced differentiation of L6A1 myoblasts. *Endocrinology* **135**(5): 2168-2176.

Bach LA, Hsieh S, Sakano K, Fujiwara H, Perdue JF, Rechler MM. 1993. Towards identification of a binding site on insulin-like growth factor-II for IGF-binding proteins. *Adv Exp Med Biol* **343**: 55-61.

Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM. 2009. Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol* **27**(4): 361-368.

Bandiera R, Vidal VP, Motamedi FJ, Clarkson M, Sahut-Barnola I, von Gise A, Pu WT, Hohenstein P, Martinez A, Schedl A. 2013. WT1 Maintains Adrenal-Gonadal Primordium Identity and Marks a Population of AGP-like Progenitors within the Adrenal Gland. *Dev Cell* **27**(1): 5-18.

Beck S. 2010. Taking the measure of the methylome. *Nat Biotechnol* **28**(10): 1026-1028.

Bernichtein S, Alevizaki M, Huhtaniemi I. 2008a. Is the adrenal cortex a target for gonadotropins? *Trends EndocrinolMetab* **19**(7): 231-238.

Bernichtein S, Peltoketo H, Huhtaniemi I. 2009. Adrenal hyperplasia and tumours in mice in connection with aberrant pituitary-gonadal function. *Mol Cell Endocrinol* **300**(1-2): 164-168.

Bernichtein S, Petretto E, Jamieson S, Goel A, Aitman TJ, Mangion JM, Huhtaniemi IT. 2007. Adrenal gland tumorigenesis after gonadectomy in mice is a complex genetic trait driven by epistatic loci. *Endocrinology* **149**(2): 651-661.

-. 2008b. Adrenal gland tumorigenesis after gonadectomy in mice is a complex genetic trait driven by epistatic loci. *Endocrinology* **149**(2): 651-661.

Bessman MJ, Lehman IR, Adler J, Zimmerman SB, Simms ES, Kornberg A. 1958. Enzymatic Synthesis of Deoxyribonucleic Acid. Iii. The Incorporation of Pyrimidine and Purine Analogues into Deoxyribonucleic Acid. *Proc Natl Acad Sci U S A* **44**(7): 633-640.

Best D, Sahlender DA, Walther N, Peden AA, Adams IR. 2008. Sdmg1 is a conserved transmembrane protein associated with germ cell sex determination and germline-soma interactions in mice. *Development (Cambridge, England)* **135**(8): 1415-1425.

Bestor T, Laudano A, Mattaliano R, Ingram V. 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of molecular biology* **203**(4): 971-983.

Bestor TH, Ingram VM. 1983. Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc Natl Acad Sci U S A* **80**(18): 5559-5563.

Beuschlein F, Galac S, Wilson DB. 2012. Animal models of adrenocortical tumorigenesis. *Mol Cell Endocrinol* **351**(1): 78-86.
Bielinska M, Genova E, Boime I, Parviainen H, Kiiveri S, Leppaluoto J, Rahman N, Heikinheimo M, Wilson DB. 2005. Gonadotropin-induced adrenocortical neoplasia in NU/J nude mice. *Endocrinology* **146**(9): 3975-3984.
Bielinska M, Kiiveri S, Parviainen H, Mannisto S, Heikinheimo M, Wilson DB. 2006a. Gonadectomy-induced adrenocortical neoplasia in the domestic ferret (Mustela putorius furo) and laboratory mouse. *VetPathol* **43**(2): 97-117.
-. 2006b. Gonadectomy-induced adrenocortical neoplasia in the domestic ferret (Mustela putorius furo) and laboratory mouse. *Veterinary pathology* **43**(2): 97-117.
Bielinska M, Parviainen H, Kiiveri S, Heikinheimo M, Wilson DB. 2009. Review paper: origin and molecular pathology of adrenocortical neoplasms. *Veterinary pathology* **46**(2): 194-210.
Bielinska M, Parviainen H, Porter-Tinge SB, Kiiveri S, Genova E, Rahman N, Huhtaniemi IT, Muglia LJ, Heikinheimo M, Wilson DB. 2003a. Mouse strain susceptibility to gonadectomy-induced adrenocortical tumor formation correlates with the expression of GATA-4 and luteinizing hormone receptor. *Endocrinology* **144**: 4123-4133.
-. 2003b. Mouse strain susceptibility to gonadectomy-induced adrenocortical tumor formation correlates with the expression of GATA-4 and luteinizing hormone receptor. *Endocrinology* **144**(9): 4123-4133.
Bird A. 2002. DNA methylation patterns and epigenetic memory. *Genes Dev* **16**(1): 6-21.
Bird A, Taggart M, Frommer M, Miller OJ, Macleod D. 1985. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* **40**(1): 91-99.
Bird AP, Wolffe AP. 1999. Methylation-induced repression--belts, braces, and chromatin. *Cell* **99**(5): 451-454.
Bock C, Tomazou EM, Brinkman AB, Muller F, Simmer F, Gu H, Jager N, Gnirke A, Stunnenberg HG, Meissner A. 2010. Quantitative comparison of genome-wide DNA methylation mapping technologies. *Nat Biotechnol* **28**(10): 1106-1114.
Bonilla E, Xu EY. 2008. Identification and characterization of novel mammalian spermatogenic genes conserved from fly to human. *Molecular human reproduction* **14**(3): 137-142.
Borgel J, Guibert S, Li Y, Chiba H, Schubeler D, Sasaki H, Forne T, Weber M. 2010. Targets and dynamics of promoter DNA methylation during early mouse development. *Nat Genet* **42**(12): 1093-1100.
Bourc'his D, Bestor TH. 2004. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431**(7004): 96-99.
Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH. 2001. Dnmt3L and the establishment of maternal genomic imprints. *Science* **294**(5551): 2536-2539.
Boyes J, Bird A. 1991. DNA methylation inhibits transcription indirectly via a methyl-CpG binding protein. *Cell* **64**(6): 1123-1134.
Branco MR, Ficz G, Reik W. 2012. Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet* **13**(1): 7-13.
Brenet F, Moh M, Funk P, Feierstein E, Viale AJ, Socci ND, Scandura JM. 2011. DNA methylation of the first exon is tightly linked to transcriptional silencing. *PLoS One* **6**(1): e14524.
Brinkman AB, Simmer F, Ma K, Kaan A, Zhu J, Stunnenberg HG. 2010. Whole-genome DNA methylation profiling using MethylCap-seq. *Methods* **52**(3): 232-236.
Broske AM, Vockentanz L, Kharazi S, Huska MR, Mancini E, Scheller M, Kuhl C, Enns A, Prinz M, Jaenisch R et al. 2009. DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat Genet* **41**(11): 1207-1215.
Buaas FW, Gardiner JR, Clayton S, Val P, Swain A. 2012. In vivo evidence for the crucial role of SF1 in steroid-producing cells of the testis, ovary and adrenal gland. *Development* **139**(24): 4661-4570.
Buiting K. 2010. Prader-Willi syndrome and Angelman syndrome. *American journal of medical genetics Part C, Seminars in medical genetics* **154C**(3): 365-376.
Burke RE, Glenn LL. 1996. Horseradish peroxidase study of the spatial and electrotonic distribution of group Ia synapses on type-identified ankle extensor motoneurons in the cat. *The Journal of comparative neurology* **372**(3): 465-485.

Carrascal L, Nieto-Gonzalez JL, Cameron WE, Torres B, Nunez-Abades PA. 2005. Changes during the postnatal development in physiological and anatomical characteristics of rat motoneurons studied in vitro. *Brain research Brain research reviews* **49**(2): 377-387.
Chahrour M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J, Zoghbi HY. 2008. MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320**(5880): 1224-1229.
Chen IB, Rathi VK, DeAndrade DS, Jay PY. 2013. Association of genes with physiological functions by comparative analysis of pooled expression microarray data. *Physiol Genomics* **45**(2): 69-78.
Chen WY, Zeng X, Carter MG, Morrell CN, Chiu Yen RW, Esteller M, Watkins DN, Herman JG, Mankowski JL, Baylin SB. 2003. Heterozygous disruption of Hic1 predisposes mice to a gender-dependent spectrum of malignant tumors. *Nat Genet* **33**(2): 197-202.
Chestnut BA, Chang Q, Price A, Lesuisse C, Wong M, Martin LJ. 2011. Epigenetic regulation of motor neuron cell death through DNA methylation. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **31**(46): 16619-16636.
Christman JK, Price P, Pedrinan L, Acs G. 1977. Correlation between hypomethylation of DNA and expression of globin genes in Friend erythroleukemia cells. *European journal of biochemistry / FEBS* **81**(1): 53-61.
Chrusciel M, Vuorenoja S, Mohanty B, Rivero-Muller A, Li X, Toppari J, Huhtaniemi I, Rahman NA. 2013. Transgenic GATA-4 expression induces adrenocortical tumorigenesis in C57Bl/6 mice. *J Cell Sci* **126**(Pt 8): 1845-1857.
Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF. 1997. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* **277**(5334): 1996-2000.
Colak D, Zaninovic N, Cohen MS, Rosenwaks Z, Yang WY, Gerhardt J, Disney MD, Jaffrey SR. 2014. Promoter-bound trinucleotide repeat mRNA drives epigenetic silencing in fragile X syndrome. *Science* **343**(6174): 1002-1005.
Coulondre C, Miller JH, Farabaugh PJ, Gilbert W. 1978. Molecular basis of base substitution hotspots in Escherichia coli. *Nature* **274**(5673): 775-780.
Cullheim S, Fleshman JW, Glenn LL, Burke RE. 1987. Membrane area and dendritic structure in type-identified triceps surae alpha motoneurons. *The Journal of comparative neurology* **255**(1): 68-81.
Dalerba P, Kalisky T, Sahoo D, Rajendran PS, Rothenberg ME, Leyrat AA, Sim S, Okamoto J, Johnston DM, Qian D et al. 2011. Single-cell dissection of transcriptional heterogeneity in human colon tumors. *Nat Biotechnol* **29**(12): 1120-1127.
Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, Davies H, Edkins S, Hardy C, Latimer C et al. 2010. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**(7279): 360-363.
Dawlaty MM, Breiling A, Le T, Raddatz G, Barrasa MI, Cheng AW, Gao Q, Powell BE, Li Z, Xu M et al. 2013. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell* **24**(3): 310-323.
Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW, Gao Q, Kim J, Choi SW, Page DC et al. 2011. Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell stem cell* **9**(2): 166-175.
Day JJ, Sweatt JD. 2010. DNA methylation and memory formation. *Nature neuroscience* **13**(11): 1319-1323.
de Greef JC, Wang J, Balog J, den Dunnen JT, Frants RR, Straasheijm KR, Aytekin C, van der Burg M, Duprez L, Ferster A et al. 2011. Mutations in ZBTB24 are associated with immunodeficiency, centromeric instability, and facial anomalies syndrome type 2. *American journal of human genetics* **88**(6): 796-804.
de Jong MK, Schoemaker NJ, Mol JA. 2013. Expression of sfrp1 and activation of the Wnt pathway in the adrenal glands of healthy ferrets and neutered ferrets with hyperadrenocorticism. *Vet J* **196**(2): 176-180.
de Jong MK, ten Asbroek EEM, Sleiderink AJ, Conley AJ, Mol JA, Schoemaker NJ. 2014. Gonadectomy-related adrenocortical tumors in ferrets demonstrate increased expression of androgen and estrogen synthesizing enzymes together with high inhibin expression. *Domestic Animal Endocrinology* **48**(0): 42-47.
Deng J, Shoemaker R, Xie B, Gore A, LeProust EM, Antosiewicz-Bourget J, Egli D, Maherali N, Park IH, Yu J et al. 2009. Targeted bisulfite sequencing reveals changes in DNA methylation associated with nuclear reprogramming. *Nat Biotechnol* **27**(4): 353-360.

Denolet E, De Gendt K, Allemeersch J, Engelen K, Marchal K, Van Hummelen P, Tan KA, Sharpe RM, Saunders PT, Swinnen JV et al. 2006. The effect of a sertoli cell-selective knockout of the androgen receptor on testicular gene expression in prepubertal mice. *Mol Endocrinol* **20**(2): 321-334.
Desrosiers RC, Mulder C, Fleckenstein B. 1979. Methylation of Herpesvirus saimiri DNA in lymphoid tumor cell lines. *Proc Natl Acad Sci U S A* **76**(8): 3839-3843.
Dhawan S, Georgia S, Tschen SI, Fan G, Bhushan A. 2011a. Pancreatic beta cell identity is maintained by DNA methylation-mediated repression of Arx. *Dev Cell* **20**(4): 419-429.
-. 2011b. Pancreatic beta cell identity is maintained by DNA methylation-mediated repression of Arx. *Dev Cell* **20**(4): 419-429.
Diep D, Plongthongkum N, Gore A, Fung HL, Shoemaker R, Zhang K. 2012. Library-free methylation sequencing with bisulfite padlock probes. *Nat Methods* **9**(3): 270-272.
Dietrich D, Lesche R, Tetzner R, Krispin M, Dietrich J, Haedicke W, Schuster M, Kristiansen G. 2009. Analysis of DNA methylation of multiple genes in microdissected cells from formalin-fixed and paraffin-embedded tissues. *J Histochem Cytochem* **57**(5): 477-489.
Doghman M, Lalli E. 2009. A matter of dosage: SF-1 in adrenocortical development and cancer. *Ann Endocrinol (Paris)* **70**(3): 148-152.
Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D, Danenberg PV, Laird PW. 2000. MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic acids research* **28**(8): E32.
Easwaran HP, Schermelleh L, Leonhardt H, Cardoso MC. 2004. Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO reports* **5**(12): 1181-1186.
Eckhardt F, Lewin J, Cortese R, Rakyan VK, Attwood J, Burger M, Burton J, Cox TV, Davies R, Down TA et al. 2006. DNA methylation profiling of human chromosomes 6, 20 and 22. *Nat Genet* **38**(12): 1378-1385.
Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, Gehrke C. 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic acids research* **10**(8): 2709-2721.
Elliott E, Ezra-Nevo G, Regev L, Neufeld-Cohen A, Chen A. 2010. Resilience to social stress coincides with functional DNA methylation of the Crf gene in adult mice. *Nature neuroscience* **13**(11): 1351-1353.
Espina V, Wulfkuhle JD, Calvert VS, VanMeter A, Zhou W, Coukos G, Geho DH, Petricoin EF, 3rd, Liotta LA. 2006. Laser-capture microdissection. *Nat Protoc* **1**(2): 586-603.
Esteller M. 2007. Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet* **8**(4): 286-298.
Fan G, Martinowich K, Chin MH, He F, Fouse SD, Hutnick L, Hattori D, Ge W, Shen Y, Wu H et al. 2005. DNA methylation controls the timing of astrogliogenesis through regulation of JAK-STAT signaling. *Development (Cambridge, England)* **132**(15): 3345-3356.
Feinberg AP, Ohlsson R, Henikoff S. 2006a. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* **7**(1): 21-33.
-. 2006b. The epigenetic progenitor origin of human cancer. *NatRevGenet* **7**(1): 21-33.
Feinberg AP, Vogelstein B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301**(5895): 89-92.
Feng J, Zhou Y, Campbell SL, Le T, Li E, Sweatt JD, Silva AJ, Fan G. 2010. Dnmt1 and Dnmt3a maintain DNA methylation and regulate synaptic function in adult forebrain neurons. *Nature neuroscience* **13**(4): 423-430.
Fidler IJ, Gersten DM, Kripke ML. 1979. Influence of immune status on the metastasis of three murine fibrosarcomas of different immunogenicities. *Cancer Res* **39**(10): 3816-3821.
Figueroa ME, Skrabanek L, Li Y, Jiemjit A, Fandy TE, Paietta E, Fernandez H, Tallman MS, Greally JM, Carraway H et al. 2009. MDS and secondary AML display unique patterns and abundance of aberrant DNA methylation. *Blood* **114**(16): 3448-3458.
Fradin A, Manley JL, Prives CL. 1982. Methylation of simian virus 40 Hpa II site affects late, but not early, viral gene expression. *Proc Natl Acad Sci U S A* **79**(17): 5142-5146.
Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. 1992. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* **89**(5): 1827-1831.

Garcia-Ortiz JE, Pelosi E, Omari S, Nedorezov T, Piao Y, Karmazin J, Uda M, Cao A, Cole SW, Forabosco A et al. 2009. Foxl2 functions in sex determination and histogenesis throughout mouse ovary development. *BMC Dev Biol* **9**: 36.
Gardiner-Garden M, Frommer M. 1987. CpG islands in vertebrate genomes. *J Mol Biol* **196**(2): 261-282.
Georges A, Auguste A, Bessiere L, Vanet A, Todeschini AL, Veitia RA. 2013. FOXL2: a central transcription factor of the ovary. *J Mol Endocrinol*.
Gertz J, Varley KE, Reddy TE, Bowling KM, Pauli F, Parker SL, Kucera KS, Willard HF, Myers RM. 2011. Analysis of DNA methylation in a three-generation family reveals widespread genetic influence on epigenetic regulation. *PLoS Genet* **7**(8): e1002228.
Gilbert SF. 2012. Commentary: 'The epigenotype' by C.H. Waddington. *International journal of epidemiology* **41**(1): 20-23.
-. 2013. *Development Biology*. Sinauer Associates, Inc., Sunderland, Massachusetts.
Glenn CC, Nicholls RD, Robinson WP, Saitoh S, Niikawa N, Schinzel A, Horsthemke B, Driscoll DJ. 1993. Modification of 15q11-q13 DNA methylation imprints in unique Angelman and Prader-Willi patients. *Hum Mol Genet* **2**(9): 1377-1382.
Glister C, Satchell L, Bathgate RA, Wade JD, Dai Y, Ivell R, Anand-Ivell R, Rodgers RJ, Knight PG. 2013. Functional link between bone morphogenetic proteins and insulin-like peptide 3 signaling in modulating ovarian androgen production. *Proc Natl Acad Sci U S A* **110**(15): E1426-1435.
Globisch D, Munzel M, Muller M, Michalakis S, Wagner M, Koch S, Bruckl T, Biel M, Carell T. 2010. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One* **5**(12): e15367.
Goelz SE, Vogelstein B, Hamilton SR, Feinberg AP. 1985. Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science* **228**(4696): 187-190.
Gold M, Hurwitz J, Anders M. 1963. The Enzymatic Methylation of Rna and DNA, Ii. On the Species Specificity of the Methylation Enzymes. *Proc Natl Acad Sci U S A* **50**(1): 164-169.
Gowher H, Liebert K, Hermann A, Xu G, Jeltsch A. 2005. Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J Biol Chem* **280**(14): 13341-13348.
Grippo P, Iaccarino M, Parisi E, Scarano E. 1968. Methylation of DNA in developing sea urchin embryos. *Journal of molecular biology* **36**(2): 195-208.
Gruenbaum Y, Cedar H, Razin A. 1982. Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* **295**(5850): 620-622.
Gruenbaum Y, Naveh-Many T, Cedar H, Razin A. 1981. Sequence specificity of methylation in higher plant DNA. *Nature* **292**(5826): 860-862.
Gu H, Bock C, Mikkelsen TS, Jager N, Smith ZD, Tomazou E, Gnirke A, Lander ES, Meissner A. 2010. Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat Methods* **7**(2): 133-136.
Guo JU, Ma DK, Mo H, Ball MP, Jang MH, Bonaguidi MA, Balazer JA, Eaves HL, Xie B, Ford E et al. 2011a. Neuronal activity modifies the DNA methylation landscape in the adult brain. *Nature neuroscience* **14**(10): 1345-1351.
Guo JU, Su Y, Zhong C, Ming GL, Song H. 2011b. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145**(3): 423-434.
Guy J, Gan J, Selfridge J, Cobb S, Bird A. 2007. Reversal of neurological defects in a mouse model of Rett syndrome. *Science* **315**(5815): 1143-1147.
Hansen KD, Timp W, Bravo HC, Sabunciyan S, Langmead B, McDonald OG, Wen B, Wu H, Liu Y, Diep D et al. 2011. Increased methylation variation in epigenetic domains across cancer types. *Nat Genet* **43**(8): 768-775.
Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405**(6785): 486-489.
Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y et al. 2010. Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotechnol* **28**(10): 1097-1105.
Hartung T, Zhang L, Kanwar R, Khrebtukova I, Reinhardt M, Wang C, Therneau TM, Banck MS, Schroth GP, Beutler AS. 2012. Diametrically opposite methylome-transcriptome relationships in high- and

low-CpG promoter genes in postmitotic neural rat tissue. *Epigenetics : official journal of the DNA Methylation Society* **7**(5): 421-428.
Heglind M, Cederberg A, Aquino J, Lucas G, Ernfors P, Enerback S. 2005. Lack of the central nervous system- and neural crest-expressed forkhead gene Foxs1 affects motor function and body weight. *Mol Cell Biol* **25**(13): 5616-5625.
Heikinheimo M, Scandrett JM, Wilson DB. 1994. Localization of transcription factor GATA-4 to regions of the mouse embryo involved in cardiac development. *Dev Biol* **164**: 361-373.
Heikkila M, Peltoketo H, Leppaluoto J, Ilves M, Vuolteenaho O, Vainio S. 2002. Wnt-4 deficiency alters mouse adrenal cortex function, reducing aldosterone production. *Endocrinology* **143**(11): 4358-4365.
Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular cell* **38**(4): 576-589.
Hellman A, Chess A. 2007. Gene body-specific methylation on the active X chromosome. *Science* **315**(5815): 1141-1143.
Hemberger M, Dean W, Reik W. 2009. Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. *Nature reviews Molecular cell biology* **10**(8): 526-537.
Heppner GH. 1984. Tumor heterogeneity. *Cancer Res* **44**(6): 2259-2265.
Hernandez DG, Nalls MA, Gibbs JR, Arepalli S, van der Brug M, Chong S, Moore M, Longo DL, Cookson MR, Traynor BJ et al. 2011. Distinct DNA methylation changes highly correlated with chronological age in the human brain. *Hum Mol Genet* **20**(6): 1164-1172.
Herrmann A, Haake A, Ammerpohl O, Martin-Guerrero I, Szafranski K, Stemshorn K, Nothnagel M, Kotsopoulos SK, Richter J, Warner J et al. 2011. Pipeline for large-scale microdroplet bisulfite PCR-based sequencing allows the tracking of hepitype evolution in tumors. *PLoS One* **6**(7): e21332.
Hoivik EA, Bjanesoy TE, Bakke M. 2013. Epigenetic regulation of the gene encoding steroidogenic factor-1. *Mol Cell Endocrinol* **371**(1-2): 133-139.
Hoivik EA, Bjanesoy TE, Mai O, Okamoto S, Minokoshi Y, Shima Y, Morohashi K, Boehm U, Bakke M. 2011. DNA methylation of intronic enhancers directs tissue-specific expression of steroidogenic factor 1/adrenal 4 binding protein (SF-1/Ad4BP). *Endocrinology* **152**(5): 2100-2112.
Holmes RL. 1961. The adrenal glands of the ferret. *JAnat* **95**: 325-336.
Hu L, Monteiro A, Johnston H, King P, O'Shaughnessy PJ. 2007. Expression of Cyp21a1 and Cyp11b1 in the fetal mouse testis. *Reproduction* **134**(4): 585-591.
Irizarry RA, Ladd-Acosta C, Carvalho B, Wu H, Brandenburg SA, Jeddeloh JA, Wen B, Feinberg AP. 2008. Comprehensive high-throughput arrays for relative methylation (CHARM). *Genome Res* **18**(5): 780-790.
Ivell R, Anand-Ivell R. 2011. Biological role and clinical significance of insulin-like peptide 3. *Curr Opin Endocrinol Diabetes Obes* **18**(3): 210-216.
Ivell R, Wade JD, Anand-Ivell R. 2013. INSL3 as a Biomarker of Leydig Cell Functionality. *Biol Reprod* **88**: 147.
Iwamoto K, Bundo M, Ueda J, Oldham MC, Ukai W, Hashimoto E, Saito T, Geschwind DH, Kato T. 2011. Neurons show distinctive DNA methylation profile and higher interindividual variations compared with non-neurons. *Genome Res* **21**(5): 688-696.
Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, Ramsahoye B. 2004. Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Molecular and cellular biology* **24**(20): 8862-8871.
Jakovcevski M, Akbarian S. 2012. Epigenetic mechanisms in neurological disease. *Nat Med* **18**(8): 1194-1204.
Jiao Y, Shi C, Edil BH, de Wilde RF, Klimstra DS, Maitra A, Schulick RD, Tang LH, Wolfgang CL, Choti MA et al. 2011. DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* **331**(6021): 1199-1203.
Jin B, Tao Q, Peng J, Soo HM, Wu W, Ying J, Fields CR, Delmas AL, Liu X, Qiu J et al. 2008. DNA methyltransferase 3B (DNMT3B) mutations in ICF syndrome lead to altered epigenetic modifications and aberrant expression of genes regulating development, neurogenesis and immune function. *Hum Mol Genet* **17**(5): 690-709.

Jin SG, Wu X, Li AX, Pfeifer GP. 2011. Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic acids research* **39**(12): 5015-5024.
Johnsen IK, Slawik M, Shapiro I, Hartmann MF, Wudy SA, Looyenga BD, Hammer GD, Reincke M, Beuschlein F. 2006a. Gonadectomy in mice of the inbred strain CE/J induces proliferation of sub-capsular adrenal cells expressing gonadal marker genes. *JEndocrinol* **190**(1): 47-57.
-. 2006b. Gonadectomy in mice of the inbred strain CE/J induces proliferation of sub-capsular adrenal cells expressing gonadal marker genes. *The Journal of endocrinology* **190**(1): 47-57.
Johnson TB, Coghill RD. 1925. RESEARCHES ON PYRMIDINES. C111. THE DISCOVERY OF 5-METHYL-CYTOSINE IN TUBERCULINIC ACID, THE NUCLEIC ACID OF THE TUBERCLE BACILLUS. *Journal of the American Chemical Society* **47**(11): 7.
Jones PA, Baylin SB. 2002. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* **3**(6): 415-428.
-. 2007. The epigenomics of cancer. *Cell* **128**(4): 683-692.
Jones PA, Taylor SM. 1980. Cellular differentiation, cytidine analogs and DNA methylation. *Cell* **20**(1): 85-93.
Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J, Wolffe AP. 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nat Genet* **19**(2): 187-191.
Kaminskas E, Farrell AT, Wang YC, Sridhara R, Pazdur R. 2005. FDA drug approval summary: azacitidine (5-azacytidine, Vidaza) for injectable suspension. *The oncologist* **10**(3): 176-182.
Khulan B, Thompson RF, Ye K, Fazzari MJ, Suzuki M, Stasiek E, Figueroa ME, Glass JL, Chen Q, Montagna C et al. 2006. Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res* **16**(8): 1046-1055.
Kim JS, Kubota H, Kiuchi Y, Doi K, Saegusa J. 1997. Subcapsular cell hyperplasia and mast cell infiltration in the adrenal cortex of mice: comparative study in 7 inbred strains. *Exp Anim* **46**(4): 303-306.
Kim K, Doi A, Wen B, Ng K, Zhao R, Cahan P, Kim J, Aryee MJ, Ji H, Ehrlich LI et al. 2010. Epigenetic memory in induced pluripotent stem cells. *Nature* **467**(7313): 285-290.
Kim M, Habiba A, Doherty JM, Mills JC, Mercer RW, Huettner JE. 2009. Regulation of mouse embryonic stem cell neural differentiation by retinoic acid. *Dev Biol* **328**(2): 456-471.
Klein CJ, Botuyan MV, Wu Y, Ward CJ, Nicholson GA, Hammans S, Hojo K, Yamanishi H, Karpf AR, Wallace DC et al. 2011. Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. *Nat Genet* **43**(6): 595-600.
Kohli RM, Zhang Y. 2013. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* **502**(7472): 472-479.
Kozlenkov A, Roussos P, Timashpolsky A, Barbu M, Rudchenko S, Bibikova M, Klotzle B, Byne W, Lyddon R, Di Narzo AF et al. 2014. Differences in DNA methylation between human neuronal and glial cells are concentrated in enhancers and non-CpG sites. *Nucleic acids research* **42**(1): 109-127.
Krachulec J, Vetter M, Schrade A, Lobs AK, Bielinska M, Cochran R, Kyronlahti A, Pihlajoki M, Parviainen H, Jay PY et al. 2012. GATA4 is a critical regulator of gonadectomy-induced adrenocortical tumorigenesis in mice. *Endocrinology* **153**(6): 2599-2611.
Kumari D, Usdin K. 2009. Chromatin remodeling in the noncoding repeat expansion diseases. *J Biol Chem* **284**(12): 7413-7417.
Laird PW. 2010a. Principles and challenges of genome-wide DNA methylation analysis. *Nat Rev Genet* **11**(3): 191-203.
-. 2010b. Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet* **11**(3): 191-203.
LaPlant Q, Vialou V, Covington HE, 3rd, Dumitriu D, Feng J, Warren BL, Maze I, Dietz DM, Watts EL, Iniguez SD et al. 2010. Dnmt3a regulates emotional behavior and spine plasticity in the nucleus accumbens. *Nature neuroscience* **13**(9): 1137-1143.
Laufer E, Kesper D, Vortkamp A, King P. 2012. Sonic hedgehog signaling during adrenal development. *Mol Cell Endocrinol* **351**(1): 19-27.
Leonhardt H, Page AW, Weier HU, Bestor TH. 1992. A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71**(5): 865-873.

Levenson JM, Roth TL, Lubin FD, Miller CA, Huang IC, Desai P, Malone LM, Sweatt JD. 2006. Evidence that DNA (cytosine-5) methyltransferase regulates synaptic plasticity in the hippocampus. *J Biol Chem* **281**(23): 15763-15773.
Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, Kandoth C, Payton JE, Baty J, Welch J et al. 2010. DNMT3A mutations in acute myeloid leukemia. *The New England journal of medicine* **363**(25): 2424-2433.
Li D, Mukai K, Suzuki T, Suzuki R, Yamashita S, Mitani F, Suematsu M. 2007. Adrenocortical zonation factor 1 is a novel matricellular protein promoting integrin-mediated adhesion of adrenocortical and vascular smooth muscle cells. *The FEBS journal* **274**(10): 2506-2522.
Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**(6): 915-926.
Li H, Ruan J, Durbin R. 2008a. Mapping short DNA sequencing reads and calling variants using mapping quality scores. *Genome Res* **18**(11): 1851-1858.
Li H, Zhong X, Chau KF, Williams EC, Chang Q. 2011. Loss of activity-induced phosphorylation of MeCP2 enhances synaptogenesis, LTP and spatial memory. *Nature neuroscience* **14**(8): 1001-1008.
Li XJ, Hu BY, Jones SA, Zhang YS, Lavaute T, Du ZW, Zhang SC. 2008b. Directed differentiation of ventral spinal progenitors and motor neurons from human embryonic stem cells by small molecules. *Stem Cells* **26**(4): 886-893.
Li Y, Brewer D, Burke RE, Ascoli GA. 2005. Developmental changes in spinal motoneuron dendrites in neonatal mice. *The Journal of comparative neurology* **483**(3): 304-317.
Lin RK, Hsu HS, Chang JW, Chen CY, Chen JT, Wang YC. 2007. Alteration of DNA methyltransferases contributes to 5'CpG methylation and poor prognosis in lung cancer. *Lung Cancer* **55**(2): 205-213.
Lister R, Mukamel EA, Nery JR, Urich M, Puddifoot CA, Johnson ND, Lucero J, Huang Y, Dwork AJ, Schultz MD et al. 2013. Global epigenomic reconfiguration during mammalian brain development. *Science* **341**(6146): 1237905.
Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**(7271): 315-322.
Lodish HF. 2013. *Molecular cell biology*. W.H. Freeman and Co., New York.
Lofton-Day C, Model F, Devos T, Tetzner R, Distler J, Schuster M, Song X, Lesche R, Liebenberg V, Ebert M et al. 2008. DNA methylation biomarkers for blood-based colorectal cancer screening. *Clinical chemistry* **54**(2): 414-423.
Lubin FD, Roth TL, Sweatt JD. 2008. Epigenetic regulation of BDNF gene transcription in the consolidation of fear memory. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **28**(42): 10576-10586.
Ma DK, Jang MH, Guo JU, Kitabatake Y, Chang ML, Pow-Anpongkul N, Flavell RA, Lu B, Ming GL, Song H. 2009. Neuronal activity-induced Gadd45b promotes epigenetic DNA demethylation and adult neurogenesis. *Science* **323**(5917): 1074-1077.
Martin JP, Bell J. 1943. A Pedigree of Mental Defect Showing Sex-Linkage. *Journal of neurology and psychiatry* **6**(3-4): 154-157.
Martin-Caraballo M, Greer JJ. 1999. Electrophysiological properties of rat phrenic motoneurons during perinatal development. *Journal of neurophysiology* **81**(3): 1365-1378.
Martinowich K, Hattori D, Wu H, Fouse S, He F, Hu Y, Fan G, Sun YE. 2003. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. *Science* **302**(5646): 890-893.
Mazzoni EO, Mahony S, Closser M, Morrison CA, Nedelec S, Williams DJ, An D, Gifford DK, Wichterle H. 2013. Synergistic binding of transcription factors to cell-specific enhancers programs motor neuron identity. *Nature neuroscience* **16**(9): 1219-1227.
McCreedy DA, Rieger CR, Gottlieb DI, Sakiyama-Elbert SE. 2012. Transgenic enrichment of mouse embryonic stem cell-derived progenitor motor neurons. *Stem cell research* **8**(3): 368-378.
McGhee JD, Ginder GD. 1979. Specific DNA methylation sites in the vicinity of the chicken beta-globin genes. *Nature* **280**(5721): 419-420.
McGowan PO, Sasaki A, D'Alessio AC, Dymov S, Labonte B, Szyf M, Turecki G, Meaney MJ. 2009. Epigenetic regulation of the glucocorticoid receptor in human brain associates with childhood abuse. *Nature neuroscience* **12**(3): 342-348.

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB et al. 2008. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454**(7205): 766-770.
Michor F, Polyak K. 2010. The origins and implications of intratumor heterogeneity. *Cancer Prev Res (Phila)* **3**(11): 1361-1364.
Miles GB, Yohn DC, Wichterle H, Jessell TM, Rafuse VF, Brownstone RM. 2004. Functional properties of motoneurons derived from mouse embryonic stem cells. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **24**(36): 7848-7858.
Millar DS, Warnecke PM, Melki JR, Clark SJ. 2002. Methylation sequencing from limiting DNA: embryonic, fixed, and microdissected cells. *Methods* **27**(2): 108-113.
Miller CA, Gavin CF, White JA, Parrish RR, Honasoge A, Yancey CR, Rivera IM, Rubio MD, Rumbaugh G, Sweatt JD. 2010. Cortical DNA methylation maintains remote memory. *Nature neuroscience* **13**(6): 664-666.
Miller LA, Fagerstone KA, Wagner RA, Finkler M. 2013. Use of a GnRH vaccine, GonaCon, for prevention and treatment of adrenocortical disease (ACD) in domestic ferrets. *Vaccine* **31**: 4619-4623.
Mohandas T, Sparkes RS, Shapiro LJ. 1981. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* **211**(4480): 393-396.
Montelius A, Marmigere F, Baudet C, Aquino JB, Enerback S, Ernfors P. 2007. Emergence of the sensory nervous system as defined by Foxs1 expression. *Differentiation* **75**(5): 404-417.
Moretti P, Levenson JM, Battaglia F, Atkinson R, Teague R, Antalffy B, Armstrong D, Arancio O, Sweatt JD, Zoghbi HY. 2006. Learning and memory and synaptic plasticity are impaired in a mouse model of Rett syndrome. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **26**(1): 319-327.
Mork L, Maatouk DM, McMahon JA, Guo JJ, Zhang P, McMahon AP, Capel B. 2012. Temporal differences in granulosa cell specification in the ovary reflect distinct follicle fates in mice. *Biol Reprod* **86**(2): 37.
Morohashi K, Zubair M. 2011. The fetal and adult adrenal cortex. *Molecular and cellular endocrinology* **336**(1-2): 193-197.
Mukai K, Mitani F, Nagasawa H, Suzuki R, Suzuki T, Suematsu M, Ishimura Y. 2003. An inverse correlation between expression of a preprocathepsin B-related protein with cysteine-rich sequences and steroid 11beta -hydroxylase in adrenocortical cells. *J Biol Chem* **278**(19): 17084-17092.
Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* **393**(6683): 386-389.
Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D et al. 2011. Tumour evolution inferred by single-cell sequencing. *Nature* **472**(7341): 90-94.
Navin N, Krasnitz A, Rodgers L, Cook K, Meth J, Kendall J, Riggs M, Eberling Y, Troge J, Grubor V et al. 2010. Inferring tumor progression from genomic heterogeneity. *Genome Res* **20**(1): 68-80.
Nelson ED, Kavalali ET, Monteggia LM. 2008. Activity-dependent suppression of miniature neurotransmission through the regulation of DNA methylation. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **28**(2): 395-406.
Nguyen S, Meletis K, Fu D, Jhaveri S, Jaenisch R. 2007. Ablation of de novo DNA methyltransferase Dnmt3a in the nervous system leads to neuromuscular defects and shortened lifespan. *Developmental dynamics : an official publication of the American Association of Anatomists* **236**(6): 1663-1676.
Nunez-Abades PA, Cameron WE. 1995. Morphology of developing rat genioglossal motoneurons studied in vitro: relative changes in diameter and surface area of somata and dendrites. *The Journal of comparative neurology* **353**(1): 129-142.
O'Rand MG, Widgren EE, Hamil KG, Silva EJ, Richardson RT. 2011. Functional studies of eppin. *Biochem Soc Trans* **39**(5): 1447-1449.
Ohlsson R, Renkawitz R, Lobanenkov V. 2001. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends in genetics : TIG* **17**(9): 520-527.
Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**(3): 247-257.

Okano M, Xie S, Li E. 1998. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat Genet* **19**(3): 219-220.
Oliveira AM, Hemstedt TJ, Bading H. 2012. Rescue of aging-associated decline in Dnmt3a2 expression restores cognitive abilities. *Nature neuroscience* **15**(8): 1111-1113.
Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD et al. 2007. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448**(7154): 714-717.
Ottolenghi C, Omari S, Garcia-Ortiz JE, Uda M, Crisponi L, Forabosco A, Pilia G, Schlessinger D. 2005. Foxl2 is required for commitment to ovary differentiation. *Hum Mol Genet* **14**(14): 2053-2062.
Payne AH, Hales DB. 2004. Overview of steroidogenic enzymes in the pathway from cholesterol to active steroid hormones. *Endocrine reviews* **25**(6): 947-970.
Penagarikano O, Mulle JG, Warren ST. 2007. The pathophysiology of fragile x syndrome. *Annual review of genomics and human genetics* **8**: 109-129.
Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R. 1972. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *The Biochemical journal* **126**(4): 781-790.
Peterson RA, Kiupel M, Bielinska M, Kiiveri S, Heikinheimo M, Capen CC, Wilson DB. 2004. Transcription factor GATA-4 is a marker of anaplasia in adrenocortical neoplasms of the domestic ferret (Mustela putorius furo). *VetPathol* **41**(4): 446-449.
Peterson RA, Kiupel M, Capen CC. 2003. Adrenal cortical carcinomas with myxoid differentiation in the domestic ferret (Mustela putorius furo). *VetPathol* **40**(2): 136-142.
Pihlajoki M, Gretzinger E, Cochran R, Kyronlahti A, Schrade A, Hiller T, Sullivan L, Shoykhet M, Schoeller EL, Brooks MD et al. 2013a. Conditional Mutagenesis of Gata6 in SF1-Positive Cells Causes Gonadal-Like Differentiation in the Adrenal Cortex of Mice. *Endocrinology* **154**(5): 1754-1767.
Pihlajoki M, Heikinheimo M, Wilson DB. 2013b. Never underestimate the complexity of remodeling. *Endocrinology* **154**(12): 4446-4449.
Polo JM, Liu S, Figueroa ME, Kulalert W, Eminli S, Tan KY, Apostolou E, Stadtfeld M, Li Y, Shioda T et al. 2010. Cell type of origin influences the molecular and functional properties of mouse induced pluripotent stem cells. *Nat Biotechnol* **28**(8): 848-855.
Portela A, Esteller M. 2010. Epigenetic modifications and human disease. *Nat Biotechnol* **28**(10): 1057-1068.
Pruitt KD, Tatusova T, Brown GR, Maglott DR. 2012. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res* **40**(Database issue): D130-135.
Qian X, Shen Q, Goderie SK, He W, Capela A, Davis AA, Temple S. 2000. Timing of CNS cell generation: a programmed sequence of neuron and glial cell production from isolated murine cortical stem cells. *Neuron* **28**(1): 69-80.
Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. 2000. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A* **97**(10): 5237-5242.
Reik W. 2007. Stability and flexibility of epigenetic gene regulation in mammalian development. *Nature* **447**(7143): 425-432.
Reik W, Collick A, Norris ML, Barton SC, Surani MA. 1987. Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature* **328**(6127): 248-251.
Reik W, Lewis A. 2005. Co-evolution of X-chromosome inactivation and imprinting in mammals. *Nat Rev Genet* **6**(5): 403-410.
Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, Fujita PA, Diekhans M, Smith KE, Rosenbloom KR, Raney BJ et al. 2010. The UCSC Genome Browser database: update 2010. *Nucleic Acids Res* **38**(Database issue): D613-619.
Rhee I, Jair KW, Yen RW, Lengauer C, Herman JG, Kinzler KW, Vogelstein B, Baylin SB, Schuebel KE. 2000. CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature* **404**(6781): 1003-1007.
Robert MF, Morin S, Beaulieu N, Gauthier F, Chute IC, Barsalou A, MacLeod AR. 2003. DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells. *Nat Genet* **33**(1): 61-65.
Robertson KD. 2005. DNA methylation and human disease. *Nat Rev Genet* **6**(8): 597-610.

Robinton DA, Daley GQ. 2012. The promise of induced pluripotent stem cells in research and therapy. *Nature* **481**(7381): 295-305.
Rohan RM, Ricciarelli E, Kiefer MC, Resnick CE, Adashi EY. 1993. Rat ovarian insulin-like growth factor-binding protein-6: a hormonally regulated theca-interstitial-selective species with limited antigonadotropic activity. *Endocrinology* **132**(6): 2507-2512.
Ryan RJ, Bernstein BE. 2012. Molecular biology. Genetic events that shape the cancer epigenome. *Science* **336**(6088): 1513-1514.
Sapienza C, Peterson AC, Rossant J, Balling R. 1987. Degree of methylation of transgenes is dependent on gamete of origin. *Nature* **328**(6127): 251-254.
Sasaki M, Anast J, Bassett W, Kawakami T, Sakuragi N, Dahiya R. 2003. Bisulfite conversion-specific and methylation-specific PCR: a sensitive technique for accurate evaluation of CpG methylation. *Biochemical and biophysical research communications* **309**(2): 305-309.
Satchell L, Glister C, Bleach EC, Glencross RG, Bicknell AB, Dai Y, Anand-Ivell R, Ivell R, Knight PG. 2013. Ovarian Expression of Insulin-Like Peptide 3 (INSL3) and Its Receptor (RXFP2) During Development of Bovine Antral Follicles and Corpora Lutea and Measurement of Circulating INSL3 Levels During Synchronized Estrous Cycles. *Endocrinology* **154**(5): 1897-1906.
Sato Y, Baba T, Zubair M, Miyabayashi K, Toyama Y, Maekawa M, Owaki A, Mizusaki H, Sawamura T, Toshimori K et al. 2008. Importance of forkhead transcription factor Fkhl18 for development of testicular vasculature. *Molecular reproduction and development* **75**(9): 1361-1371.
Sauvageot CM, Stiles CD. 2002. Molecular mechanisms controlling cortical gliogenesis. *Current opinion in neurobiology* **12**(3): 244-249.
Schauwaers K, De Gendt K, Saunders PT, Atanassova N, Haelens A, Callewaert L, Moehren U, Swinnen JV, Verhoeven G, Verrijdt G et al. 2007. Loss of androgen receptor binding to selective androgen response elements causes a reproductive phenotype in a knockin mouse model. *Proc Natl Acad Sci U S A* **104**(12): 4961-4966.
Schillebeeckx M, Schrade A, Lobs AK, Pihlajoki M, Wilson DB, Mitra RD. 2013. Laser capture microdissection-reduced representation bisulfite sequencing (LCM-RRBS) maps changes in DNA methylation associated with gonadectomy-induced adrenocortical neoplasia in the mouse. *Nucleic acids research* **41**(11): e116.
Schmidt D, Ovitt CE, Anlag K, Fehsenfeld S, Gredsted L, Treier AC, Treier M. 2004. The murine winged-helix transcription factor Foxl2 is required for granulosa cell differentiation and ovary maintenance. *Development* **131**(4): 933-942.
Schoemaker NJ, Teerds KJ, Mol JA, Lumeij JT, Thijssen JH, Rijnberk A. 2002. The role of luteinizing hormone in the pathogenesis of hyperadrenocorticism in neutered ferrets. *MolCell Endocrinol* **197**(1-2): 117-125.
Sen GL, Reuter JA, Webster DE, Zhu L, Khavari PA. 2010. DNMT1 maintains progenitor function in self-renewing somatic tissue. *Nature* **463**(7280): 563-567.
Serre D, Lee BH, Ting AH. 2010. MBD-isolated Genome Sequencing provides a high-throughput and comprehensive survey of DNA methylation in the human genome. *Nucleic Acids Res* **38**(2): 391-399.
Shackleton M, Quintana E, Fearon ER, Morrison SJ. 2009. Heterogeneity in cancer: cancer stem cells versus clonal evolution. *Cell* **138**(5): 822-829.
Sharma S, Kelly TK, Jones PA. 2010. Epigenetics in cancer. *Carcinogenesis* **31**(1): 27-36.
Shima Y, Miyabayashi K, Baba T, Otake H, Oka S, Zubair M, Morohashi K. 2012. Identification of an enhancer in the Ad4BP/SF-1 gene specific for fetal Leydig cells. *Endocrinology* **153**(1): 417-425.
Shirane K, Toh H, Kobayashi H, Miura F, Chiba H, Ito T, Kono T, Sasaki H. 2013. Mouse oocyte methylomes at base resolution reveal genome-wide accumulation of non-CpG methylation and role of DNA methyltransferases. *PLoS genetics* **9**(4): e1003439.
Siegmund KD, Connor CM, Campan M, Long TI, Weisenberger DJ, Biniszkiewicz D, Jaenisch R, Laird PW, Akbarian S. 2007. DNA methylation in the human cerebral cortex is dynamically regulated throughout the life span and involves differentiated neurons. *PLoS One* **2**(9): e895.
Silva EJ, Patrao MT, Tsuruta JK, O'Rand MG, Avellar MC. 2012. Epididymal protease inhibitor (EPPIN) is differentially expressed in the male rat reproductive tract and immunolocalized in maturing spermatozoa. *Mol Reprod Dev* **79**(12): 832-842.
Silverman LR, Demakos EP, Peterson BL, Kornblith AB, Holland JC, Odchimar-Reissig R, Stone RM, Nelson D, Powell BL, DeCastro CM et al. 2002. Randomized controlled trial of azacitidine in

patients with the myelodysplastic syndrome: a study of the cancer and leukemia group B. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **20**(10): 2429-2440.
Simon DP, Hammer GD. 2012. Adrenocortical stem and progenitor cells: implications for adrenocortical carcinoma. *Mol Cell Endocrinol* **351**(1): 2-11.
Sivashanmugam P, Hall SH, Hamil KG, French FS, O'Rand MG, Richardson RT. 2003. Characterization of mouse Eppin and a gene cluster of similar protease inhibitors on mouse chromosome 2. *Gene* **312**: 125-134.
Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* **8**(4): 272-285.
Slott VL, Suarez JD, Poss PM, Linder RE, Strader LF, Perreault SD. 1993. Optimization of the Hamilton-Thorn computerized sperm motility analysis system for use with rat spermatozoa in toxicological studies. *Fundam Appl Toxicol* **21**(3): 298-307.
Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. 2012. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**(7394): 339-344.
Stadler MB, Murr R, Burger L, Ivanek R, Lienert F, Scholer A, van Nimwegen E, Wirbelauer C, Oakeley EJ, Gaidatzis D et al. 2011. DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**(7378): 490-495.
Stein R, Gruenbaum Y, Pollack Y, Razin A, Cedar H. 1982. Clonal inheritance of the pattern of DNA methylation in mouse cells. *Proc Natl Acad Sci U S A* **79**(1): 61-65.
Stevens M, Cheng JB, Li D, Xie M, Hong C, Maire CL, Ligon KL, Hirst M, Marra MA, Costello JF et al. 2013. Estimating absolute methylation levels at single-CpG resolution from methylation enrichment and restriction enzyme sequencing methods. *Genome Res* **23**(9): 1541-1553.
Stratton MR. 2011. Exploring the genomes of cancer cells: progress and promise. *Science* **331**(6024): 1553-1558.
Suzuki H, Watkins DN, Jair KW, Schuebel KE, Markowitz SD, Chen WD, Pretlow TP, Yang B, Akiyama Y, Van Engeland M et al. 2004. Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nat Genet* **36**(4): 417-422.
Svingen T, Beverdam A, Bernard P, McClive P, Harley VR, Sinclair AH, Koopman P. 2007. Sex-specific expression of a novel gene Tmem184a during mouse testis differentiation. *Reproduction (Cambridge, England)* **133**(5): 983-989.
Taby R, Issa JP. 2010. Cancer epigenetics. *CA Cancer J Clin* **60**(6): 376-392.
Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L et al. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**(5929): 930-935.
Taiwo O, Wilson GA, Morris T, Seisenberger S, Reik W, Pearce D, Beck S, Butcher LM. 2012. Methylome analysis using MeDIP-seq with low DNA concentrations. *Nat Protoc* **7**(4): 617-636.
Takazawa T, Croft GF, Amoroso MW, Studer L, Wichterle H, Macdermott AB. 2012. Maturation of spinal motor neurons derived from human embryonic stem cells. *PLoS One* **7**(7): e40154.
Takizawa T, Nakashima K, Namihira M, Ochiai W, Uemura A, Yanagisawa M, Fujita N, Nakao M, Taga T. 2001. DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain. *Dev Cell* **1**(6): 749-758.
Tan L, Xiong L, Xu W, Wu F, Huang N, Xu Y, Kong L, Zheng L, Schwartz L, Shi Y et al. 2013. Genome-wide comparison of DNA hydroxymethylation in mouse embryonic stem cells and neural progenitor cells by a new comparative hMeDIP-seq method. *Nucleic acids research* **41**(7): e84.
Tate PH, Bird AP. 1993. Effects of DNA methylation on DNA-binding proteins and gene expression. *Current opinion in genetics & development* **3**(2): 226-231.
Uda M, Ottolenghi C, Crisponi L, Garcia JE, Deiana M, Kimber W, Forabosco A, Cao A, Schlessinger D, Pilia G. 2004. Foxl2 disruption causes mouse ovarian failure by pervasive blockage of follicle development. *Hum Mol Genet* **13**(11): 1171-1181.
Uemura M, Tamura K, Chung S, Honma S, Okuyama A, Nakamura Y, Nakagawa H. 2008. Novel 5 alpha-steroid reductase (SRD5A3, type-3) is overexpressed in hormone-refractory prostate cancer. *Cancer science* **99**(1): 81-86.
Uhlenhaut NH, Jakob S, Anlag K, Eisenberger T, Sekido R, Kress J, Treier AC, Klugmann C, Klasen C, Holter NI et al. 2009. Somatic sex reprogramming of adult ovaries to testes by FOXL2 ablation. *Cell* **139**(6): 1130-1142.

Val P, Jeays-Ward K, Swain A. 2006. Identification of a novel population of adrenal-like cells in the mammalian testis. *Dev Biol* **299**(1): 250-256.
Vanyushin BF, Tkacheva SG, Belozersky AN. 1970. Rare bases in animal DNA. *Nature* **225**(5236): 948-949.
Varley KE, Gertz J, Bowling KM, Parker SL, Reddy TE, Pauli-Behn F, Cross MK, Williams BA, Stamatoyannopoulos JA, Crawford GE et al. 2013. Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Res* **23**(3): 555-567.
Venolia L, Gartler SM, Wassman ER, Yen P, Mohandas T, Shapiro LJ. 1982. Transformation with DNA from 5-azacytidine-reactivated X chromosomes. *Proc Natl Acad Sci U S A* **79**(7): 2352-2354.
Vertino PM, Yen RW, Gao J, Baylin SB. 1996. De novo methylation of CpG island sequences in human fibroblasts overexpressing DNA (cytosine-5-)-methyltransferase. *Molecular and cellular biology* **16**(8): 4555-4565.
Waalwijk C, Flavell RA. 1978. DNA methylation at a CCGG sequence in the large intron of the rabbit beta-globin gene: tissue-specific variations. *Nucleic acids research* **5**(12): 4631-4634.
Waddington C. 1940. *Organisers and Genes*. Cambridge University Press, Cambridge.
Waddington CH. 1939. *An Introduction to Modern Genetics*. George Allen & Unwin Ltd, London.
Wade PA. 2001. Methyl CpG-binding proteins and transcriptional repression. *BioEssays : news and reviews in molecular, cellular and developmental biology* **23**(12): 1131-1137.
Wagner S, Kiupel M, Peterson RA, Heikinheimo M, Wilson DB. 2008. Cytochrome b5 expression in gonadectomy-induced adrenocortical neoplasms of the domestic ferret (Mustela putorius furo). *VetPathol* **45**(4): 439-442.
Walsh CP, Chaillet JR, Bestor TH. 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nat Genet* **20**(2): 116-117.
Warnecke PM, Stirzaker C, Melki JR, Millar DS, Paul CL, Clark SJ. 1997. Detection and measurement of PCR bias in quantitative methylation analysis of bisulphite-treated DNA. *Nucleic acids research* **25**(21): 4422-4426.
Watt F, Molloy PL. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* **2**(9): 1136-1143.
Weber MM, Fottner C, Schmidt P, Brodowski KM, Gittner K, Lahm H, Engelhardt D, Wolf E. 1999. Postnatal overexpression of insulin-like growth factor II in transgenic mice is associated with adrenocortical hyperplasia and enhanced steroidogenesis. *Endocrinology* **140**(4): 1537-1543.
Weinberg RA. 2007. *The Biology of Cancer*. Garland Science, New York.
Wheeler HL, Johnson TB. 1904. Researches on pyrimidine derivatives: 5-Methylcytosine. *American Chemical Journal* **31**: 1.
Wichterle H, Lieberam I, Porter JA, Jessell TM. 2002. Directed differentiation of embryonic stem cells into motor neurons. *Cell* **110**(3): 385-397.
Wichterle H, Peljto M. 2008. Differentiation of mouse embryonic stem cells to spinal motor neurons. *Current protocols in stem cell biology* **Chapter 1**: Unit 1H 1 1-1H 1 9.
Willems A, De Gendt K, Allemeersch J, Smith LB, Welsh M, Swinnen JV, Verhoeven G. 2010. Early effects of Sertoli cell-selective androgen receptor ablation on testicular gene expression. *Int J Androl* **33**(3): 507-517.
Wilson AS, Power BE, Molloy PL. 2007. DNA hypomethylation and human diseases. *Biochim Biophys Acta* **1775**(1): 138-162.
Wood MA, Acharya A, Finco I, Swonger JM, Elston MJ, Tallquist MD, Hammer GD. 2013. Fetal adrenal capsular cells serve as progenitor cells for steroidogenic and stromal adrenocortical cell lineages in M. musculus. *Development* **140**(22): 4522-4532.
Wood MA, Hammer GD. 2011. Adrenocortical stem and progenitor cells: unifying model of two proposed origins. *Mol Cell Endocrinol* **336**(1-2): 206-212.
Wyatt GR. 1950. Occurrence of 5-methylcytosine in nucleic acids. *Nature* **166**(4214): 237-238.
Xi Y, Bock C, Muller F, Sun D, Meissner A, Li W. 2012. RRBSMAP: a fast, accurate and user-friendly alignment tool for reduced representation bisulfite sequencing. *Bioinformatics* **28**(3): 430-432.
Xie W, Barr CL, Kim A, Yue F, Lee AY, Eubanks J, Dempster EL, Ren B. 2012. Base-resolution analyses of sequence and parent-of-origin dependent DNA methylation in the mouse genome. *Cell* **148**(4): 816-831.

Xiong Z, Laird PW. 1997. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic acids research* **25**(12): 2532-2534.
Yang WH, Gutierrez NM, Wang L, Ellsworth BS, Wang CM. 2010. Synergistic activation of the Mc2r promoter by FOXL2 and NR5A1 in mice. *Biol Reprod* **83**(5): 842-851.
Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. 2007. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* **39**(1): 61-69.
Ziller MJ, Gu H, Muller F, Donaghey J, Tsai LT, Kohlbacher O, De Jager PL, Rosen ED, Bennett DA, Bernstein BE et al. 2013. Charting a dynamic DNA methylation landscape of the human genome. *Nature* **500**(7463): 477-481.
Ziskind-Conhaim L. 1988. Electrical properties of motoneurons in the spinal cord of rat embryos. *Dev Biol* **128**(1): 21-29.
Zovkic IB, Guzman-Karlsson MC, Sweatt JD. 2013. Epigenetic regulation of memory formation and maintenance. *Learn Mem* **20**(2): 61-74.
Zubair M, Ishihara S, Oka S, Okumura K, Morohashi K. 2006. Two-step regulation of Ad4BP/SF-1 gene transcription during fetal adrenal development: initiation by a Hox-Pbx1-Prep1 complex and maintenance via autoregulation by Ad4BP/SF-1. *Mol Cell Biol* **26**(11): 4111-4121.
Zubair M, Parker KL, Morohashi K. 2008. Developmental links between the fetal and adult zones of the adrenal cortex revealed by lineage tracing. *Mol Cell Biol* **28**(23): 7030-7040.