Distribution agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____

Benjamin G. Barwick Date

Epigenetic regulation of B cell differentiation

Benjamin G. Barwick
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Genetics and Molecular Biology

Jeremy M. Boss, Ph.D
Advisor

Xiaodong Cheng, Ph.D.
Committee Member

Joshy Jacob, Ph.D.
Committee Member

Paula M. Vertino, Ph.D.
Committee Member

David J. Katz, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Epigenetic regulation of B cell differentiation

By

Benjamin G. Barwick

B.S., Georgia Institute of Technology, 2003

M.S., Georgia Institute of Technology 2006

Advisor: Jeremy M. Boss, Ph.D.

An Abstract of

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate School of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Graduate Division of Biological and Biomedical Sciences

Genetics and Molecular Biology

2016

Abstract

Appropriate regulation of B cell differentiation is essential to humoral immunity and controlled through transcription factors that dictate B cell fate and function. Transcription factor control of gene expression is restrained by the epigenetic accessibility of target loci, a process that is mediated, in part, by DNA methylation at CpG dinucleotides. Here, DNA methylation changes during plasma cell differentiation were studied in mice three days post challenge with LPS. As compared to naïve B cells, plasma cells were found to change DNA methylation state at 10% of CpG loci, 99.7% of which were demethylation events. Losses in DNA methylation tended to be distal to transcription start sites but overlapped enhancers of transcription. Chemical inhibition of DNA methylation augmented plasma cell commitment in a cell division-dependent manner, suggesting DNA methylation losses must occur prior to plasma cell differentiation. Analysis of *in vivo* differentiating B cells separated by the number of cellular divisions the B cell had undergone revealed that highly divided cells increased mRNA levels fivefold as compared to undivided B cells. This was coupled to a hierarchy of DNA hypomethylation events at specific transcription factor motifs in a cell division dependent manner. Conversely, accumulation of DNA methylation was less apparent and mice that lack Dnmt3a and Dnmt3b enzymes, which deposit *de novo* DNA methylation, had phenotypically normal B cell development and similar plasma cell responses to LPS over three days. However, upon extended immunologic challenge, Dnmt3ab-deficient mice accumulated more antigen-specific B cells and exhibited a fivefold increase in splenic and bone marrow plasma cells. Molecular analysis revealed that Dnmt3ab-deficient bone marrow plasma cells failed to repress gene expression at several hundred genes to the same levels as Dnmt3ab-sufficient mice. This was coupled with a failure of Dnmt3ab-deficient germinal center B cells and plasma cells to gain and/or maintain DNA methylation at several thousand loci that were clustered in enhancers of genes that function in B cell activation and homing. These data provide mechanistic insight into cell-division coupled transcriptional and epigenetic reprogramming and suggest that while DNA hypomethylation events reflect the recent cis-regulatory history of plasma cells, extended responses are controlled by *de novo* DNA methylation that function to help maintain plasma cell homeostasis.

Epigenetic regulation of B cell differentiation

By

Benjamin G. Barwick

B.S., Georgia Institute of Technology, 2003

M.S., Georgia Institute of Technology 2006

Advisor: Jeremy M. Boss, Ph.D.

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate School of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Graduate Division of Biological and Biomedical Sciences

Genetics and Molecular Biology

2016

Acknowledgements

The work contained herein would not have been possible without the guidance and support of my advisor Dr. Jeremy M. Boss and numerous lab members including Dr. Christopher D. Scharer, and Alexander Bally. I would like to thank my committee members Drs. Xiaodong Cheng, Joshy Jacob, David J. Katz, and Paula M. Vertino for their time, guidance, and world renowned expertise in epigenetics and B cell biology. I would like to thank Ryan J. Martinez for critical expertise and Drs. Jeremy M. Boss, Jake Kohlmeier, Brian Evavold and Emory University for the acquisition and maintenance of resources integral for these studies. I would like to thank all Boss laboratory members including Royce Butler, Robert H. Haines, Joshua Lee, Dr. Parimal Majumder, Dr. Qing Wang, Muyao Guo, Madeline Price, Anna K. Kania and Dillon Patterson, as well as former Boss laboratory members Drs. Nancy Choi, Peiyuan Lu, Wendy Zinzow-Kramer, Sarah Lohsen, James W. Austin, and Valerie Tang. I would like to thank my family for support including my wife Freedom B. David, my mother Elizabeth A. Cohen, and my step-father Richard Stone. I would also like to thank collaborators and colleagues that provided feedback in the publication of these studies including Drs. Paula M. Vertino, Lawrence H. Boise, Paul A. Wade, and Karen N. Conneely.

Table of Contents

Chapter 1 : General Introduction

A brief history of B cells and antibodies

Well before B cells or antibodies were discovered, observations from John Fewster and Edward Jenner realized individuals previously infected with smallpox were immune to subsequent infection [1]. Edward Jenner later discerned that milkmaids who acquired sores from cowpox – a related virus – were also immune. In 1796, Jenner described inoculating patients with the pus from the milkmaids hands and correlated this with protection from smallpox [1]. A century later, in 1890, Emil von Behring and Kitasato Shibasaburo showed that the serum of animals immunized with sub-lethal doses of dipetheria and tetanus contained an antitoxin [2]. This discovery proved the existence of an adaptive and humoral component in the immune system. The following year Paul Ehrlich described this antitoxin component as an "antibody" and in his 1908 Nobel laurate speech predicted the existence of cells that recognize these toxins using a "toxin receptor" and amazingly anticipated that "the antitoxin is nothing else but discharged components of the cell, namely receptors discharged in excess" [3]. Yet it would be more than 40 years before the cellular source of this immunity was discovered through contemporaneous observations.

In the late 1940s the formation of plasma cells was correlated with the production of antibody *in vitro* [4,5]. Yet the source of these plasma cells remained a mystery until several convergent findings aided in the discovery and delineation of lymphocytes. One such discovery was aided by the advent of antibiotics, which not only provided an enormous advancement in the treatment of bacterial infections, but also allowed for the

survival of genetically immune compromised patients that were unable to mount antibacterial responses. These cases – termed agammaglobulinaemia – were identified by Ogden Bruton [6] and subsequently others observed that these patients lacked plasma cells and germinal centers – a lymphoid structure in the white pulp of the spleen and lymph nodes [7]. Around the same time, removal of the chicken hindgut lymphoid organ – named the Bursa of Fabricius – resulted in animals that failed to mount antibody responses [8]. It is this organ from which B cells get their namesake. These observations were synthesized in seminal work by Max Cooper that showed a bifurcation of lymphoid cell types, where B cells controlled germinal center formation and antibody response, and thymic derived cells controlled cell mediated immunity [9]. These remarkable findings helped César Milstein and Georges Köhler fuse individual B cells with a transformed plasma cell – derived from patients with the plasma cell cancer multiple myeloma [10]. The fused "hybridomas" produced large amounts of clonal antibodies encoded by the B cell using the machinery from the myeloma cell line. This technique has revolutionized the production of biologic therapeutics with unprecedented specificity and aided in the successful treatment of autoimmunity [11,12] and cancer [13–15]. In the years since these discoveries, the genetic and cellular mechanisms that control antigen receptor rearrangement and diversity have been elucidated [16–18], revealing how T cells and B cells rearrange their receptors and recognize a diverse range of antigens, a concept that underlies the central dogma of adaptive immunity. Further, the molecular programming of B cells has been extensively studied, exposing an exquisite regulation of humoral immunity (reviewed in [19,20]). Despite this, many aspects of lymphocyte biology in health

and disease remain an enigma, justifying the continued study of the adaptive immune system.

B cells in health and disease

B cells provide humoral immunity by differentiating into plasma cells that secrete antibody which circulate through the lymphatics and into the mucosa. Antibodies can bind, neutralize, and thereby protect hosts from pathogens. This process is essential for protective immunity and vaccination. The correct balance of B cell quiescence vs. activation and differentiation provides the body with the right amount and repertoire of circulating antibody. However, this process is too often dysregulated. For instance, infant B cell responses are suppressed by maternal antibody [21] leaving them susceptible to disease with a reduced ability to develop immunological memory in response to vaccination. Women, are differentially affected by autoimmune diseases such as systemic lupus erythematosus (**SLE**), where pathogenic antibody repertoires become autoreactive [22]. Finally, B-cell receptor (**BCR**) rearrangement can result in genetic aberrations and are a leading cause of lymphoma (reviewed in [23]). Understanding both the normal and diseased B cell processes will provide insight into how to repress pathogenic B cell expansion and augment insufficient B cell responses.

Hematopoiesis and B cell development

B cells, like all immune cells, develop in primary lymphoid tissues during the processes of hematopoiesis. Hematopoietic stem cells (**HSC**) in the bone marrow [24,25] or fetal liver [26] differentiate into multi-potent progenitors (**MPP**) prior to commitment to the common lymphoid progenitor (**CLP**; **Figure 1-1**). CLP specification requires the transcription factors PU.1 [27] and E2A [28], and CLPs are denoted by expression of IL-7Rα [29] – required

for receipt of growth factor signals – and the recombination activated genes, RAG1 and RAG2 [30], which facilitate antigen receptor recombination [31,32]. Recently, CLPs have also been found in the gut lamina propria where it is believed B cell development also occurs [33].

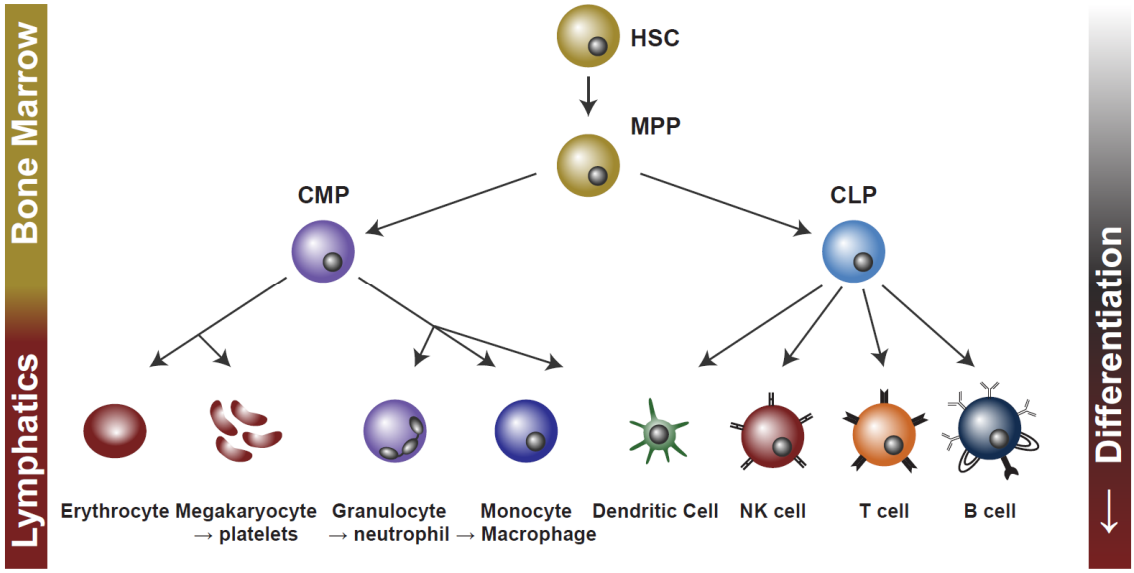
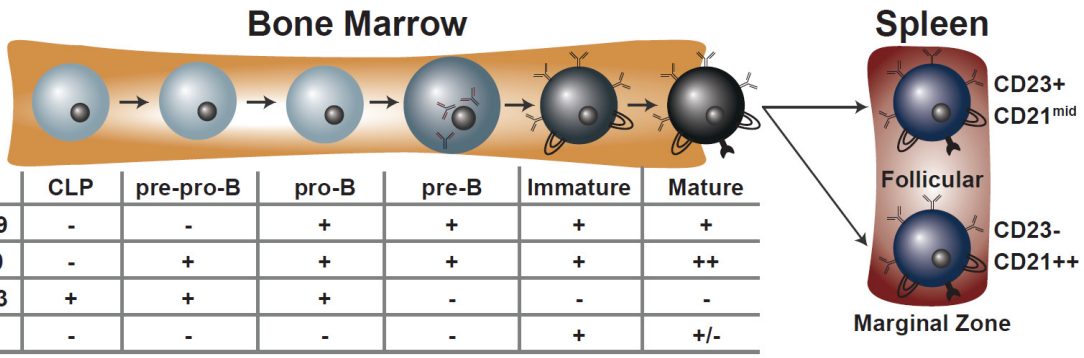


Figure 1-1: Hematopoiesis. Hematopoietic stem cells (**HSC**) give rise to multipotent progenitors (**MPP**) prior to commitment to either the common myeloid progenitor (**CMP**) or common lymphoid progenitor (**CLP**). The CLP can differentiate into dendritic cells, NK cells, T cells, but the most common fate are B cells.

Successive stages in B cell development require distinct signals that progressively rearrange the B cell receptor, thereby providing receptor diversity. The earliest B cell lineage specification is termed pre-pro-B (also known as fraction A) [34] and requires expression of Early B cell Factor 1 (**EBF1**) [35], which is dependent upon E2A, PU.1 and IL-7R (reviewed in [19]). It is during this stage when the RAG enzymes start to rearrange the immunoglobulin heavy chain (**IgH**). The IgH locus is stratified into Variable ($\mathbf{V_H}$), Diversity ($\mathbf{D_H}$), Joining ($\mathbf{J_H}$) and Constant regions (reviewed in [23]). The V_H, D_H, and J_H

regions are somatically rearranged during B cell development to create an intact heavy chain with the Constant μ region (**IgM**). During the pre-pro-B cell stage D_H-J_H recombination begins to occur [34]. Mechanistically, RAG enzymes recognize recombination signal sequences [36], and subsequently induce DNA breaks which can result in inversion or excision of a section of genomic DNA [31,32]. The end of the pre-pro-B cell stage is marked by completion of D_H-J_H recombination. At this stage EBF1 activates Pax5 – an essential gene for B cell development and function [37], and Pax5 in turn activates CD19 [38] which is a necessary for effective B cell responses [39] (**Figure 1-2**). During the pro-B stage V_H-DJ_H recombination occurs resulting in a fully rearranged IgM locus that begins to be transcribed (reviewed in [40]). Expressed IgM pairs with a surrogate light chain, marking the transition to the pre-B cell stage [34]. During the pre-B cell stage the light chain – which is composed of Variable ($\mathbf{V_L}$), Joining ($\mathbf{J_L}$), and Constant regions – begins to somatically rearrange [33]. Upon completion of VJ_L rearrangement, the light chain is expressed and pairs with the heavy chain to express surface IgM marking the beginning of the immature B cell stage [34]. B cells subsequently mature in the bone marrow and periphery to express IgM and the splice variant IgD. Mature B cells home to the spleen and circulate through the lymphatics. In an antigen naïve state, mature B cells rarely undergo mitosis [41] and generally express basal levels of transcripts [42]. This may reflect both an evolutionary origin of energy conservation and a mechanism to prevent autoimmunity.



	CLP	pre-pro-B	pro-B	pre-B	Immature	Mature
CD19	-	-	+	+	+	+
B220	-	+	+	+	+	++
CD43	+	+	+	-	-	-
IgM	-	-	-	-	+	+/-

Figure 1-2: B cell Development. B cells development progresses in sequential stages that are demarcated by expression of key signaling molecules and V(D)J recombination. As B cells mature they exit the bone marrow / fetal liver and enter the lymphatics where they can home to the spleen and lymph nodes and circulate throughout the body.

B cell subset function

Murine B cells are generally divided into three major subsets: Follicular (**Fo**), Marginal Zone (**MZ**), and B1 B cells (reviewed in [43]). MZ and B1 B cells compose 10-20% of the total B cell population and are anatomically located in the marginal zone of the white pulp in the spleen and the peritoneal cavity, respectively. These non-circulating B cell subsets are believed to be programmed to preferentially respond to circulating bacterial antigens such as lipopolysaccharides (**LPS**) in a T cell-independent manner [44]. MZ B cells have also been reported to facilitate T cell priming and T cell help through antigen shuttling mediated by CD21 [45]. Subsequently, MZ B cells can be distinguished from Fo B cells based on expression of CD21 and CD23 (**Figure 1-2**). Fo B cells account for 80-90% of mature B cells and circulate throughout the blood stream, lymphatics, spleen and even the bone marrow [46]. Due to their frequency and ability to circulate, Fo B cells are the primary subset believed to participate in germinal center formation, affinity

maturation, and longer-lived plasma cell responses. Thus, the stratification of B cell subtypes is a reflection of anatomical location and immunological function.

The cellular and molecular programming of B cell activation

B cell activation is generally characterized into T-cell independent (**TI**) and dependent (**TD**) pathways (**Figure 1-3**). TI B cell activation by definition does not receive direct T cell cytokine stimulation, an interaction mediated by presentation of protein antigen peptides on the major histocompatibility complex class II (**MHCII**) to T cells [47]. Thus, TI antigens are primarily non-protein antigens. TI antigens can be further bifurcated into type I and II. TI type I B cell responses occur through pattern recognition receptors (**PRRs**) that recognize pathogen epitopes. Examples of PRRs include the Toll-Like Receptors (**TLR**). TLRs were originally discovered to control fungal infections in flies [48] and later were found to have mammalian homologs [49–51]. TLRs expressed on B cells [52,53] include TLRs 1, 2, 4, 5, and 6 that recognize lipopeptides [51], peptidoglycans [54,55], lipopolysaccharide [49], flagellin [56], and lipoproteins, as well as TLRs 3, 7, and 9 that recognize double-stranded RNA [57], single-stranded RNA [58–60], and unmethylated CpG DNA [61,62], respectively (reviewed in [63]). When activated, TLRs provide mitogenic signals through Myeloid Differentiation Primary-Response Protein 88 (**MyD88**) [64] and in some cases TRIF [65] and TRAM [66]. These adaptor proteins potentiate transcription factors, including Nuclear Factor Kappa Light Chain Enhancer of Activated B cells (**NF-κB**), Activator Protein 1 (**AP-1**), and Interferon Regulatory Factors (**IRF**) affecting downstream gene expression programs including immunoglobulin synthesis [67]. This signaling cascade is extremely important and potent in mediating both innate and adaptive immunity. This is exemplified in B cells, where the MZ and B1 more innate-

like B cell subsets preferentially respond to TI type I antigens [44]; yet TD immune responses are still largely dependent upon TLR signaling for humoral immune responses [68,69].

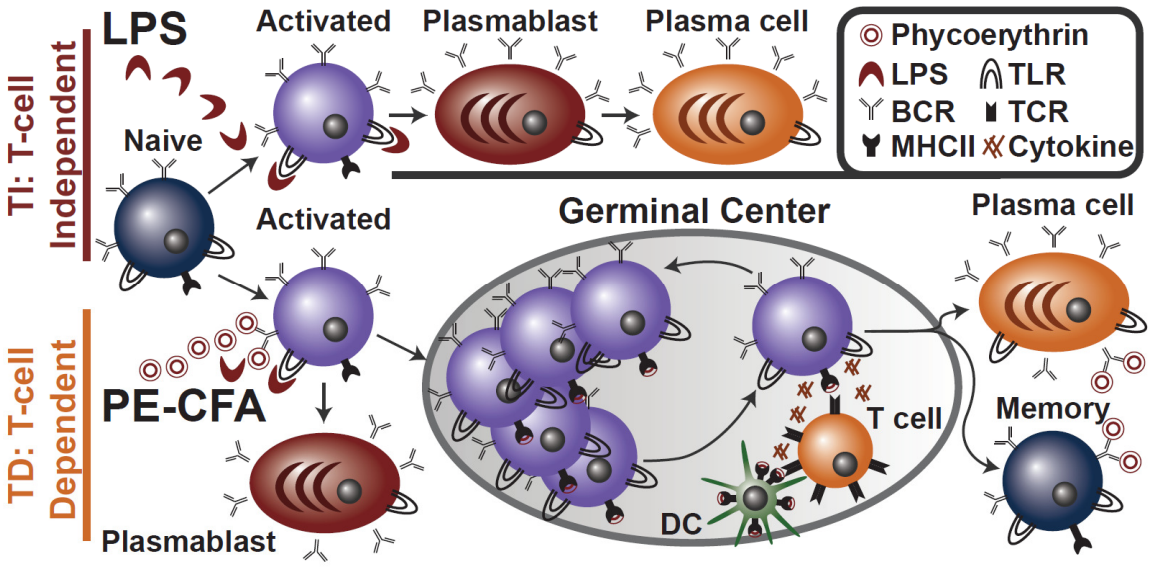


Figure 1-3: Pathways of B Cell Differentiation. Both T-cell independent (**TI**, top) and dependent (**TD**, bottom) mechanisms can results in B cell differentiation. TI pathways often include activation by pathogen epitopes through pattern recognition receptors such as LPS and TLR4, respectively. TD pathways involve protein antigens and cytokine help from T helper cells. See text for details.

TI type II antigens include polysaccharides and signal through the BCR to induce B cell activation and differentiation but do not receive direct T cell cytokine stimulation. Membrane bound BCRs self-cluster with the help of Igα and Igβ [70,71]. Upon BCR ligation by antigen, the Src-family protein tyrosine kinases FYN, BLK, and LYN [72] phosphorylate the immunoreceptor tyrosine-based activation motifs (**ITAMs**) on the cytoplasmic tails of Igα and Igβ [70,71,73]. Subsequently, SYC [74] and BTK [75,76] signal through BLNK [77–80] to phosphorylate PLCγ2 [78,81] and VAV [82,83] – a guanine nucleotide

exchange factor that signals through RAC1 – a Rho family guanosine triphosphatase (**GTPase**) [78]. CD19 functions as an adaptor transmitting VAV phosphorylation to the Phosphatidylinositol 3-Kinase (**PI3K**) p85 [84] that is required for B cell development and activation [85]. During this process PLCγ2 phosphorylates phosphatidylinositol-4,5-bisphosphate to make inositol-1,4,5-trisphosphate and diacylglycerol (**DAG**) which is required for calcium signaling and activation of Nuclear Factor of activated T cells (**NFAT**), AP-1 factors through Mitogen Activated Protein Kinases (**MAPK**), and NF-κB pathways [86,87]. Thus, although BCR and TLR signaling mechanisms are distinct, they result in activation of similar pathways. This may explain the strong influence of TLR agonists on TD B cell responses [68].

Traditionally, TD responses are believed to result in more robust humoral longevity whereas TI responses – both type I and II – are believed to result in shorter-lived B cell responses [88]. However, recent evidence has found that both B cell memory [89] and long-lived plasma cells [90] results from at least type II TI antigens. These observations are supported by the efficacy of Pneumovax, a vaccine based on immunity against pneumococcal polysaccharide [91].

TD antigens involve interaction and cytokine stimulation from CD4 T helper (**T_H**) cells. TD antigens crosslink the BCR which results in the downstream signaling pathways described above [92]. This also induces the chemokine receptor CCR7 [93], which in conjunction with expression of CXCR5 [94] and the down regulation of the Sphingosine-1-Phosphate Receptor (**S1PR**) [95,96], facilitates B cell migration to the T cell zone boundary where it can interact with T cells in either the spleen or lymph nodes. B cell and T cell interactions can result in T cell help in the form of cytokines. Mechanistically

this occurs when the B cell presents processed peptide antigen via the MHCII to a cognate TCR [92]. This interaction creates an "immunological synapse" between the T cell and B cell and initiates a signal cascade that results in a polar release of cytokines by the T cell. The MHCII-TCR interaction also requires B cell expression of CD40, which interacts with CD40 ligand (**CD40L**), expressed on T cells [92,97–99]. CD40-CD40L interactions provide growth and differentiation signals for both B cells and T cells and is necessary for TD B cell responses [97–99]. The MHCII-TCR and CD40-CD40L interactions initiate cytokine production by the T cell. The specific cytokine components secreted by the T cell depends on the T_H subtype and subsequently determines the type and magnitude of the B cell response (reviewed in [100,101]).

Naïve CD4 T cells upon activation produce Interleukin 2 (**IL2**) and differentiate into T_H cells that are classified into at least the T_{H1}, T_{H2}, T_{H17}, and T regulatory ($\mathbf{T_{reg}}$) subtypes [100,102]. The most common and first discovered T_H types are T_{H1} and T_{H2} cells [103]. T_{H1} cells were found to produce interferon gamma (**IFNγ**) and other macrophage stimulating cytokines that result in an inflammatory immune responses and activation of cellular immunity [103]. T_{H2} cells produce interleukin 4 (**IL4**) – which enhances the B cell mediated antibody response [103,104] – and interleukin 10 (**IL10**) – which has anti-inflammatory effects [105,106]. It has been subsequently shown that IL4 stimulates B cells to proliferate and differentiate [107–109] through the Signal Transducer and Activator of Transcription 6 (**STAT6**) [110], a signaling mechanism required for B cell responses.

T_{H17} cells are primed by IL23 [111,112], and mediate pro-inflammatory effects through the secretion of IL17 [113–115]. T_{H17} cells also provide effective B cell help [116]. Finally, T_{reg} cells suppress immune responses [117] and are controlled by the transcription

factor FoxP3 [118,119]. T_{reg} cells exert their affect by secretion of IL10 and other immunosuppressive cytokines, as well as by directly interacting with activated B cells [120]. Thus, the type of T cell help that occurs during a TD B cell response strongly influences the B cell reaction and immune resolution.

The interaction of an activated B cell with a T_H cell can result in rapid B cell proliferation in the splenic or lymph node follicles, which ultimately forms a circular structure called a Germinal Center [98] (**Figure 1-3**). Such T cells, regardless of subtype, are referred to as T follicular helper ($\mathbf{T_{FH}}$) cells. During the Germinal Center reaction B cells rapidly proliferate and cycle through the Germinal Center, which is composed of light and dark zones. In the light zone, B cells – referred to as centrocytes – interact with T cells and receive cytokines causing them to migrate to the dark zone and rapidly proliferate, at which point they are referred to as centroblasts (reviewed in [121]). It is at this stage that the antibody receptor is diversified through two distinct mechanisms. First, somatic hypermutation occurs in the V_H region of the Ig heavy chain [18]. Second, Class Switch Recombination (**CSR**) results in somatic deletion of parts of the constant region, resulting in new antibody isotypes. Both of these processes are dependent upon the gene Activation Induced Cytidine Deaminase (**AICDA**) [122].

Transcriptional regulation of B cell differentiation

Upon activation through either the BCR or TLR, B cells rapidly divide [123], increase global levels of gene expression [42], and gain the potential to differentiate into mitotically cycling plasmablasts, post-mitotic terminally differentiated plasma cells, or memory B cells [124]. Plasmablasts and plasma cells actively secrete Ig. Memory B cells do not secrete Ig but have the potential to more rapidly differentiate into antibody secreting cells

upon subsequent antigen exposure. Plasmablasts and plasma cells provide early and long-term humoral protection, respectively, and plasma cells have the potential to survive and secrete antigen-specific antibody for the lifetime of the host (**Figure 1-3**) [125,126]. Thus, B cell differentiation is a requisite step in the development of humoral immunity.

A dichotomy of mutually repressive transcription factors that either maintain the B cell fate or enforce the plasma cell gene expression program have been extensively studied [20,123]. Factors, such as Pax5 [37], Irf8 [127], PU.1 [128], Bach2 [129], and Bcl6 [130], simultaneously maintain the B cell gene expression program and repress plasma cell differentiation. As mentioned above, Pax5 is required for B cell development [37,131], but this factor must be silenced upon plasma cell differentiation [132,133]. This is presumably due to Pax5 activating B cell genes [38] and repressing the plasma cell genes including Xbp1 [134]. PU.1 contributes to B cell development [128,135] by forming a transcription factor complex with Irf8 [127] to establish the mature B cell gene expression program. When this complex is genetically ablated, B cells show greater potential to differentiate into plasma cells [127]. Bach2 directly represses IgH 3' enhancer activity [136]; and similarly, its deletion also increases plasma cell differentiation [129,137]. Thus, these key factors help establish and/or maintain the B cell gene expression program while simultaneously repressing plasma cell specific gene expression.

Conversely, several necessary transcription factors for plasma cell differentiation activate plasma cell genes while repressing the B cell program. One of the first discovered requisite transcription factors for plasma cell differentiation was Blimp-1 [138], which is encoded by *Prdm1*. Blimp-1 functions by repressing B cell target genes including Pax5 [139], the MHCII transactivator Ciita [140], and SpiB [141]. More recently, the

genomic targets of Blimp-1 have been identified through the creation of Biotin tagged Blimp-1 knock-in mouse [142,143], revealing that Blimp-1 directly regulates the secreted versus membrane bound forms of Ig. The transcription factor Inteferon Regulatory Factor 4 (**IRF4**) was discovered as a transcription factor important in plasma cell biology through translocations in multiple myeloma [144]. In depth study of Irf4 has shown that it is a potent regulator of plasma cell differentiation that directly binds to and induces transcription of *Prdm1*, and thereby induces Blimp-1 expression [145–147]. Other factors specifically requisite for plasma cell differentiation include Oct2 which is important for B cell activation and proliferation [148,149] and Xbp1, which functions downstream of Blimp-1 to help plasma cells expand their endoplasmic reticulum (**ER**) to process antibody [150,151]. Thus, plasma cell differentiation is accomplished through a regulatory cascade such that upon Irf4 activation, Blimp-1, and Xbp1 become upregulated enabling plasma cell biology. These functions include repression of the B cell program thereby helping to commit activated and mitotically dividing B cells, known as B lymphoblasts, to the plasma cell lineage. Furthermore, these transcription factors facilitate Ig secretion, and allow plasma cells to cope with ER stress. These and other studies have broadly characterized the transcription factor programming required for plasma cell differentiation.

Despite the extensive study of B cell and plasma cell transcriptional programming, the mechanisms that cause an activated B cell to commit to the plasma cell fate versus a memory B cell fate or apoptosis remain poorly understood. Work to date has demonstrated a clear dependence of differentiation on cellular division [123,127]. Yet, the number of divisions per se is not sufficient to predict plasma cell differentiation as not

all equally divided B cells experience the same fate [152]. This has led to a stochastic model of differentiation that is highly variable for individual B cells, but leads to balanced progeny fates at a population level [124,152,153]. One mechanism that could contribute to such cellular variability is epigenetic heterogeneity. Epigenetic modifications can enhance or repress gene transcription and are mitotically heritable [154,155]. Thus a more complete understanding of the epigenetic changes that occur during B cell differentiation may provide insight into fate choices in humoral immunity.

The conception of epigenetics

The term epigenotype was coined by Conrad Waddington in 1942 to describe "the mechanism by which the genes of the genotype bring about phenotypic effects" [156]. Years later Conrad Waddington proved that environmental conditions produced sustained phenotypic effects independent of genotype using heat shock stimulus in *Drosophila melanogaster* [157]. In the years since, epigenetics has become known collectively as the sum of chemical modifications that occur directly on DNA – such as 5-methylcytosine (**5mC**) – or on the nucleosomes that package DNA – such as histone acetylation – as well as other molecules that affect the three dimensional conformation of DNA and subsequently the way encoded genes are expressed including non-coding RNAs such as microRNAs [158] and long non coding RNAs. Currently epigenetics is defined as a trait that is a "stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence" [159].

Concurrent with Waddington's observations, biochemical work was being conducted on nucleic acids that would ultimately lead to identification of what is today one of the best understood epigenetic mechanisms: 5mC, also commonly referred to as

DNA methylation. In 1960 Eduardo Scarano and others made observations of 5-methylcytidine in ribonucleic acid found in tissues from a wide range of eukaryotes [160,161]. These findings were validated and extended in 1962 when an RNA methyltransferase was identified [162], and again in 1965 and 1968 when 5mC was also found in DNA from developing sea urchins. This was primarily localized to the fraction of DNA containing CpG dinucleotides [163,164]. A decade later Holliday and Pugh proposed that DNA methylation was involved in developmental gene regulation and proposed that enzymes must exist to maintain DNA methylation through cell replication as well as to deposit it *de novo* in other locations [165].

Mechanisms of DNA methylation reprogramming

The first DNA methyltransferase (**Dnmt**) to be isolated became known as Dnmt1 and was found to be expressed in all of the mouse cell types tested at the time [166]. The human orthologue was cloned a few years later and displayed 80% homology [167]. It would later be discovered that this enzyme showed a preference for hemi-methylated DNA [168] and localized to replication foci [169], leading to the deduction that this gene encoded the maintenance DNA methyltransferase that would reciprocally methylate hemi-methylated CpGs that result from a newly synthesized and unmethylated daughter strand pairing with a methylated parent strand. Later, genetic ablation of Dnmt1 showed that this enzyme is required for embryonic viability [170]. Another class of Dnmts was discovered through homology searches of Dnmt1 in expressed sequence tag databases [171]. These genes were deemed Dnmt3a and Dnmt3b. A year later, genetic deletion of these enzymes in mice showed that they mediated *de novo* DNA methylation but do not impact the maintenance

of DNA methylation patterns through cell division, and that both are required for mammalian development [172].

In the years since the discovery and characterization of the DNA methyltransferases, the genomic revolution has provided us with a map of the human [173] and mouse [174] genomes. In total, allelic polymorphisms notwithstanding, there are 28,217,448 CpGs in the human genome and 21,722,958 in the mouse. This is far less than the expected dinucleotide frequency given the size of the genomes and average GC content. Further analyses of the genomic sequences have confirmed the asymmetrical distribution of CpG dinucleotides, an observation that was made more than 30 years ago based on restriction enzymes [175,176]. Most CpG dinucleotides fall in dense regions or 'islands' that are in contrast to the rest of mammalian genomes, which are depleted of CpG dinucleotides – a phenomenon that likely occurs because unmethylated cytosines are prone to deaminate into uracils which are then subject to base-excision repair mechanisms that can result in the incorrect repair to an AT base pair. CpGs in CpG islands have been conserved over evolutionary time, presumably due to their role in gene regulation. Consistent with this, more than half of all CpG islands in the mammalian genome overlap promoters and this is particularly pronounced for housekeeping and developmental genes [177]. In this context, CpG islands are generally constitutively unmethylated in contrast to the rest of the genome which is mostly methylated dependent upon the cell type.

The sequencing of mammalian genomes has also revealed that more than 44% of the genome is composed of repetitive elements [173]. These elements include Long Interspersed Nuclear Elements (**LINES**), Short Interspersed Nuclear Elements (**SINES**),

and Long Terminal Repeats (**LTRs**). Examples of such repetitive elements include retrotransposons, which when activated can rapidly mutate the genome. These elements are heavily methylated and repressed, thus preserving the integrity of the genome. This is one of the most important functions of DNA methylation [177].

The characterization of the DNA methylome has also been aided by advances in genomics. The combination of sodium bisulfite chemical treatment – which deaminates unmethylated cytosine to uracil that is subsequently PCR amplified as thymine – with high throughput sequencing – which facilitates the analysis of hundreds of millions of short reads in a rapid amount of time [178], has enabled rapid profiling of the DNA methylome. Data generated from such techniques can be informatically mapped to the *in silco* bisulfite converted genome to decipher the position and fraction of DNA methylation at base resolution [179]. Still, whole genome bisulfite sequencing (**WGBS**) is expensive. Techniques that reduce the representation of genome by using restriction enzymes to digest the genome prior to sequencing adapter ligation have enabled cost effective genome-wide DNA methylation profiling of approximately 10% of the DNA methylome at a fraction of the cost and are referred to as reduced representation bisulfite sequencing (**RRBS**) [180,181]. These next generation DNA methylation analyses have confirmed observations made decades before [175] that CpG islands are primarily constitutively unmethylated and that cytosine methylation is overwhelmingly found in the CpG context [181]. Notable exceptions can occur during development, where human embryonic stem cells show significant amounts of CHH and CHG methylation (H is any base except for G) [182]. This is often attributed to very high levels of Dnmt3a or Dnmt3b expression [177,182] as compared to cell types that lack CHH and CHG methylation. Finally,

the genomic targets and functions of Dnmt3a and Dnmt3b have been aided through methylation profiling in a large number of cell types and through the creation cell lines with biotin tagged genes that can be more precisely mapped to the genome. These studies have revealed that Dnmt3a and Dnmt3b have both overlapping and distinct functions with respect to the regions that they methylate and the phenotypic consequences of their deletion [183–185]. Dnmt3b was found to localize in gene bodies and contribute to the high levels of gene body DNA methylation found in highly transcribed genes [185]. These studies and others have broadly characterized how DNA methylation is gained and maintained across the genome, but what has been less clear until recently was how DNA methylation was lost.

DNA demethylation can occur by cell replication coupled or 'passive' loss of DNA methylation (*i.e.* failure to maintain DNA methylation), as well as 'active' removal of cytosines containing DNA methylation. Active loss of DNA methylation was suspected to occur long before the mechanisms were realized, as the paternal genome undergoes a genome-wide loss of DNA methylation after fertilizing an egg and prior to zygote formation and cellular division [186,187]. Despite intensive searches for the molecular mechanisms, early reports were plagued by irreproducibility [188]. In 2009, this changed when it was shown that the Ten-Eleven Translocation (**TET**) family of 2-oxoglutarate, and Fe(II)-dependent enzymes could oxidize 5mC into 5-hydroxymethylcytosine (**5hmC**) [189]. Shortly after, it was realized that TET enzymes can further oxidize 5hmC to 5-formylcytosine (**5fC**) and 5-carboxylcytosine (**5caC**) [190,191], and these bases can be removed through the base excision repair pathway by the enzyme Thymine DNA Glycosylase (**TDG**) [192].

Although it is accepted that Tet-mediated oxidation of 5mC and subsequent removal through TDG can result in active demethylation, it is less clear what are the individual contributions of the active versus replication-coupled passive pathways of DNA demethylation in any given biological process. For instance, it has been shown that overexpression of factors such as the CCCTC-binding factor (**CTCF**) can enhance the demethylation of cognate DNA binding sites while neighboring sites are unchanged [193,194], but it is unclear if this DNA demethylation is mediated by i) exclusion of the maintenance methyltransferase at these sites through replication *i.e.* failure to maintain, ii) through Tet-mediated oxidation and active removal of 5hmC, 5fC, or 5caC by TDG, iii) Tet-mediated oxidation of 5mC and subsequent inability to maintain oxidized forms of 5mC through replication, or iv) some combination of the aforementioned. Supporting the notion that Tet-mediated oxidation results in DNA demethylation through both active and passive mechanisms, the maintenance methyltransferase seems to specifically recognize hemi-methylated 5mC and not hemi-methylated 5hmC [195]. Thus, although much has been learned about DNA methylation dynamics over the past half century resolving the kinetics and mechanisms of DNA methylation changes is still a big challenge.

Histone modifications and control of gene expression

As noted above, DNA methylation is just one mode of epigenetic regulation used to program cell identity. DNA is wrapped around nucleosomes composed of histones that typically include two copies of H2A, H2B, H3 and H4 packaged as an octamer. Epigenetic modifications to these core histones are important for gene regulation. These include, but are not limited to, histone acetylation, methylation, phosphorylation, and

ubiquitination [196]. Histone acetylation is generally associated with open chromatin that is permissive to transcription; an observation that was made over 50 years ago [197]. Acetylation reduces or removes the positive charge on histones resulting in negatively charged DNA that is wound less tightly around histones, thus facilitating 'open' chromatin [198]. Methylation modifications are charge neutral and mediate repressive or permissive gene expression effects by modifying the epitopes with which chromatin remodeling factors and transcription factor complexes interact.

Some of the well-studied histone modifications include histone H3 lysine 4 mono- (**H3K4me1**) and di-methylation (**H3K4me2**) found at enhancers [199,200], and H3 lysine 4 tri-methylation (**H3K4me3**) at the promoters of active genes [201]. Conversely, histone H3 lysine 9 methylation (**H3K9me**) induces heterochromatin and gene silencing [202]. Histone H3 lysine 27 acetylation (**H3K27ac**) demarcates active from poised enhancers and is also found at promoters [200]. H3K27ac antagonizes the Polycomb repressive mark H3 lysine 27 tri-methylation (**H3K27me3**) [203], which is also found at highly conserved bivalent promoters that contain both active (H3K4me3) and repressive (H3K27me3) histone modifications, and are expressed at low levels but are "poised" for gene activation [204]. Histone lysine 36 methylation (**H3K36me**) is found in the gene body of highly expressed genes, a phenomenon mediated by the interaction of the H3K36 methyltransferase Set2 interacting with the carboxyl-terminal tail of RNA polymerase II [205]. Finally, Histone H4 lysine 20 tri-methylation (**H4K20me3**) occurs at transcriptionally incompetent regions [206], a phenomenon at least partially explained by the pausing of RNA polymerase II at such regions [207].

The "Epigenetic Code"

Even prior to completion of the human and mouse genome sequencing projects, the combinatorial modifications to histones were envisioned to compose a "histone code" that regulated all chromatin templated processes [208,209]. Mechanistic and crystallographic structural studies have revealed how specific histone modifications interact with specific DNA methylation states (reviewed in [210]) (**Fig 1-4**). Some notable discoveries include the understanding that Dnmt3l – which recruits the *de novo* DNA methyltransferases [211,212] – specifically interacts only with nucleosomes that are unmethylated H3K4 (**H3K4me0**) [213,214], thus explaining the lack of DNA methylation at H3K4me marked promoters. As noted above H3K36me3 is found in the gene body of highly transcribed genes due to the interaction of RNA polymerase II and the H3K36 methyltransferase, Set2 [205]. These regions also contain high levels of DNA methylation, which is attributed to the interaction of the PWWP domain of Dnmt3b with H3K36me3 [185]. Finally, H3K9me is associated with heterochromatin and high levels of DNA methylation, which may be explained by interaction between the H3K9 methyltransferase, G9a, and the *de novo* DNA methyltransferase Dnmt3a [215].

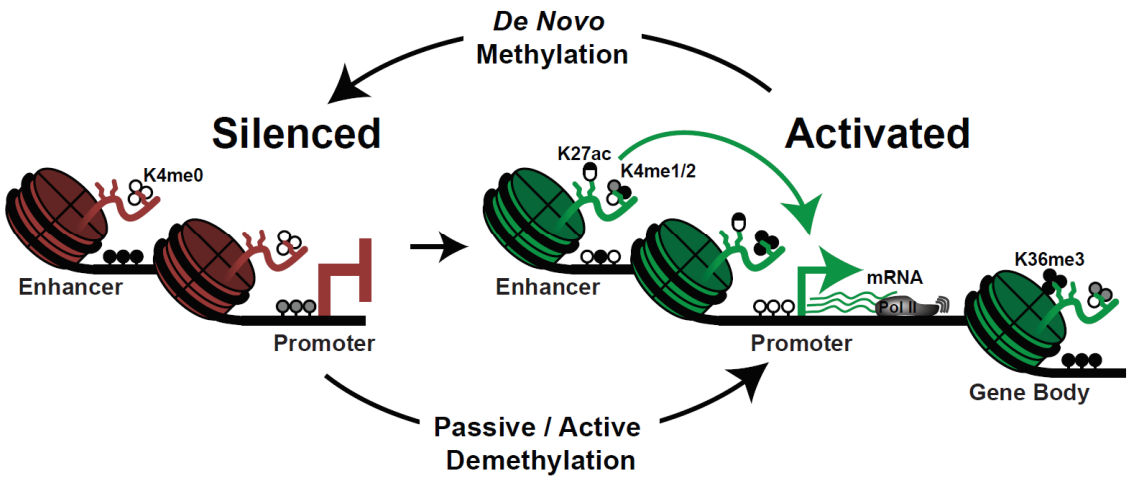


Figure 1-4: Epigenetic Mechanisms Influencing Gene Expression. Gene silencing is marked by a lack of H3K4me at promoters which may or may not contain DNA methylation depending upon the promoter CpG content (left). Active enhancers contain H3K4me1/2 and H2K27ac and transcribed promoters have H3K4me3 and lack DNA methylation (right). Gene bodies of highly transcribed genes typically have high levels of H3K36me3, DNA methylation, and may contain intragenic enhancers marked by H3K4me1/2 (far right). CpGs potentially containing DNA methylation are denoted by lollipops with black representing methylated CpGs, white denoting unmethylated regions and gray indicating regions that may or may not be methylated.

The epigenetic code has been further realized with the advent of high throughput sequencing in combination with enrichment techniques such as chromatin immunoprecipitation (**ChIP-seq**), DNase I hypersensitivity (**DNase-seq**), and the Assay for Transposase Accessible Chromatin (**ATAC-seq**). Large international consortiums have systematically mapped epigenetic data in a wide range of cell types revealing a cartography of epigenetic regulation in normal development and disease [216–233]. These data have revealed hundreds of thousands of enhancers in over a hundred cell types [231] and have shown that intermediate levels of DNA methylation are associated with these distal regulatory elements [193,229]. Further, these data are beginning to show how

environmental stimulus and genetic information influence combinatorial epigenetic states that control gene expression [230].

Epigenetic programming in hematopoiesis

Like most differentiation processes, hematopoiesis involves epigenetic reprogramming, which has been intensively studied over the past few decades. Early data found high expression of the DNA methyltransferases in hematopoietic stem cells (**HSCs**), which were progressively downregulated in differentiated cell types, especially those of the myeloid lineage [234,235]. Comparison of hematopoietic cells from healthy individuals to those from patients with Acute Myeloid Leukemia (**AML**) or Myelodysplastic Syndrome (**MDS**) correlated overexpression of the DNA methyltransferases and hypermethylation of the cell cycle regulator $p15^{INK4B}$ with disease [234,236]. Genetic disruption of *Dnmt1* in HSCs yielded discrepant results with one group reporting severe cytopenia and cell-autonomous mortality within 2 weeks [237] while the other reported normal peripheral blood and survival for at least 12 weeks [238]. Both studies used a common genetic model [239,240] and reported similar efficiencies of deletion, but other experimental details may have been inconsistently described [238]. Nonetheless, both studies found ablation of differentiated lymphoid and myeloid cell types and reported that *Dnmt1* deficient cells were outcompeted by control cells in the bone marrow. Additional analysis of a hypomorphic mutation in *Dnmt1* allowed a more detailed analysis showing HSC self-renewal was impaired and differentiation was skewed towards myeloerythroid fates without proper Dnmt1 expression [237]. This was later corroborated by treatment of HSCs with the DNA methylation inhibitor 5-aza-2'-deoxycytidine, which increased myeloid progeny at the expense of lymphoid differentiation, and also by genome-wide analysis of

myeloid and lymphoid development showing a genome-wide loss of DNA methylation in the myeloid versus lymphoid lineage [241].

Deletions of the *de novo* DNA methyltransferases were found to have differing effects on HSC self-renewal and differentiation. The first report deleted *Dnmt3a* and/or *Dnmt3b* in HSCs using a retroviral expressed *Cre* recombinase *ex vivo* and found HSCs were still capable of *in vitro* differentiation and single mutants were capable of HSC self-renewal in transplantation assays, whereas double mutants where not [242]. Deletion of *Dnmt3a* and *Dnmt3b* using Mx-*Cre* also found that double mutants had impaired differentiation after serial transplantations, but unlike the first report, this analysis found that double mutant cells had expanded HSC self-renewal [184]. This discrepancy may be partially explained by timing of deletion where one group deleted cells *ex vivo* prior to transplantation, while the other induced deletion 4 weeks post transfer. Deletion of *Dnmt3a* alone also increased HSC self-renewal and differentiated cells failed to repress the HSC gene expression program [243], were prone to aberrant DNA methylation [244], and hematological malignancies [245]. These phenotypic effects of *Dnmt3a* deletion are consistent with *Dnmt3a* mutations found in AML [246,247] and T cell lymphoma [248]. Cumulatively, these data indicate that *de novo* DNA methylation facilitates hematopoietic differentiation and without it, HSC self-renewal is perturbed.

Recently mutations and/or epigenetic silencing of *Tet1* and *Tet2* have been reported in a number of hematological malignancies including AML [249–252], MDS [250,253], and lymphomas [254,255]. Subsequent genetic analysis has shown that deletion of *Tet2* results in increased HSC self-renewal and myeloproliferation [254,256,257]. These studies have also found the inhibition of the Tet enzymes can result in increased differentiation

of both lymphoid [257,254] and myeloid [256,258] lineages, suggesting a general hyperproliferative effect upon Tet inhibition. Despite these intensive studies it is still unclear exactly how Tet deficiency induces oncogenesis.

Epigenetic programming in B cell development and differentiation

Commitment to the B cell lineage is associated with epigenetic modifications that position lineage factors for expression. Prior to B cell commitment, regulatory elements near lineage specific genes are decorated with the enhancer mark H3K4me1/2 in both macrophages and B cells. This process was found to be both dependent upon PU.1 expression at the multipotent progenitor (**MPP**) stage and coincident with PU.1 binding [259], suggesting that PU.1 may be recruiting MLL proteins that catalyze H3K4me1. Upon lymphoid restriction, which requires PU.1 activation of IL7R [19], E2A becomes expressed [28]. The number of E2A bound sites substantially increases during the transition from pre-pro-B to pro-B stage, which is coincident with these sites gaining H3K4me1/2 [260]. These sites included putatively regulatory elements around the essential B cell factors *Cd19* and *Cd79a* [260]. E2A activates EBF1 [261], which helps further poise the chromatin structure of key B cell genes including *Cd79b* and *Pax5* through chromatin remodeling complexes [260,262,263]. Thus, one function of many B cell lineage factors maybe to help specify cell fate by recruiting complexes that epigenetically prime regulatory elements near requisite B cell genes.

In concert with epigenetic and transcriptional specification of B cell fate, the immunoglobulin heavy and light chains are somatically rearranged – a process also dictated by epigenetic accessibility. This was first evidenced through the transfection of an accessible T cell receptor variable region that underwent rearrangement in a B cell line

whereas the endogenous T cell receptor did not and was not accessible [264]. This was supported by studies of the Igκ light chain that found the active allele lost DNA methylation dependent upon an intronic enhancer bound by the Nuclear Factor Kappa-light-chain-enhancer of activated B cells (**NF-κB**) [265,266]. Genetic deletion of the *de novo* DNA methyltransferases found enhanced Igκ rearrangement suggesting DNA methylation played a mechanistic role in antigen receptor rearrangement [267]. The active Igκ allele is also demarcated with H3K4me3 [268], which is now known to be mechanistically required for recombination due to the RAG enzymes containing a plant homeodomain that specifically recognizes H3K4me2/3 [269,270]. Thus antigen receptor diversity and expression are preceded by epigenetic reprogramming of cis-regulatory elements that control humoral immune responses.

Despite the intensive study of epigenetics in hematopoiesis and B cell development, only recently have studies started to address the role of epigenetic programming in B cell differentiation. Many of these studies have been motivated by the occurrence of B cell lymphomas with mutations in epigenetic enzymes. For instance, mutations that change the specificity of Enhancer of Zest 2 (**Ezh2**) are common in lymphomas of germinal center origin [271] and subsequent functional studies have shown that Ezh2 is required for germinal center formation and altered function contributes to lymphomagenesis [272]. The importance of Ezh2 in germinal centers has been attributed to its ability to repress cell cycle proliferation and plasma cell differentiation through direct repression of Cdkn1a and Blimp-1, respectively [272]. The promoters of these genes were shown to contain both H3K4me3 and H3K27me3 modifications, a chromatin state referred to as bivalent, and associated with genes that can be rapidly induced.

Identification of Blimp-1 binding targets also revealed that a subset of repressed Blimp-1 targets gained H3K27me3 and that Blimp-1 co-immunoprecipitated with Ezh2, suggesting a functional role in plasma cells [142].

Seminal work found aberrant DNA methylation in hematological malignancies and subsequently DNA methylation has been studied in both healthy and diseased B cell differentiation settings. Comparison of DNA methylation in naïve and germinal center B cells indicated dramatic rearrangement of the DNA methylome in germinal center B cells where most differential regions were hypomethylated [273]. Germinal center B cells were also found to be reliant upon *Dnmt1* expression, not a surprising result given the rapid proliferation that germinal center B cells undergo [273]. Germinal center demethylation has been attributed to the cytidine deaminase AID as germinal centers in AID-deficient mice underwent reduced DNA methylation changes [274]. However, it is unclear if this is a direct effect of AID removing DNA methylation or an artefact of perturbed B cell activation in AID deficient mice [275,276]. Lending doubt to the hypothesis that AID is a driving factor in B cell DNA methylation changes, other groups have reported no DNA methylation differences in similar experiments [277] and the authors of the original study reported a large decrease in both hypo- and hyper-methylated CpGs in AID-deficient mice suggesting the observed AID-dependent differences may be technical in nature.

Further analysis of human B cell subsets corroborated a loss of DNA methylation upon B cell activation. Analysis of B cell subsets from human tonsils showed that the DNA methylome of germinal center B cells, memory B cells, and plasma cells were dramatically and similarly rearranged when compared to naïve B cells, despite unique transcriptional states in each cell type [278]. Many of the regions that lost DNA methylation

contained transcription factor binding motifs suggesting potential cis-regulatory activity at these regions. The DNA methylation states on a range of developing and differentiating B cells were definitively shown using WGBS. These studies revealed high levels of DNA methylation in developing and naïve B cells that were reduced in germinal center B cells, memory B cells, and plasma cells [279]. The authors reported 30% of DNA methylome was remodeled during this process and demethylation corresponded with enhancers whereas regions that gained DNA methylation correlated with polycomb marked regions [279]. Thus, while the DNA methylation states of different developing and differentiating B cell subsets are now known, the timing and function of these epigenetic changes in relation to B cell differentiation are still poorly understood.

To better understand the activation-induced DNA methylation changes that occur during B cell differentiation and it's relation to gene expression and other epigenetic marks, mice were challenged with T-cell independent (**TI**) and dependent (**TD**) model antigens and the immune responses were characterized by flow cytometry, gene expression, and DNA methylation analysis. The results show that there is robust genome-wide DNA demethylation at ~ 10% of CpG loci interrogated, that occurs within three days of a TI B cell activation by LPS which is dependent upon cell division and concurrent with a more than 5-fold transcriptional amplification. Conversely, only ~0.01% of CpGs assayed gained DNA methylation after three days of TI activation. During a 30 day TD challenge, similar results were found in that both germinal center B cells and bone marrow plasma cells showed pervasive DNA demethylation. However, approximately 0.18% of CpGs assayed gained DNA methylation. The majority of these were dependent upon the *de novo* DNA methyltransferases: *Dnmt3a* and *Dnmt3b*.

Furthermore, these mice had a marked expansion of germinal center B cells and splenic and bone marrow plasma cells, indicating these subtle gains in DNA methylation are important in limiting B cells responses. Cumulatively, these data provide a model whereby the DNA methylome of activated B cells is dramatically reduced during activation-induced division, but that over time distinct gains in DNA methylation help to control B cell homeostasis.

Chapter 2 : Plasma cell differentiation is coupled to division-dependent DNA hypomethylation and gene regulation

Abstract

The epigenetic processes that regulate the formation of antibody secreting plasma cells are not well understood. Here, analysis of in vivo plasma cell differentiation revealed a robust transcriptional induction and targeted DNA hypomethylation of 10% of CpG loci that were overrepresented at enhancers. Inhibition of DNA methylation enhanced plasma cell commitment in a cell division-dependent manner. Examination of in vivo differentiating B cells stratified by cell division revealed a 5-fold transcriptional amplification coupled to DNA hypomethylation. These division-dependent DNA demethylation events were ordered, occurring first at NF-κB and AP-1 sites and later at IRF and Oct-2 motifs that were coincident with activation and differentiation gene expression programs. Thus, DNA methylation changes reflected the cis-regulatory history of plasma cell differentiation. The data therefore provide mechanistic insight into the cell-division coupled transcriptional and epigenetic reprogramming events during plasma cell differentiation.

Introduction

B cell differentiation is required for humoral immunity. Resting naïve B cells rarely undergo mitosis [41], do not secrete immunoglobulins (Ig), and express only basal levels of transcripts [42]. Upon activation through the B-cell or Toll-like receptors, B cells rapidly divide [123] and differentiate into mitotically cycling plasmablasts, post-mitotic terminally differentiated plasma cells, or memory B cells [124,280]. Plasmablasts and plasma

cells actively secrete Ig whereas memory B cells do not but have the potential to more rapidly differentiate upon subsequent antigen exposure.

Despite the extensive study of B cell and plasma cell transcriptional programming [20,123], many mechanisms that govern differentiation remain unknown. Work to date has demonstrated that differentiation requires cell division [124,280]. Yet, the number of divisions does not solely determine plasma cell fate [124,152]. This has led to a stochastic model of differentiation that is highly variable for individual B cells but leads to balanced progeny fates at a population level [124,152,153]. One mechanism that could contribute to such cellular heterogeneity is epigenetic variability. Epigenetic modifications, such as DNA methylation or histone modification, can enhance or repress gene transcription and are mitotically heritable [154,155]. Epigenetic studies have found that DNA methylation is necessary for hematopoietic stem cell renewal, restricts myeloid differentiation, and allows for B cell commitment [237]. Comparatively few studies focused on understanding epigenetic programming during a B cell immune response [273,278,279]. In these studies, DNA methylation was remodeled in germinal center and memory B cells and plasma cells. Recently, analysis of the plasma cell transcription factor Blimp-1 binding targets has revealed epigenetic changes associated with Blimp-1 binding in plasmablasts [142]. However, the breadth, timing, and function of these epigenetic changes in response to an *in vivo* stimulus are incompletely understood.

To gain insight into the epigenetic mechanisms that govern B cell differentiation, *in vivo* models were used to determine direct relationships between DNA methylation, gene expression, and cell division. *In vivo* B cell differentiation was associated with targeted DNA hypomethylation and increased gene expression. Cell division was

accompanied by a hierarchy of DNA hypomethylation events at cis-regulatory elements that corresponded with division-specific expression. These results define a step-wise process of division-coupled epigenomic remodeling that allows B cells to adopt a new transcriptional program and cell fate.

LPS mediated B-cell differentiation is coupled to unique transcriptional states

An inducible *in vivo* model of B cell differentiation was used to examine the molecular events that could be traced to a defined stimulus. Mice challenged with the mitogen lipopolysaccharide (LPS) *i.v.* exhibited splenomegaly and a three-fold expansion of total B cells (B220+), with activated B cells (GL7+) increasing from 2% to 35% of splenocytes three days post-challenge (**Supplementary Fig. 2-1a-c**). Extrapolation of these data indicated that there were approximately 120 million new B cells in the splenic compartment, including more than 60 million newly activated B cells (**Supplementary Fig. 2-1d-f**). Analysis of CD138+ differentiating cells found an admixture of cells expressing mid to low levels of B220 (**Fig. 2-1a**). B220 expression on plasma cells (CD138+) has been identified as a marker of rapid cellular turnover in the spleen [281] and bone marrow [282], whereas $B220^{low}$ CD138+ cells are believed to compose a post-mitotic population [281]. Both $B220^{mid}$ and $B220^{low}$ CD138+ populations were strongly induced three days post-challenge with LPS, and are herein referred to as plasmablasts and plasma cells, respectively (**Fig. 2-1b-c**).

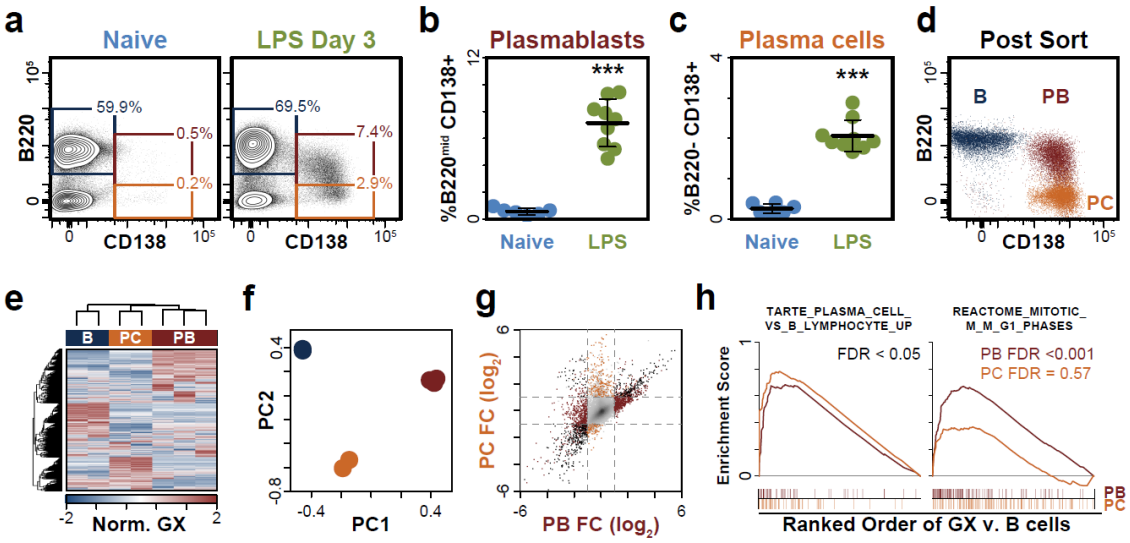


Figure 2-1. B-cell differentiation is coupled to unique transcriptional states. **a-c)** Flow cytometry analysis of B220 and CD138 expression on splenocytes from naïve and LPS-challenged mice on day 3. Cells represent size-gated and CD11b- CD11c- populations. **b)** Quantitation of B220^{mid} CD138+ (plasmablasts) and **c)** B220^{low} CD138+ (plasma cells). Mean and standard deviation are represented, ***$P \leq 0.001$, Welch's *t*-test. **d)** Post sort purity of B cells from naïve mice (B; blue) and plasmablasts (PB; burgundy) and plasma cells (PC; gold). All 3 samples are projected onto one plot in different colors. **e)** Expression of 16,181 genes in naïve B cells, plasmablasts and plasma cells represented as a heatmap with both samples and genes clustered. **f)** Principle components analysis of all expression data shown in **e)**. **g)** Scatterplot of gene expression fold-changes (FC) in plasmablasts (PB) and plasma cells (PC), each as compared to naïve B cells. Differentially expressed genes (FDR ≤ 0.01, ≥ 2 fold-change; **Supplementary Table 1**) are shown in burgundy (plasmablasts), gold (plasma cells), or black (both). Dashed gray lines indicate fold-changes of 2. **h)** Gene set enrichment analysis of plasmablast and plasma cell gene expression changes indicated an enrichment of human plasma cell genes published by Tarte *et al.* [283] (left) in both plasmablasts and plasma cells and an enrichment of the Reactome Pathway *Mitotic M-M/G1 phases* specifically in plasmablasts. Enrichment score is shown on top for both plasmablasts (burgundy) and plasma cells (gold). Below is the overlap of genes from each gene set with the ordered expression changes of plasmablasts and plasma cells relative to naïve B cells. Overlap is denoted in color. Data in a-c are from 15 mice from two experiments. Microarray analysis in d-f are from an independent experiment of 3 naïve and 3 LPS challenged mice and microarray analysis was performed on 2 independent biological replicates of B cells, 3 plasmablasts from LPS-challenged mice, and 2 plasma cells from the same mice.

To characterize transcriptional stages in B cell differentiation, B cells from naïve mice were compared to plasmablasts and plasma cells isolated from mice three days post-challenge with LPS (**Fig. 2-1d**). RNA was analyzed using Illumina MouseRef-8 BeadArrays covering 16,181 genes after quality control. Hierarchical clustering of all expression data stratified samples by cell type, indicating each cell type was transcriptionally unique (**Fig. 2-1e**). This observation was supported by principle components analysis (PCA), which stratified each biological replicate by cell type (**Fig. 2-1f**). Compared to naïve B cells, plasmablasts and plasma cells up-regulated 937 and 567 genes, and down-regulated 1,016 and 501 genes, respectively (FDR≤0.01, ≥2 fold-change; **Supplementary Table 2-1**). Despite the distinct gene expression programs of each cell type, 544 genes exhibited similar regulation in plasmablasts and plasma cells as compared to B cells (**Fig. 2-1g**). Gene ontology annotation of differentially expressed genes (DEGs) indicated common down-regulated genes were involved in hematopoiesis (e.g. *Hhex*), immune system development (*Bcl2*, *Irf8*), and antigen presentation (*Ciita*, *H2-Aa*, *H2-Ab1*, *H2-Eb1*). Up-regulated genes in plasmablasts were associated with mitosis and cellular division (**Supplementary Table 2-2**). These observations were supported by Gene Set Enrichment Analysis (GSEA), which indicated plasmablast and plasma cell gene expression changes were similar to those previously reported in humans [283] and identified mitotic pathways as selectively enriched in plasmablasts (**Fig. 2-1h**, **Supplementary Table 2-3**).

Plasma cells undergo targeted DNA hypomethylation

To directly measure epigenetic changes during B cell differentiation, DNA from B cells, plasmablasts, and plasma cells was extracted, and the DNA methylation state was

determined using Reduced Representation Bisulfite Sequencing (RRBS) [180]. In total, ≥10X coverage was achieved at 911,004 common CpGs per cell type (**Supplementary Table 2-4**). Hierarchical clustering of the DNA methylation data separated plasmablasts and plasma cells from B cells, indicating that DNA methylation may play a significant role during B cell differentiation (**Fig. 2-2a**). PCA of DNA methylation data suggested a common epigenetic change in plasmablasts and plasma cells (**Fig. 2-2b**). The distribution of DNA methylation levels showed a characteristic bimodal curve for B cells but plasmablasts and plasma cells had a substantial shift of their highly methylated regions (≥80%) to more intermediate levels (~50-80%) (**Fig. 2-2c**). Indeed, the average DNA methylation level was reduced from B cells to plasmablasts to plasma cells (**Fig. 2-2d**), suggesting that plasmablasts and plasma cells undergo global DNA hypomethylation.

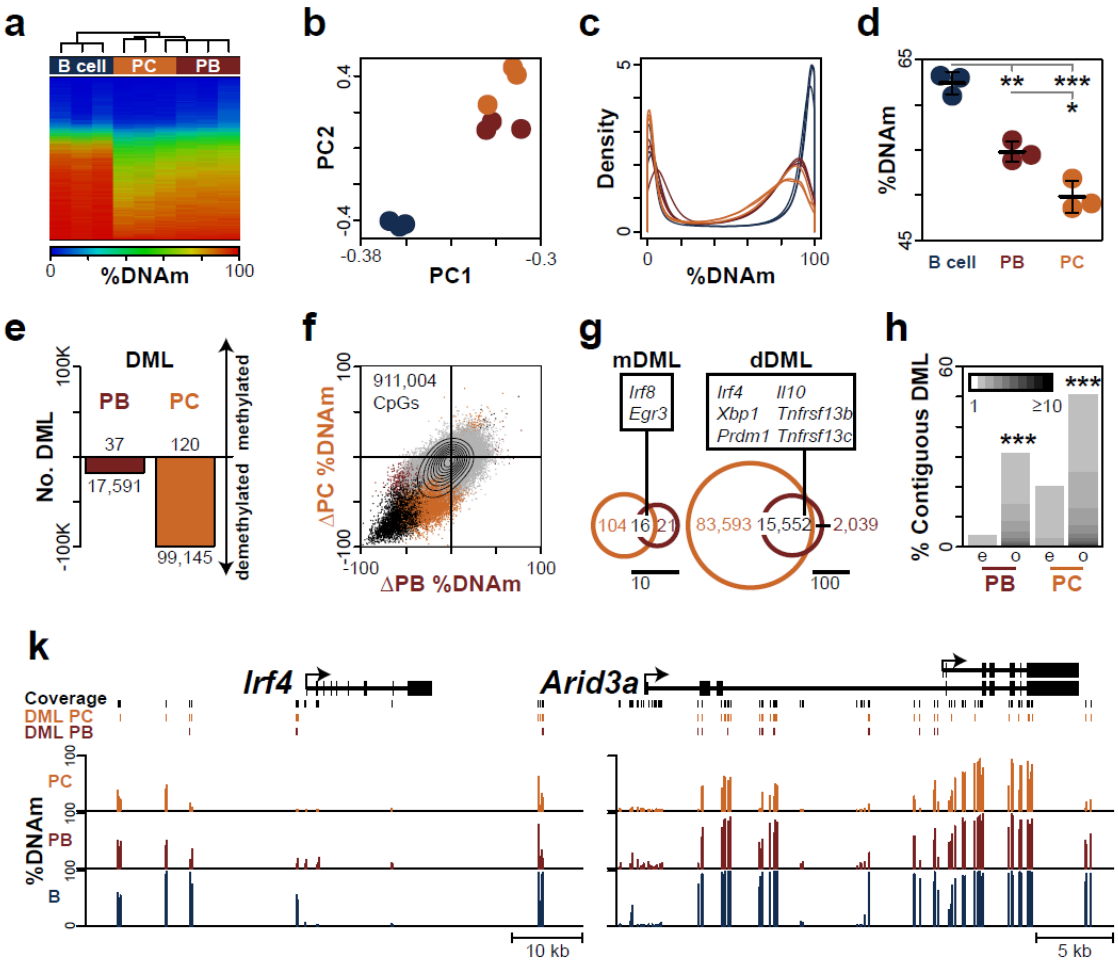


Figure 2-2. Plasmablasts and plasma cells undergo progressive and targeted DNA hypomethylation. **a)** Hierarchically clustered heatmap, **b)** principle components analysis, and **c)** probability distribution of DNA methylation (**DNAm**) values in B cells (blue), plasmablasts (PB; burgundy), and plasma cells (PC; gold). **d)** Average DNA methylation for all commonly covered CpGs. Mean is denoted by thick black line +/- the standard deviation; **P* ≤0.05, ***P* ≤0.01, ****P* ≤0.001; Welch's *t*-test. **e)** Barplot of differentially methylated loci (DML) (FDR ≤0.01; ≥10x coverage; ≥20% change). Methylated DML (mDML) are shown on a positive scale and demethylated DML (dDML) are shown as negative values. **f)** Scatter plot of changes in DNAm between B cells and PBs (x-axis) and PCs (y-axis). DML are colored in burgundy (PB only), gold (PC only) and black (both). The contour plot denotes 10 percentile densities and all covered CpGs are shown. **g)** Venn diagram of mDML (left) and dDML (right) showing overlap of PB (burgundy) and PC (gold) DML. Scale is indicated on the bottom. **h)** Bar plot showing the percent of DML clustered into contiguous regions for PBs (left) and PCs (right). The shade denotes the number of CpGs in the contiguous region, where contiguous is defined relative to

assay coverage. The expected (e) percentiles are shown to the left of the observed (o) PB and PC DML (***P<0.001; permutation test). **k)** Genome plots of *Irf4* and *Arid3a*. Coverage is indicated in black and PB- and PC-specific DML are shown in burgundy and gold, respectively. Average DNAm for B cells, plasmablasts, and plasma cells are shown below. RRBS was performed on three independent biological replicates.

Differentially methylated loci (DML) were identified using a statistical model that accounts for sequencing depth and biological variation [284]. This identified 17,628 and 99,265 DML in plasmablasts and plasma cells, respectively, as compared to B cells (FDR ≤0.01, **Supplementary Table 2-5**). In both plasmablasts and plasma cells, more than 99.7% of the DML identified were demethylated (**Fig. 2-2e**). DML near *Irf4*, *Il10*, *Tnfrsf13b*, *Tnfrsf13c*, and *Egr3* – genes implicated in plasma cell biology [146,147,285–287] – were validated using combined bisulfite restriction analysis [288] (**Supplementary Fig. 2-2**). Moreover, DNA methylation levels in bone marrow $B220^{mid}$ and $B220^{low}$ CD138+ cells were also reduced for CpG loci near *Irf4*, *Prdm1*, *Il10*, and *Cd86* genes, suggesting that the observed DML represent a common epigenetic program of plasma cells (**Supplementary Fig. 2-3**).

Several lines of evidence indicated that this DNA hypomethylation was an organized epigenetic reprogramming and not just a random event associated with rapid cell division. First, although a large number of CpGs coordinately lost DNA methylation in both plasmablasts and plasma cells, 90% of the CpGs interrogated did not significantly change their methylation state (**Fig 2-2f**, see contour indicating density). Second, there was a substantial overlap of DML between plasmablasts and plasma cells (**Fig. 2-2g**). Demethylated loci included CpGs around key genes expressed in plasma cells, such as *Irf4*, *Prdm1*, and *Xbp1*; and methylated DML included important regulators of B cell

differentiation such as *Irf8* and *Egr3*. Finally, DML were clustered into contiguous regions; that is, more than 30% of plasmablast and 50% of plasma cell DML were adjacent to at least one other DML relative to assay coverage (**Fig. 2-2h**). This is more than what would be expected by chance ($P \leq 0.001$). Examples of such contiguous DML occurred at *Irf4*, a transcription factor necessary for B cell differentiation [145,289], and *Arid3a*, a gene known to regulate Ig heavy chain transcription [290] (**Fig. 2-2k**). Other examples included regions around *Prdm1*, which encodes Blimp-1, a master regulator of plasma cell fate [138]; and *Cflar*, a gene involved in the regulation of caspase-induced apoptosis (data not shown). These DML also corresponded with increased gene expression. Together, these data identified a targeted DNA hypomethylation that occurs during B cell differentiation, impacting ~10% of CpG loci assayed.

Demethylated regions overlap enhancers and transcription factor motifs

Transcriptional enhancer regions can be globally identified based on their histone modifications [291]. To test whether plasmablast and plasma cell DML were preferentially occurring at enhancer regions, DML were compared to both active (H3K4me1+, H3K27ac+, H3K4me3-) and poised enhancer regions (H3K4me1+, H3K27ac-, H3K4me3-) [291] in B cells, splenocytes, CH12 lymphoma cells, and disparate tissues using data generated by the ENCODE consortium and others [227,292]. Overlap of DML with cell type-specific enhancers were assessed using an odds ratio, which indicated that DML were generally enriched at enhancers. This was most pronounced in active enhancers of B cells, splenocytes, and CH12 cells (**Fig. 2-3a**), but not in thymus, testis, or brain enhancers. Thus, B cell enhancers are preferentially demethylated during B cell

differentiation, suggesting that these regulatory regions are utilized to facilitate B cell differentiation and the plasma cell gene expression program.

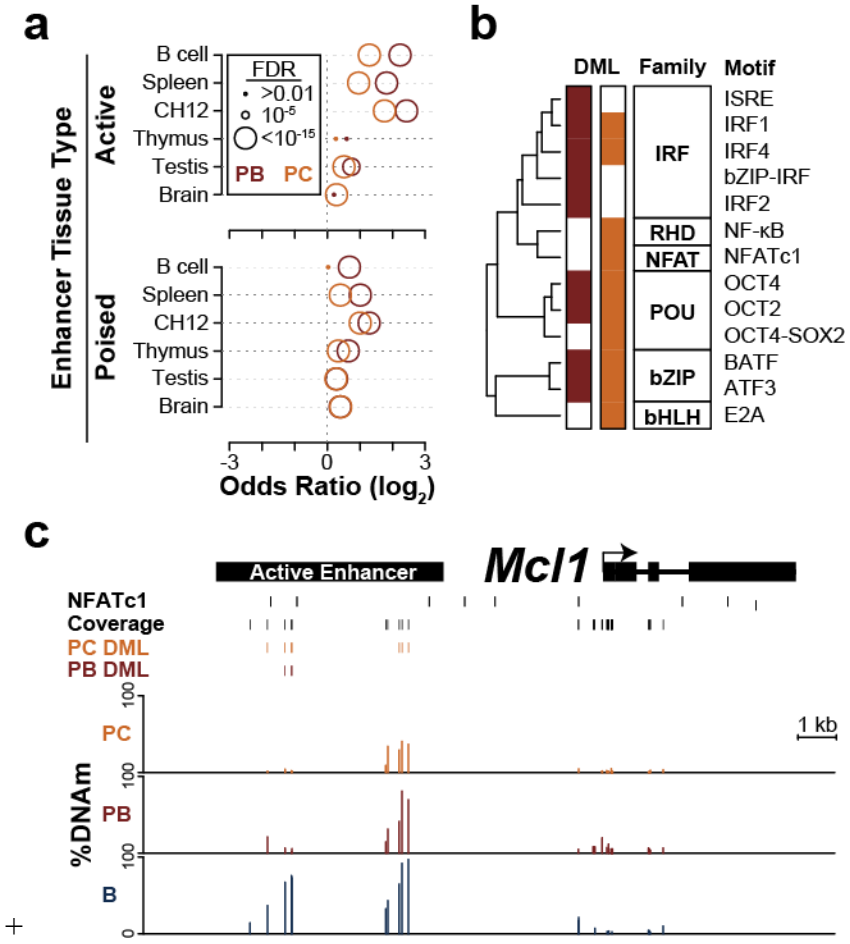


Figure 2-3. Differentially methylated loci preferentially occur at B cell enhancers and near motifs of transcription factors required for B cell differentiation. a) Plot of plasmablast (PB, burgandy) and plasma cell (PC, gold) DML overlap with active and poised enhancer regions identified in B cells, the spleen, the CH12 lymphoma cell line, thymus, testis, and whole brain. The overlap is represented as an odds-ratio (log_2) with the statistical significance represented by the size of the circle. Active enhancers were defined as H3K4me1+ H3K27ac+ H3K4me3- and poised enhacers as H3K4me1+ H3K27ac- H3K4me3- [291]. **b)** Clustering of transcription factor motifs enriched near (≤ 50bp) DML in plasmablasts and plasma cells. Motifs are clustered on the similarity of their position weight matrix. Transcription factor families are denoted in black boxes (IRF: Inteferon Regulatory Factor; RHD: Rel Homology Domain; NFAT: Nuclear Factor of Activated T-cells; POU: Pituitary Octamer Unc-86 transcription factor; bZIP: basic Leucine

Zipper Domain; bHLH: basic Helix-Loop-Helix). The HOMER [259] motif name is labelled on the right. Significance (FDR ≤0.05) was determined from HOMER [259] software. **c)** Genome plot for *Mcl1* with the an active enhancer defined from Sabo *et al*.[292]. RRBS coverage is shown in black with plasmablast and plasma cell DML illustrated as above. Average DNA methylation levels for B cells (B), plasmablasts and plasma cells are shown below.

To determine which transcription factors or families may be active during B cell differentiation, regions within 50bp of DML were searched for enriched transcription factor motifs relative to assay coverage. Transcription factor families enriched at DML included the basic leucine zipper domain (bZIP), the Rel homology domain (RHD; NF-κB), the POU homeobox domain (POU), NFATc1, and interferon regulatory factors (IRF). To illustrate these findings, each transcription factor was clustered by the similarity of its motif with an annotation bar to denote statistical significance in plasmablasts and plasma cells (**Fig. 2-3b**). Several of these factors were previously shown to be critical determinants of B cell fate, including Oct2 [148,149], Irf4 [145,147], Fra1 [293], and NF-κB [294]. An example of DML found in a B cell enhancer includes a region 3.6 kb upstream of the TSS for *Mcl1*, an anti-apoptotic gene required for plasma cell survival [295] (**Fig. 2-3c**). DML within this enhancer corresponded with NFATc1 motifs. These data therefore suggest that DNA demethylation occurs at B cell enhancers and may regulate and/or be regulated by specific transcription factor binding and enhancer usage during B cell differentiation.

Inhibition of DNA methylation facilitates plasma cell differentiation

The above analyses identified dramatic changes in DNA methylation during B cell differentiation and suggest that specific losses of DNA methylation were necessary for

plasma cell formation. To address the function of DNA methylation during plasma cell differentiation, an *ex vivo* model that induces differentiation using LPS, IL-2, and IL-5 [296] was used in conjunction with either carboxyfluorescein succinimidyl ester (CFSE) or cell trace violet (CTV) staining to track cellular division and the DNA methylation inhibitor 5-azacytidine (5-azaC). Splenic B cells were CFSE or CTV labeled and cultured for three days in the presence of LPS, cytokines, and 5-azaC. Flow cytometry analysis revealed a dose-dependent effect of 5-azaC that resulted in a higher frequency of CD138+ plasma cells (**Fig. 2-4a,c**). Moreover, this effect was seen with fewer cell divisions (**Fig. 2-4a,d**). Cell viability analysis demonstrated that the cells were 94-98% viable (**Fig. 2-4a,b**). These data suggest that inhibition of DNA methylation promotes plasma cell differentiation.

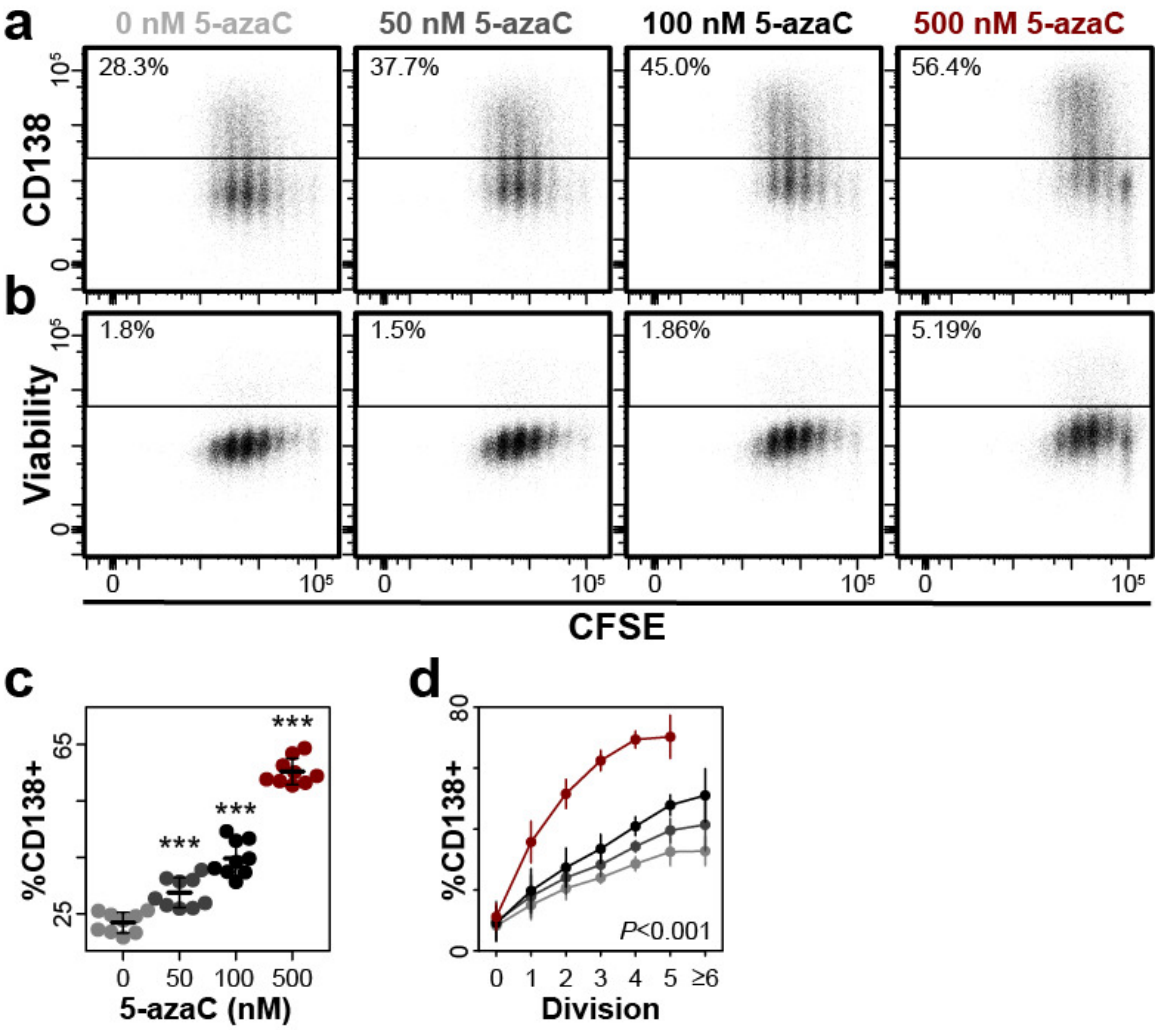


Figure 2-4: Inhibition of DNA Methylation Facilitates Plasma Cell Differentiation**. a-b)** Flow cytometric analysis of CFSE, **a)** CD138, and **b)** viability exclusion dye on B cells differentiated for three days *ex vivo* with LPS, IL-2, IL-5, and treated with increasing amounts of 5-azaC. **c)** Frequency of CD138+ cells. **d)** Frequency of CD138+ cells by cell division. Data are from 3 experiments with 1-5 mice per experiment (N=9). DNA methylation was analyzed for 1 experiment with 3 mice. Mean is denoted by thick black line +/- the standard deviation. Significance was assessed by Welch's t-test (*$P \leq 0.05$, **$P \leq 0.01$, ***$P \leq 0.001$) as compared to untreated cells except for panel **d)** which was determined by an ANOVA.

DNA demethylation and gene expression are coupled to activation-induced cell division

To determine how DNA methylation and gene expression changes occur in relation to

cell division during B cell differentiation *in vivo,* an adoptive transfer model was utilized (**Fig. 2-5a**). B cells from CD45.1+ mice were isolated and labeled with CFSE or CTV prior to being transferred into B-cell deficient CD45.2+ μMT mice [297]. Post-transfer, hosts were challenged with LPS and the splenic compartments were characterized three days post-challenge. Analysis of recipient spleens identified a clear CD45.1+ population of donor cells (**Fig. 2-5b**). CFSE staining of transferred cells identified discrete peaks, indicating cells in LPS-inoculated animals underwent many rounds of division, whereas transferred cells in mice that did not receive LPS underwent minimal division (**Fig. 2-5c,d**). Plasmablasts and plasma cells were only observed in animals that received LPS (**Fig. 2-5e-g**). $B220^{low}$ plasma cells had divided more than $B220^{mid}$ plasmablasts and no differentiation was observed prior to division six (**Fig. 2-5h,i**). B220 expression was only lost on cells that underwent many rounds of division in LPS-inoculated animals (**Fig. 2-5j,k**). These cells were viable and a large portion expressed CD138 (**Supplementary Fig. 2-4**). Unlike B220 and CD138 expression, GL7+ activated cells were present in inoculated animals even in undivided cells and progressively upregulated at later divisions (**Fig. 2-5l,m**). In agreement with previous *ex vivo* experiments [124,280], these data link differentiation to division and suggest that the observed changes in gene expression and DNA hypomethylation may be related to the many rounds of mitotic division that plasmablasts and plasma cells undergo *in vivo*.

To determine the relationship between DNA methylation and gene expression at successive stages of differentiation, cells representing divisions 0, 1, 3, 5, as well as cells that divided at least 8 times and were either CD138- or CD138+ (hereafter referred to as '8-' and '8+', respectively) were FACS isolated to ≥90% purity and subjected to paired

RNA-seq and RRBS (**Fig. 2-5n,o**). RNA-seq was performed using a specified number of cells with External RNA Controls Consortium (ERCC) synthetic mRNA spike-in controls to quantitate the average number of mRNAs on a per cell basis (see methods). A dual restriction enzyme RRBS protocol was performed (*MspI*, *TaqI*) resulting in greater coverage (*N*=1,639,598 CpGs; **Supplementary Table 2-4**). The resulting datasets were validated by direct comparison to data generated by RT-qPCR and a qPCR DNA methylation assay, on independently isolated cells (**Supplementary Fig. 2-5**).

Heatmaps of mRNA expression and DNA methylation suggested an overall increase in gene expression and decrease in DNA methylation at later cell divisions (**Fig. 2-5p,q**). Indeed, populations 5, 8-, and 8+ had on average more mRNAs per cell as compared to division 0 (**Fig. 2-5r**). This corresponded with a significant decrease in average DNA methylation in populations 8- and 8+ (**Fig. 2-5s**). The distribution of mRNA levels identified a uniform amplification across a wide range of expression levels (**Fig. 2-5t**), suggesting that the increased mRNA levels were not solely attributable to a few highly expressed mRNAs such as those encoding immunoglobulins. As expected from the above results, highly methylated loci in cell populations 8- and 8+ were shifted to more intermediate levels (**Fig. 2-5u**). PCA indicated that smaller, progressive changes in gene expression occurred at early divisions, but the greatest variation was between division 8+ and all others (**Fig. 2-5v**). Similarly, small gradual changes in DNA methylation were observed at early divisions, but there was a substantial difference after 8 divisions for DNA methylation, suggesting that some DNA methylation changes may precede expression changes (**Fig. 2-5w**). Cumulatively, these data identify a division-dependent global amplification of mRNA that corresponds with DNA hypomethylation.

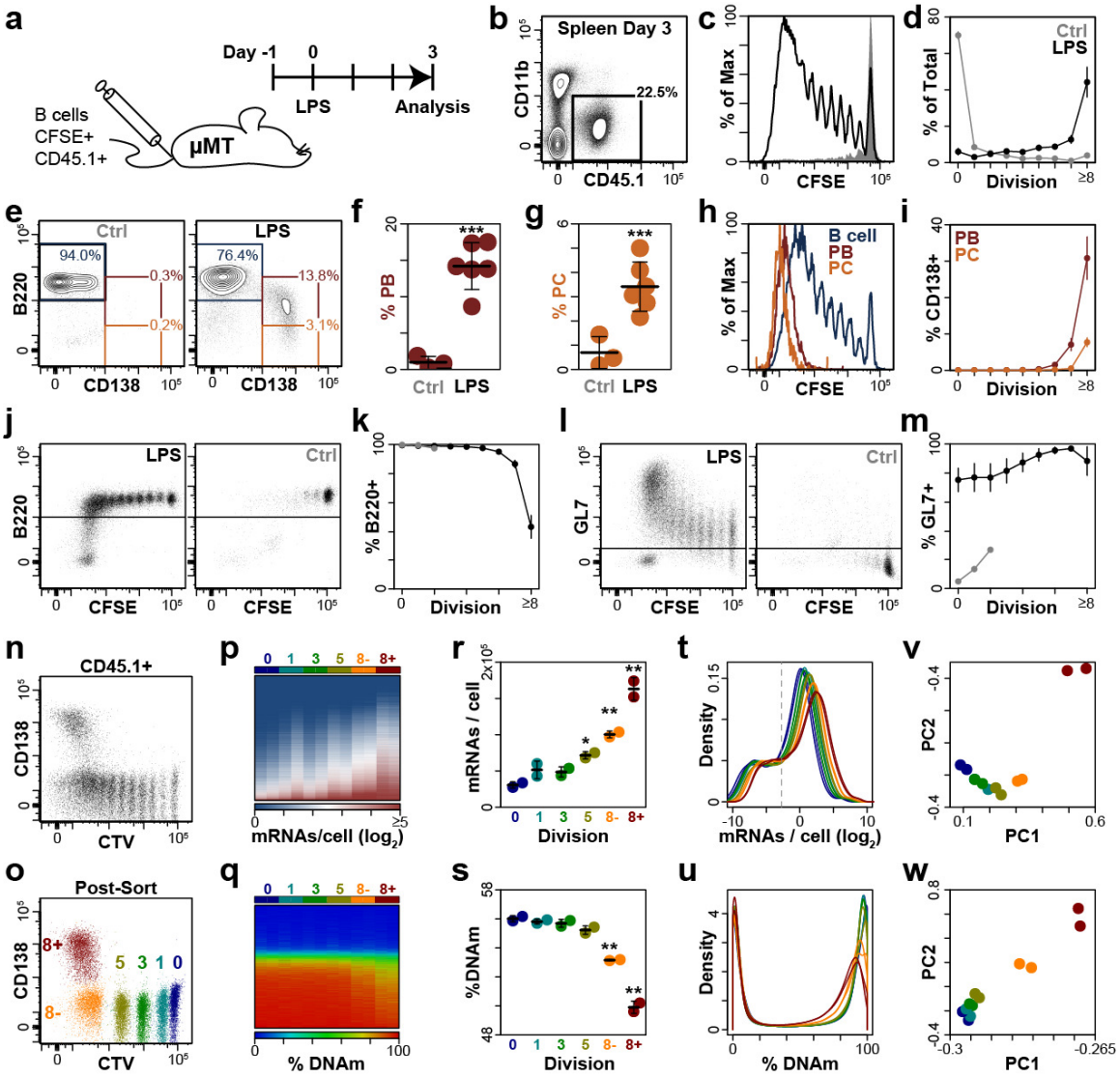


Figure 2-5: Transcriptional amplification and DNA hypomethylation coincide with cellular division. **a)** Schematic of experimental design. **b)** Flow cytometry showing representative population of CD45.1+ B cells in the spleen of µMT hosts 3 days post-LPS challenge. **c)** Representative CFSE histograms for CD45.1+ cells from an LPS challenged (black) and unchallenged control (Ctrl; gray). **d)** Quantification of divisions for LPS and control mice. Mean is represented by the circle and standard deviation by the vertical line. **e)** Representative frequencies of B cells, plasmablasts (PB; burgundy) and plasma cells (PC; gold) determined using B220 and CD138 markers for LPS challenged (right) and control mice (Ctrl; left). **f-g)** Quantification of CD138+ B220^{mid} plasmablasts (**f**) and CD138+ B220- plasma cells (**g**) shown in part **e**. **h)** Histogram of CFSE staining on B cells, plasmablasts, and plasma cells. **i)** Quantification of plasmablasts and plasma cells by division. **j)** B220 staining by division for LPS

challenged (left) and control (right) mice. **k)** Quantitation of B220+ cells by division. **l)** GL7 staining by division for LPS challenged (left) and control (right) mice. **m)** Quantitation of GL7+ cells by division. **n)** Cell Trace Violet (CTV) staining and CD138 expression for CD45.1+ adoptively transferred B cells. **o)** Representative post-sort purity shown for six populations. All samples are projected on to one plot and represented in different colors. **p-q)** Heatmaps of transcript expression **(p)** and DNA methylation (**q**) for biological replicates of the six populations shown in panel **o**. **r-s)** Quantification of average mRNAs per cell **(r)** and average DNA methylation **(s)** measured across 1,639,598 CpGs. Significance is relative to undivided cells (division 0). **t-u)** Probability density for mRNA expression **(t)** and DNA methylation **(u)**. The dashed gray line indicates the 90% detection level (0.147 molecules / cell) based on ERCC spike-in controls. **v-w)** Principle component analysis for mRNA expression **(v)** and DNA methylation **(w)**. Panels a-o represent an aggregate from three experiments (6 LPS treated, 3 control mice). Panels p-w were performed on two independent biological replicates from a separate experiment. * $P<0.05$; ** $P<0.01$; *** $P<0.001$; Welch's *t*-test.

Cell division is marked by dynamic gene regulation and progressive DNA hypomethylation.

Division-specific transcriptional and DNA methylation changes were compared to undivided B cells (division 0). Pair-wise comparisons of gene expression levels between divisions confirmed that expression was mostly upregulated, especially at later divisions (**Fig. 2-6a**). Here, regression lines are shown for both the ERCC controls (bright red) and the average expression difference between the compared divisions (dashed black line). DEGs were determined using criteria for both relative and absolute change in expression; thus, to be counted, DEGs must be upregulated 2-fold more than the average increase in expression for any division. DEGs are shown in division-specific colors that denote the first division where the gene was found to be differentially regulated. A similar analysis was performed for DNA methylation, where DML were determined relative to division 0 and colored as described for gene expression (**Fig. 2-6b**). Comparatively few DNA

methylation changes occur at early divisions, and that most changes occurred in cells that had divided many times. Finally, the correlation of gene expression and DNA methylation changes was compared by division (**Fig. 2-6c**). Here, only DML associated with a DEG were shown in color. This identified that 98% of DML associated with DEGs lost DNA methylation and gained expression through division 5. DML in divisions 8- and 8+ were also inversely correlated with gene expression (94% and 72%, respectively). These data suggest that gene expression changes are highly correlated with local DNA methylation changes.

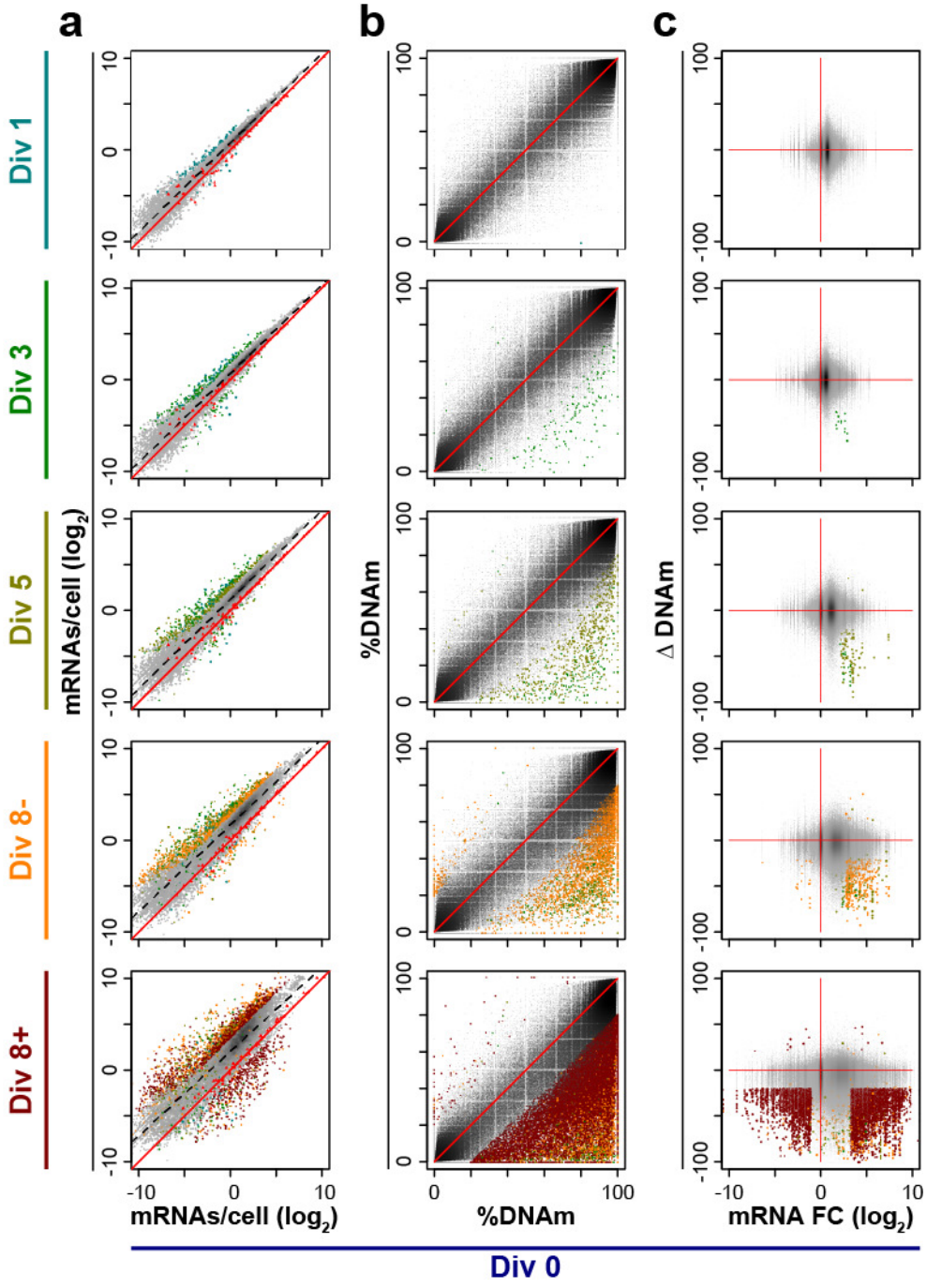


Figure 2-6: Cell division coupled changes in gene expression and DNA methylation. **a-b)** Scatter plot of gene expression **(a)** and DNA methylation **(b)** changes between undivided B cells (x-axis; division 0) and distinct divisions of differentiating B cells (y-axis). Differentially expressed genes **(a)** or differentially methylated loci **(b)** are shown in color with the color denoting the earliest division at which the gene or loci were determined significant. Spike-in ERCC controls are used to calculate average mRNAs per cell and are shown as bright red

triangles with a solid bright red regression line. The average regression between the two comparisons is shown as a black dashed line. **c)** Scatter plot of mRNA fold-change (FC; x-axis) by change in DNA methylation (y-axis). Genes or CpG loci that are both differentially expressed and differentially methylated are colored with the color denoting the division at which they first become differentially regulated. Data are the average from two biological replicates where all 6 populations were analyzed.

Quantitation of division-specific DEGs, as compared to undivided B cells, revealed that 86% of DEGs were upregulated with increasing cellular division and comparatively few genes were down-regulated until division 8+ (**Fig. 2-7a**). The functions of these DEGs were annotated using gene ontology and GSEA (**Supplementary Tables 3, 4**). The top unique ontology for up- and down-regulated genes in each division were displayed using a heatmap to compare the significance across other divisions (**Fig. 2-7b,c**). Cell cycle related genes were upregulated in divisions 5, 8-, and 8+, but division 8+ genes were also enriched for metabolic processes that were not upregulated in any other population. Ontologies associated with downregulated genes were less significant until division 8+ where activation genes, such as *Bcl6*, *Cd83*, *Cd86*, and *Lck,* were downregulated. Similarly, GSEA found a large number of gene sets associated with upregulated genes, but almost no gene sets enriched in downregulated genes until division 8+ (**Fig 2-7d**). Upregulated gene sets were associated with proliferation, cell division, previous studies of human plasma cells [283], and pathways associated with endoplasmic reticulum (ER) stress and processing of proteins. Examples of enriched gene sets included the KEGG proteasome and Myc targets (**Fig 2-7e**). Downregulated genes sets included those defined by NF-κB and type-I interferon signaling (**Fig. 2-7e**). Examples of genes that were dynamically regulated across cell division are indicated (**Fig. 2-7f**). In total, gene expression was globally upregulated with

increasing division, yet annotation of genes preferentially increased or decreased indicated that pathways important for B cell and plasma cell biology were selectively regulated.

Division-specific DNA methylation changes were compared to division 0 and the number of DML indicated that most changes were hypomethylation events at later divisions (**Fig. 2-7g**). DML at earlier divisions continued to lose DNA methylation at later divisions (**Fig. 2-7h**). As in the analyses of plasmablasts and plasma cells above, DML were organized into contiguous blocks of DML, identifying these changes as focal epigenetic events (**Fig. 2-7i**). As above, division-specific DML had a substantial overlap with enhancers in B cells, splenocytes, and CH12 cells (**Fig. 2-7j**). This was most pronounced for DML in earlier divisions, and considerably less for enhancers in thymus, testis, and brain tissue. Overlap of DML with known transcription factors was determined using HOMER [259] and results were clustered based on the transcription factor motif sequence similarity. This revealed that DML preferentially occurred near (≤50 bp) five families of transcription factors, including bZIP, IRF, MADS, POU, and RHD (**Fig. 2-7k**). DML occurred near NF-κB (RHD) and AP-1 (bZIP) motifs starting at division 3; whereas IRF and POU family members were only enriched at DML specific to divisions 8- and 8+, suggesting a hierarchy of transcription factor utilization. DML near these motifs lost more DNA methylation than those distal to such motifs (**Fig. 2-7l**). These transcription factor motifs were used to identify genes putatively regulated by these factors. Genes that contained a transcription factor motif with an associated DML were expressed at higher levels in each division than genes that contained the motif without an

associated DML (**Fig. 2-7m**). These data therefore suggest that ordered demethylation at cis-regulatory elements are driving transcriptional changes.

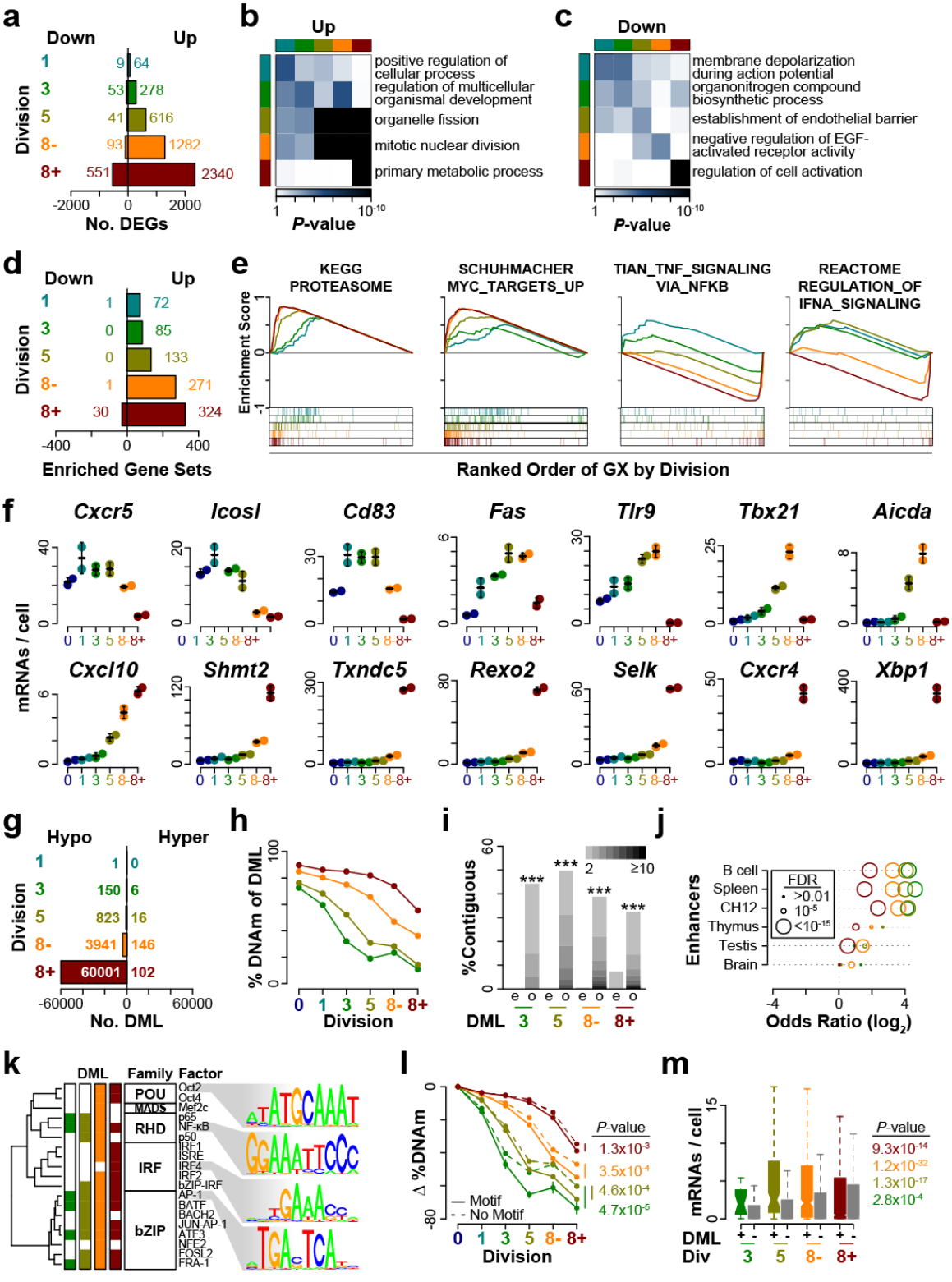


Figure 2-7: Dynamic gene expression changes correspond with a hierarchy of DNA hypomethylation reprogramming events. **a)** Barplot of the number of differentially expressed

genes (DEGs). DEGs were determined relative to division 0 expression and included criteria for both relative and absolute changes in expression (fold-change ≥2; FDR ≤0.01). **b-c)** Heatmap of gene ontology results for **b)** up- and **c)** down-regulated genes. Each row represents the top ontology for division specific gene expression changes. The significance of each ontology in other divisions is depicted by the columns. **d)** Barplot of positively and negatively correlated gene sets determined by Gene Set Enrichment Analysis (FDR ≤0.01). **e)** Examples of enriched (left) and depleted (right) gene sets. The enrichment score for division-specific expression changes are shown in the respective division's color (top). Overlap is shown on the bottom, where division-specific expression changes are indepently sorted and overlap with a gene from the genes set is shown by a colored mark. **f)** Examples of differentially expressed genes. Mean +/- standard deviation is shown. **g)** Barplot of the number of differentially methylated loci (DML; FDR ≤0.01, ≥20% change in DNA methylation). Loci that gain DNA methylation (hypermethylated) are plotted as positive values and loci that lose DNA methylation (hypomethylation) are plotted as negative values. **h)** Plot of DNA methylation levels for division specific DML. **i)** Barplot of the percent of DML that fall into contiguous blocks. The expected (e) number is shown next to the observed (o) (***P ≤0.001; permutation testing). **j)** Overlap of division-specific DML with tissue-specific enhancers is represented as an odds ratio as determined using Fisher's exact test relative to assay coverage. Tissue-specific enhancers were determined using published data from Encode and others [227,292]. **k)** Clustering of transcription factor motifs that were found significantly enriched near division-specific DML. Division-specific significance is indicated by the green (3), gold (5), orange (8-) and dark red (8+) annotation bars. The transcription factor family and HOMER factor name are shown (FDR ≤0.05). Examples of position weight motifs are given for 4 factors. **l)** Demethylation differences for DML that are near a transcription factor motif shown in part e (solid line) or do not (dashed line). *P*-values are determined by ANOVA. **m)** Gene expression for all genes that contain any transcription factor motif shown in part **k)** with or without a DML. Boxplots represent the 25%, median, and 75% quartiles. *P*-values were determined by Wilcoxon-rank sum test. Data represent 6 populations from 2 biological replicates.

Global gene expression and average DNA methylation levels were highly anti-correlated with each other (**Fig. 2-8a**). To understand individual correlations of gene expression with DNA methylation at each division, DML were annotated to the closest

gene (within 100 kb). This resulted in 11,456 DML being annotated to a DEG. The similarity of individual DML-DEG correlations were compared using a normalized Euclidean distance metric such that two DML-DEG patterns that where exactly the same would have a distance of 0, whereas the most divergent were assigned a distance of 1. Results were categorized using K-means clustering and displayed in a heatmap (**Fig. 2-8b**). The summarized correlations of change in gene expression and DNA methylation were plotted (**Fig. 2-8c**). Cluster 1 was negatively correlated and tended to lose DNA methylation prior to gaining gene expression. Clusters 2 and 3 lost DNA methylation during each division, and were transiently expressed during divisions 1 through 8- and repressed in division 8+. Cluster three was expressed at higher levels in divisions 0 through 5. Cluster 4 was the largest and most homogenous cluster, composed of DML that lost DNA methylation in proportion to increased gene expression, resulting in the highest average expression of any category. Gene ontology annotation of the genes in each cluster revealed distinct functional categories with the top ontology displayed in a heatmap (**Fig. 2-8d**). ER stress was enriched in clusters 1 and 4, but metabolic processes were exclusive to cluster 4 genes. Cluster 2 was associated with anatomical morphogenesis and cell adhesion (**Supplementary Table 2**). Cluster 3 was composed of genes involved in leukocyte activation. Examples of cluster 1 and 4 DNA methylation and gene expression changes could be found at regions near *Prdm1* and *Il10*, important regulators of plasma cell biology [138,285] (**Fig. 2-8e,f**). More novel examples of gene regulation and DNA methylation remodeling included *C1qbp*, a complement protein believed to also be involved in ER stress and mitochondrial energy metabolism [298], as well as *Rexo2*, an RNA exonuclease also thought be important in mitochondrial function

[299] (**Fig. 2-8g,h**). Examples of clusters 2 and 3 included several loci near the B cell activation genes *Cd80*, *Cd83*, *Cd86*, as well as *Aicda*, which encodes AID, a protein required for class-switch recombination and somatic hypermutation, and *Ablim1*, a molecule involved in cell adhesion and recently implicated in B cell differentiation [286] (**Fig. 2-8i-j**, and data not shown). These analyses revealed several regions where DNA methylation was lost with transient gene expression prior to silencing upon further differentiation. Together, these results suggest that DNA methylation reflects a historical footprint of gene regulation by cell division in B cell differentiation.

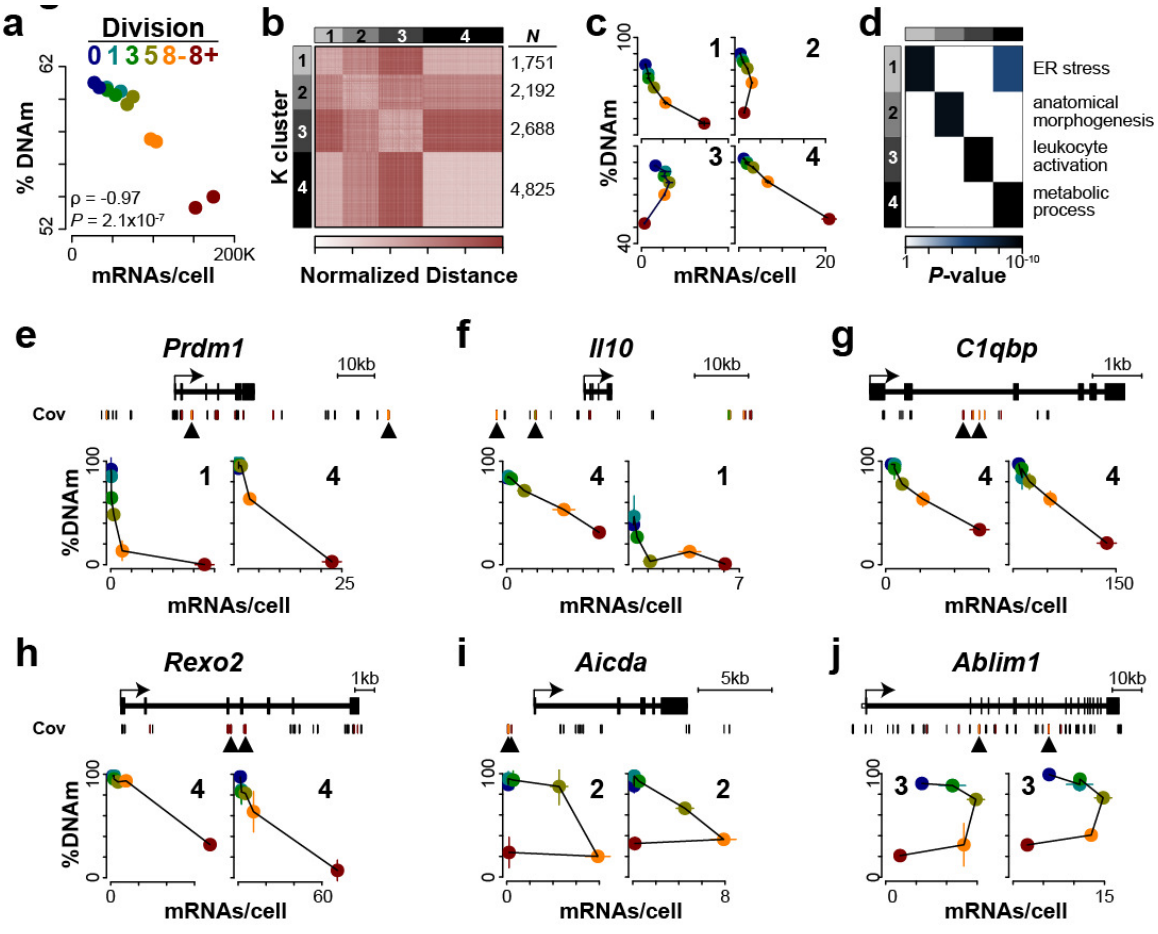


Figure 2-8: Global and local correlation of DNA methylation and gene expression. **a)** Scatterplot of average DNA methylation level by total number of mRNAs per cell. The

association *P*-value (determined using an ANOVA) and Spearman's correlation coefficient (ρ) are shown. **b)** Heatmap of DEG-DML correspondance**.** Gene expression and DNA methylation changes of 2,431 DEGs with 11,456 neighboring DML were measured using a normalized Euclidean metric and organized using K-means clustering. **c)** Average gene expression (x-axis) and DNA methylation change for the 4 K-means groups. **d)** Heatmap of the top gene ontology result for each of the top 4 K-means DEG-DML correlation groups. Each row represents the top ontology for cluster-specific gene function and how that ontology is enriched in other groups is depicted by the columns. Results are provided in **Supplementary Table 2**. **e-j)** Scaled gene plots for DEGs with DML are shown (top). Specific DEG-DML correlations are shown (below) for CpGs denoted in black arrows. Left and right arrows denoted the left and right correlations shown below each gene plot. Data represent 6 populations from two biological replicates.

Discussion

These data provide insight into the dynamic gene expression and epigenetic changes that occur during B cell differentiation *in vivo*. Plasmablasts and plasma cells had distinct gene expression programs, and both underwent focal DNA methylation changes at up to 10% of their DNA methylome. Remarkably, more than 99% of the DML represented losses in DNA methylation. Although greater in percentage, these observations are consistent with DNA methylation studies of human steady-state plasma cells, which showed 60-90% DNA methylation differences were hypomethylation events compared to B cells [278,279]. This difference in observed DNA hypomethylation may reflect the kinetics of the process, as the plasma cells analyzed here were less than three days old, whereas the half-life of steady-state plasma cells can range from several days to more than a 100 days depending on the subpopulation [125,282,300]. If this is indeed the case, then the data suggest an abrupt and targeted loss of DNA methylation occurs during the initial stages of differentiation, but this may be followed by gradual gains in DNA methylation as plasma cells age.

B cells increased global mRNA levels by more than 5-fold, through division-coupled transcriptional amplification. Surprisingly, this increase was attributable to transcriptional amplification of thousands of mRNAs and not just immunoglobulins. Such global amplification may be required to maintain the B cell fate program while the cells are undergoing massive proliferation and differentiation. Pairing this state-of-the-art gene expression analysis with DNA methylation data from matched cells provided critical insight into the functional categories of gene expression and DNA methylation changes revealing demethylation occurred at transiently expressed activation genes and highly expressed plasma cell genes. These step-wise transcriptional and epigenetic changes were found to be highly correlated.

Plasmablasts and plasma cells were found to contain a large number of partially methylated loci, a phenomenon previously attributed to regulatory regions [193,229]. Consistent with this, DML were overrepresented at B cell enhancers and binding motifs of transcription factors required for B cell differentiation [145–149,293]. This was particularly pronounced in DML of less differentiated cell types (e.g. plasmablasts versus plasma cells and Division 3 versus Division 8+ cells). This suggests that certain B cell enhancers are utilized upon activation, but more differentiated cell types increasingly make use of novel regulatory elements. This is further supported by the co-localization of transcription factor binding motifs and DML where NF-κB and AP-1 motifs were enriched starting at division 3; whereas Oct-2 and IRF motifs were only enriched in cells that divided 8 times or more. Interestingly, NF-κB and AP-1 factors are directly induced through TLR4 signaling as a result of LPS stimulation. The data therefore indicate that

the DNA hypomethylation observed represents a hierarchy of cis-regulatory events during B cell differentiation.

Currently, known mechanisms of DNA demethylation include enzymatic-dependent active [189,192] and replication-coupled passive processes. Given the rapid replication of differentiating B cells, it is likely that a passive process, partially accounts for the DNA demethylation observed here. This is supported by the genome-wide analyses of division-specific DNA methylomes that found the majority of changes occurred in cells that had undergone many rounds of division. Yet, DNA methylation changes were focal at regulatory regions, suggesting a targeted process. One possibility is that DNA demethylation is facilitated by the binding of transcription factors that protect the site from maintenance methylation during cell division [193]. Indeed, DML were enriched near NF-κB, IRF, bZIP, and POU transcription factor family binding motifs. The observed demethylation could also result from 5-hydroxymethylation of cytosines (5-hmC) in B cells, which does not have a known mechanism of maintenance through mitosis. Although there is no evidence for active demethylation during B cell differentiation, such a process cannot be ruled out.

Treatment of cells *ex vivo* with 5-azaC resulted in augmented plasma cell differentiation in a division-dependent manner, and suggested that loss of DNA methylation is limiting for plasma cell differentiation. This may be in contrast to observations that inhibition of DNA methylation reduces germinal center formation [273]. It is possible that DNA methylation functions to extend B cell activation at the cost of plasma cell differentiation. This is consistent with greater DNA hypomethylation in plasmablasts that divided at least 8 times as compared to B cells that also divided at least

8 times. These division-coupled epigenetic changes may also help explain why the number of divisions a B cell has undergone directly corresponds with the potential of that B cell to differentiate [124]. Such epigenetic mechanisms may contribute to both the variability of the B cell response at the population level, as well as the high correlation of cell fates between sibling B cells [152]. These data don't preclude other models of differentiation such as the asymmetric division reported in T cells [301], or a transcription factor centric model where substantial evidence indicates that many transcription factors drive or inhibit plasma cell formation [20]. Rather these data provide insight into an additional and critical layer of complexity that regulates plasma cell formation. The results support a model of DNA methylation remodeling that is coincident with gene expression and reflects the cis-regulatory history of plasma cells and the epigenetic reprogramming events through cell division and differentiation.

Acknowledgements

We acknowledge Royce Butler for mouse care. We thank Drs. Paula M. Vertino, Paul A. Wade, Lawrence H. Boise, Hrisavgi D. Kondilis-Mangum for helpful comments and critique in reading the manuscript. We also thank Drs. Karen N. Conneely, and Hao Wu for statistical advice and Ryan Martinez for help with flow cytometry. We thank the Genome Technology Center at NYU for expertise in Illumina sequencing, the Emory Flow Cytometry Core for expertise in FACS, and the Emory Integrated Genomics Core for running high-sensitivity DNA Bioanalyzer analysis.

Funding

This work was supported by Emory University School of Medicine institutional funds to JMB and a National Institutes of Health grant to J.M.B (RO1 GM47310; RO1 AI123733)

and a fellowship to B.G.B. (F31 AI112261), as well as a training grant to JMB and BGB (T32 GM008490). A.P.R.B was supported in part by T32 AI007610.

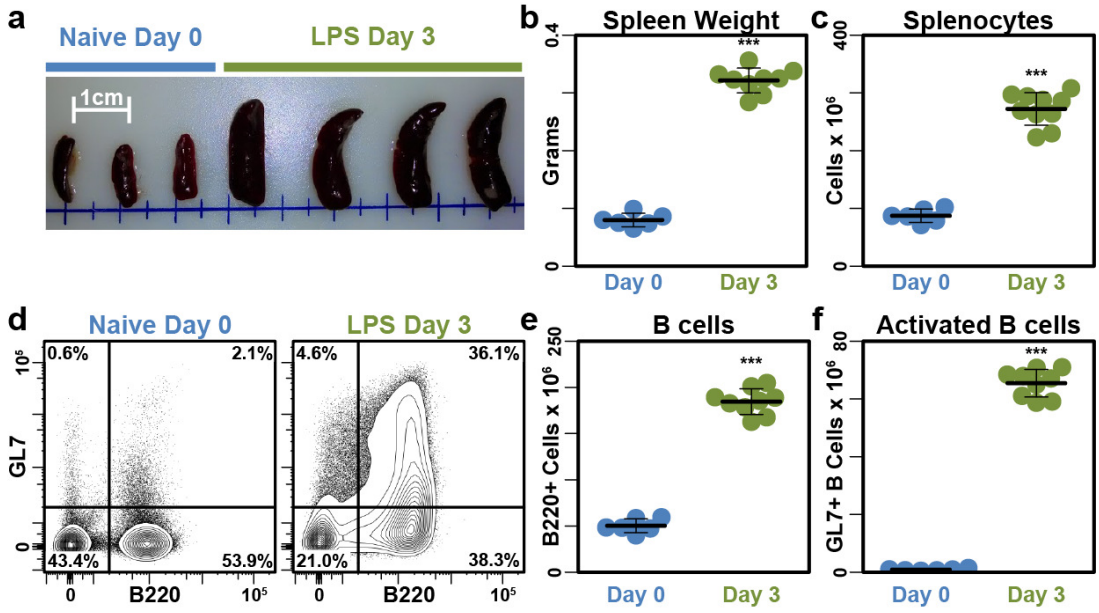
Author Contributions

B.G.B. contributed to experiment conception and design, performed the DNA methylation analyses, mouse experiments, RNA-seq analysis, bioinformatic analyses, and wrote the paper. C.D.S. contributed to experiment conception and design and performed the RNA microarray analysis. A.P.R.B. provided technical expertise with mouse experiments. J.M.B. contributed to experiment conception and design and wrote the paper. All authors provided editorial input.

Competing Financial Interests

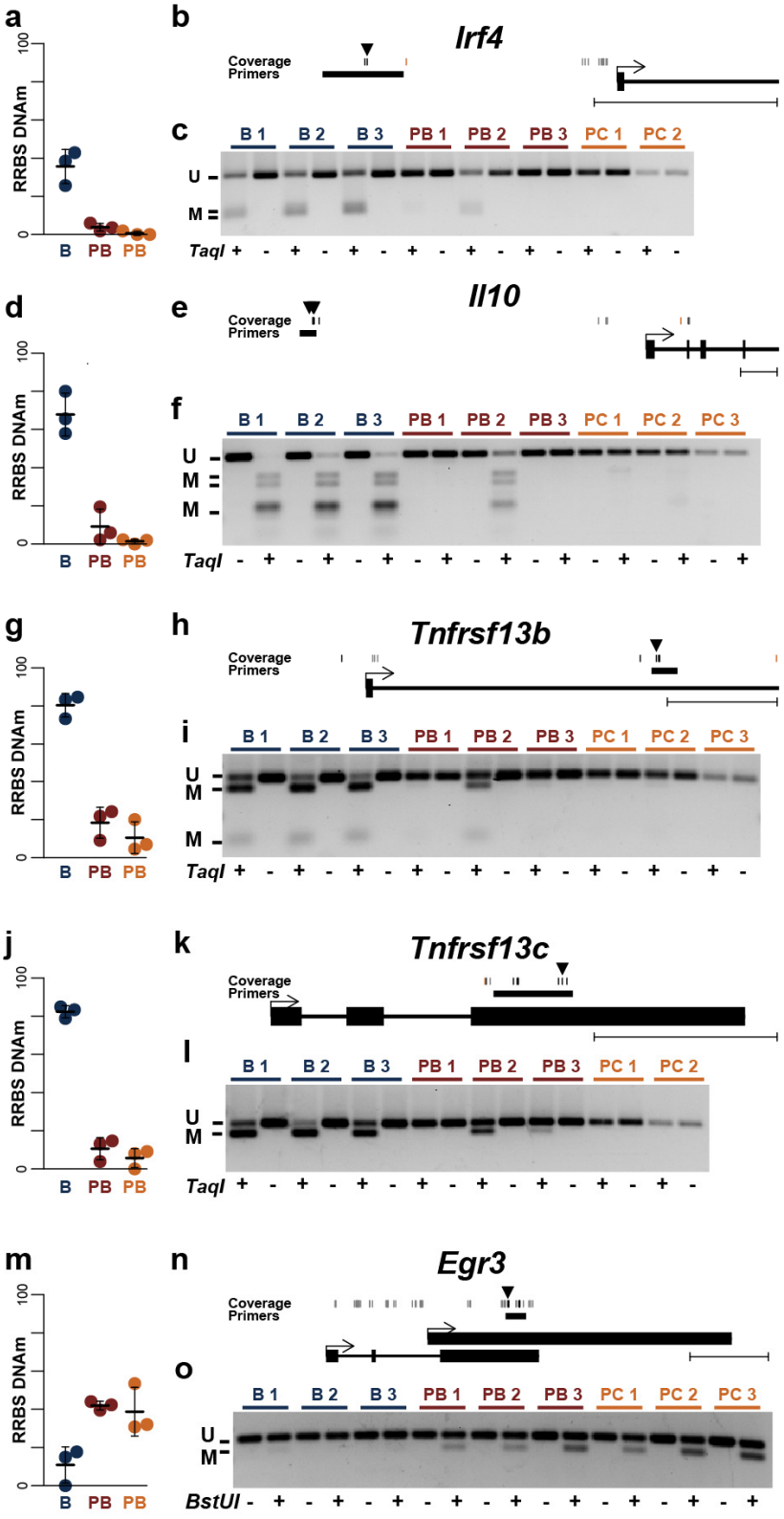
All authors declare no competing financial interests.

Supplemental Figures



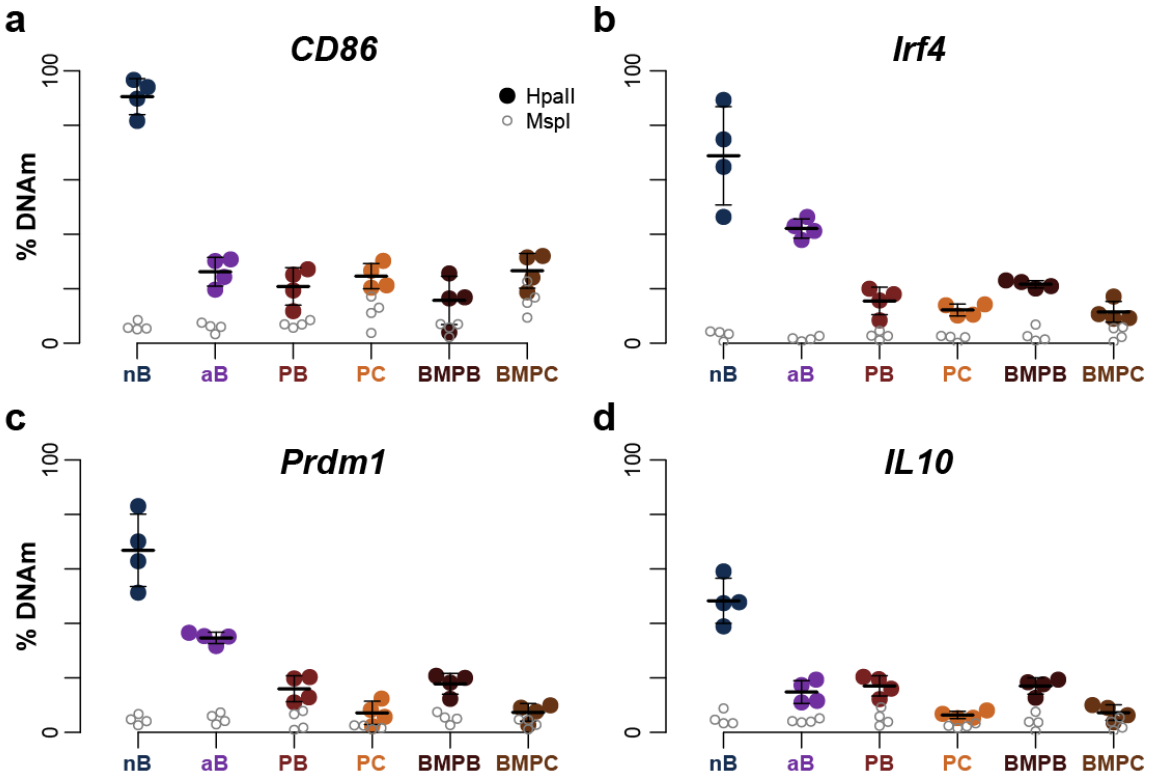
Supplementary Figure 1. LPS induces a robust B cell response. Comparison of spleen **a)** size, **b)** weight and **c)** total cells from naïve (blue) and LPS challenged mice (green). **d)** Representative

analysis of splenic B cell frequency in naïve and LPS-challenged mice indicate proportional increase in B cell frequency primarily of the GL7+ activated population, with quantitation of **e)** total B cells and **f)** GL7+ activated B cells per spleen**.** Mean and standard deviation are represented on beeswarm plots, ***$P < 0.001$, Welch's *t*-test. Data are from 2 independent experiments with 9 LPS challenged mice and 6 control mice.

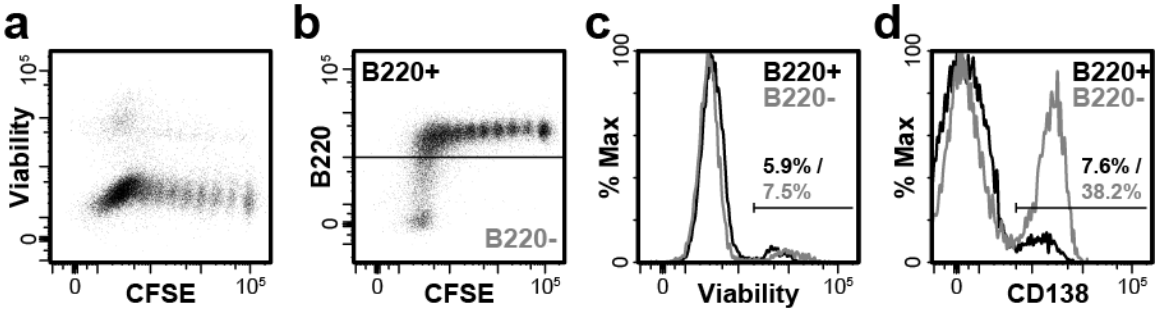


Supplementary Figure 2. Validation of statistically representative and biologically relevant differentially methylated loci (DML). **a, d, g, j, m)** RRBS data are shown for B cells (B; blue),

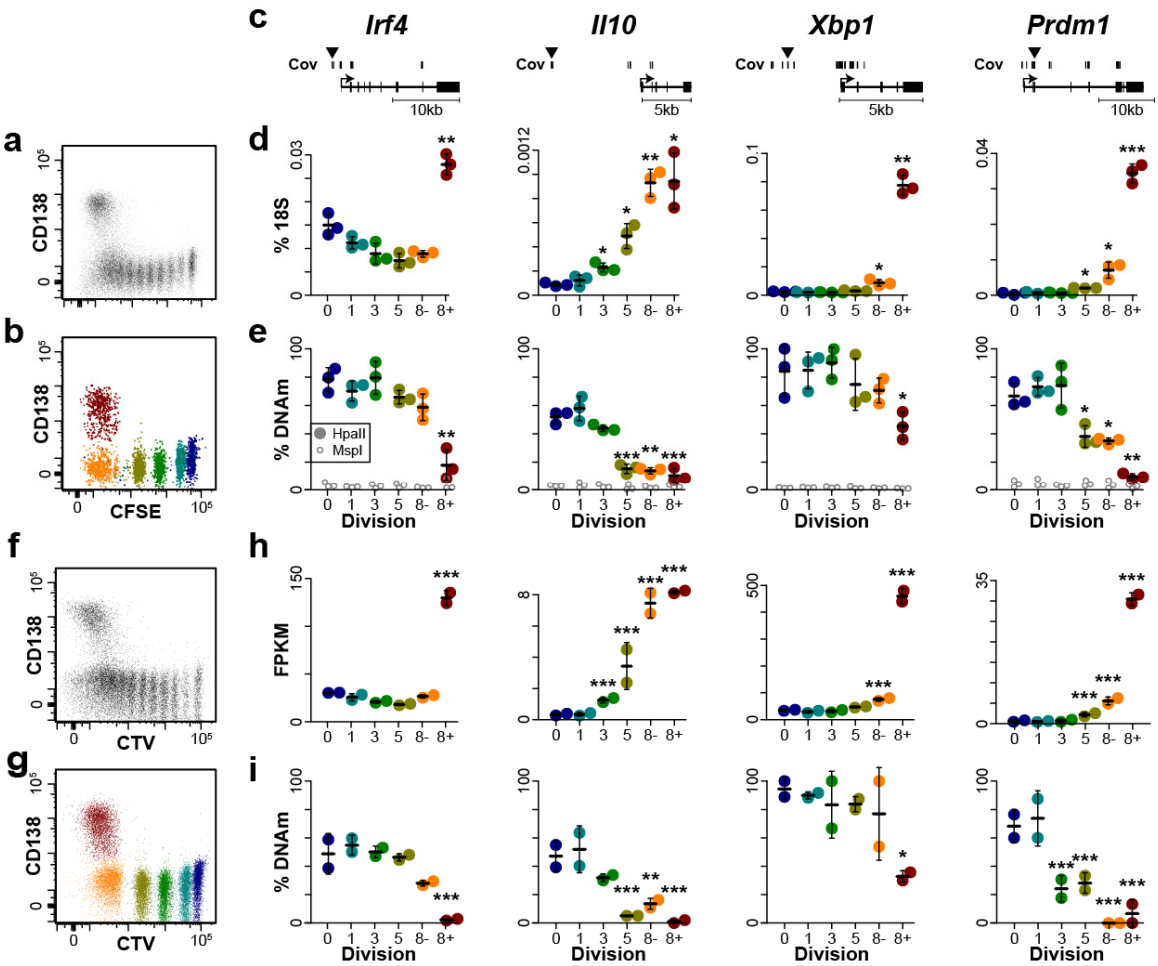
plasmablasts (PB; burgundy), and plasma cells (PC; gold) for DML near *Irf4* **(a)**, *Il10* **(d)**, *Tnfrsf13b* **(g)**, *Tnfrsf13c* **(j)**, and *Egr3* **(m).** **b, e, h, k, n)** Genome schematic of regions near the validated DML. RRBS coverage is shown with vertical black lines and the DML interrogated is indicated with a black arrow. Primer regions are shown below coverage. UCSC KnownGenes are plotted with exonic and intronic sequence denoted with thick and thin black lines, respectively. Scale on the bottom right indicates 1 kb. **c, f, i, l, o)** Combined Bisulfite Restriction Analysis (COBRA) validation of DML. Enzyme digested and mock digested samples are run adjacent on the agarose gel and indicated by a "+" and "-" at the bottom of the gel image, respectively. Biological replicates are indicated on top. Plasmablasts and plasma cells were obtained from the same mice. Six mice were used in this experiment.



Supplementary Figure 3. DNA methylation in B cell subsets. **a-d)** Plasmablast and plasma cell DML near *Cd86*, *Irf4*, *Prdm1*, and *Il10* were validated in naïve B cells (nB), GL7+ activated B cells (aB), CD138+ B220^{mid} splenic plasmablasts (PB), CD138+ B220^{low} splenic plasma cells (PC), CD138+ B220^{mid} bone marrow cells (BMPB), and CD138+ B220^{low} bone marrow cells (BMPC). DNA methylation was measured as a ratio of mock digested DNA to *HpaII* digested DNA using qPCR primers. *MspI* digested DNA serves as a negative control.



Supplementary Figure 4. Viability and CD138+ expression on B220- cells. **a)** Flow cytometry analysis of CFSE and viability exclusion dye on CD45.1+ cells transferred into µMT hosts that were challenged with LPS as described in **Fig 5**. **b)** CFSE and B220 expression showing B220+ and B220- populations. **c-d)** Histograms of **c)** viability exclusion dye and **d)** CD138 expression on B220+ and B220- populations.



Supplementary Figure 5. Validation of division-specific DNA methylation and gene expression changes. **a-b)** Flow cytometry analysis of **a)** splenic CD45.1+ cells and **b)** post-sort population

from μMT hosts 3 days post-LPS challenge. Post sort analysis represents 6 distinct populations represented in different colors, projected onto one plot. **c)** Schematic of DEGs and associated DML that were interrogated for division-linked gene expression and DNA methylation changes. The interrogated DML is denoted by the black. **d)** Gene expression by division for the genes shown in **(c)**. Expression was determined using RT-qPCR and is plotted as % 18S. Mean is denoted by thick black line +/- the standard deviation. **e)** DNA methylation by division determined for genes shown in **(c)**. DNA methylation was determined using qPCR on HpaII digested DNA relative to mock digested DNA. The methyl-insensitive enzyme MspI was also used as a negative control (see open circles). **f-g)** Flow cytometry analysis as in **a-b)** that was used to analyze specific cell divisions using RNA-seq and RRBS. **h)** Gene expression determined using RNA-seq and **i)** DNA methylation determined using RRBS for the same genes and CpGs shown in **c-e)**.

Methods

Accession codes

Gene expression and RRBS data are available in the Gene Expression Omnibus under accession GSE70294.

Mice and LPS challenge

C57/BL6J mice between 8-12 weeks of age were used for most experiments. All animals were housed by the Emory Division of Animal Resources and all protocols were approved by the Emory Institutional Animal Care and Use Committee. LPS challenge was administrated intravenously using 50 μg of *Salmonella* LPS (Enzo Life Sciences, ALX-581-008). CD45.1+ mice were of strain B6.SJL-*Ptprc*^{a} *Pepc*^{b}/BoyJ from Jackson Laboratories (#002014). B cell deficient mice (μMT) were previously described [297] and were also obtained from Jackson Laboratories (#002288, strain B6.129S2-*Ighm*^{tm1Cgn}/J). For cell division assays, 20 x 10^{6} CTV- or CFSE-labeled CD45.1+ B cells were

adoptively transferred into μMT hosts and allowed to rest for 18-24 hours prior to LPS challenge.

Cell Isolation and Flow Cytometric Analysis

Splenic cell suspensions were made by mechanically forcing spleens through a 40 μm filter and lysing red blood cells with ACK lysis buffer (0.15 M NH_4Cl, 10 mM $KHCO_3$, 0.1 mM EDTA) for 30 seconds prior to quenching the reaction with 4 volumes RPMI 1640 media (Corning Cellgro) supplemented with 10% heat-inactivated fetal bovine serum (Sigma-Aldrich), 1% MEM - non essential amino acids, 100 μM Na pyruvate (Sigma), 10 mM HEPES (Sigma), 0.0035% β-mercaptoethanol (Sigma-Aldrich). Cells were washed and resuspend at 10^7 cells / ml in PBS with 1% BSA and 2 mM EDTA. Prior to staining, cells were blocked with α-Fc (α-CD16/CD32) (Tonbo Biosciences, 2.4G2) at a final concentration of 0.25 μg / 10^6 cells for 15 minutes on ice. Staining panels included α-CD11b and α-CD11c conjugated to FITC or PerCP-Cy5.5 (Tonbo Biosciences M1/70) at a concentration of 0.25 μg / 10^6 cells to remove autofluorescent macrophages. The following stains, antibodies-fluorophore combinations were used to assess cellular phenotype: α-B220-PerCP-Cy5.5 or -PE-Cy7 (Tonbo Biosciences, RA3-6B2) at 0.05 μg / 10^6 cells; α-CD43-FITC (BD #553270) at 0.125 μg / 10^6 cells; α-CD138-PE, -BV421, or -BV711 (BD, 281-2) at 0.025 μg / 10^6 cells; α-GL7-eFluor660 (eBioscience GL-7) at 0.025 μg / 10^6 cells; Viability Violet Stain (Life Technologies L34955), CFSE (Tonbo #13-0850) and CTV (Life Technologies #C34557) both at 10 μM / 10^7 cells / ml. Cells were stained for 30 minutes and fixed using 1% paraformaldehyde prior to analysis. Staining panels included fluorescence minus one (FMO) controls to ensure that correct compensation was applied, as well as isotype controls to assess non-

specific staining. Flow cytometric analysis was collected on a Becton Dickinson (BD) LSRII and FCS files were exported using FACSDiva (v6.2). Analysis of flow cytometric data was conducted in R/Bioconductor (v.3.2.2) using the 'flowCore' (v.1.36.9) package [302] or FlowJo software (v9.7.6). Code is available upon request.

Naïve B cells were isolated using Magnetic Activated Cell Sorting (MACS) negative selection for CD43, CD4, and Ter-11 (Miltenyi #130-090-862) following the manufacturer's protocol. Purity was confirmed by flow cytometric analysis. Plasmablasts and plasma cells were isolated by first enriching the CD138 positive fraction of splenocytes and/or bone marrow using a positive MACS enrichment (Miltenyi #130-098-257) and then FACS using a BD FACS Aria II at the Emory Flow Cytometry Core Laboratory.

Ex Vivo Differentiation and 5-aza-Cytidine Treatment

B cells were differentiated *ex vivo* as previously described [296], but with the incorporation of division tracking dye. B cells were isolated using MACS as above, and stained with CTV or CFSE at a concentration of 20^6 cells / mL in PBS. Cells were differentiated at an initial concentration of 0.5×10^6 cells per ml with LPS (20 μg / ml; Sigma #L2630), IL-2 (20 ng / ml; eBioscience #14-8021), and IL-5 (5 ng / ml; eBioscience #14-8051). Half doses of LPS and cytokines were given on subsequent days. 5-azacytidine (Sigma Aldrich #A2385) was added to cultures every day at concentrations ranging from 50 nM to 500 nM as indicated.

Quantitative real-time PCR (qRT-PCR) Analysis

Division specific gene expression was analyzed using 20,000 – 500,000 cells. Cells were sorted into RLT lysis buffer and total RNA was purified using the Quick-RNA MicroPrep Kit (Zymo Research). The entire RNA yield was reverse transcribed using SuperScript II reverse transcriptase (Invitrogen), diluted prior to quantitative real-time PCR (qRT-PCR) on a CFX96 instrument (BioRad) using SYBR Green incorporation. Expression levels were expressed relative to 18S ribosomal RNA levels for each gene analyzed. A full list of qRT-PCR primers can be found in Supplementary Table 7.

Microarray Analysis

Total RNA from each cell type was extracted using the RNeasy mini prep kit (Qiagen) and used for microarray analysis on the MouseRef-8v2 BeadChip (Illumina). Gene expression data were quantile normalized using GenomeStudio v.2011.1 (Illumina) and exported for analysis. Quality control (QC) steps included mapping all probes to the mouse reference genome (mm9) using Bowtie [303] (v.1.0.0) and removing probes that had multiple alignments, or did not align to a UCSC Known Gene exon [304]. The UCSC known gene database was obtained via the R/Bioconductor package 'TxDb.Mmusculus.UCSC.mm9.knownGene' [305] (v.3.2.2). This resulted in 22,907 of 25,697 probes passing quality control and coverage of 16,181 genes. Differential expression was identified using linear regression implemented in R/Bioconductor (v3.1.3) [305]. Multiple hypothesis testing was applied to probes with signal that was detected in at least one sample (N=10,322) and with a 2-fold change in expression or greater using a Benjamini-Hochburg false discovery rate (FDR) [306]. Those genes with a FDR less than 0.01 were considered significant (Supplementary Table 1). All R code is available upon request.

RNA-seq Analysis

For RNA-seq analysis, 40,000 cells were sorted into RLT buffer (Qiagen) with 1% 2-mercaptoethanol (Sigma), vortexed and snap frozen. Prior to extraction 5 µl of 1:2000 dilution of ERCC synthetic RNAs (ThermoFisher) were added to each sample. RNA was extract using RNeasy Mini Kit (Qiagen) following the manufacturer's protocol. DNA was removed with RQ1 DNase (Promega) at 37˚C for 30 minutes and RNA was purified using 3x AMPPure XP clean up (Beckman Coulter). Stranded mRNA-seq libraries were constructed using KAPA Biosystems Stranded mRNA-Seq Kit following the manufacturer's protocol (KAPA Biosystems). Equal molar amounts of mRNA-seq libraries were amplified for 11 PCR cycles and purified using a 1.5x ratio of AMPPure XP beads (Beckman Coulter). Libraries were pooled in equal molar ratios based on the KAPA Library Quantification Kit (Kapa Biosystems) and sequenced on an Illumina HiSeq2500 with 50 bp paired-end reads.

Mapping and Quantification of RNA-seq Data

RNA-seq data was mapped back to the UCSC mouse genome mm9 using Tophat2 [307] (v.2.0.13) with the following parameters "–p 8 –N2 –max-multihits 1 –read-gap-length 1" and the UCSC Known Genes [304] mm9 transcript file as a guide. The 92 ERCC sequences were added to the mm9 genome as artificial chromosomes. PCR duplicates were determined using Picard (http://broadinstitute.github.io/picard/) and removed from subsequent analyses. Reads that uniquely overlapped mm9 exons were determined in R (v.3.2.2) using the 'summarizeOverlaps' function in mode 'IntersectionNotEmpty' of the 'GenomicAlignments' package [308] (v.1.6.3). Reads per million (RPM) were calculated for each gene based on the number reads in all potential exons for a given gene and the

total number of uniquely mappable reads per sample. Fragments per kilobase per million (FPKM) were calculated based on RPM and the total size of non-overlapping exons for a gene. The number of mRNAs per cell were calculated with the following equation:

$$\frac{mRNAs_{gene\ A}}{cell} = \frac{FPKM_{gene\ A}}{\#\ cells}\ x\ \frac{\sum Molecules_{ERCC}}{\sum FPKM_{ERCC}}$$

Differential Analysis of RNA-seq Data

Differentially expressed genes (DEGs) were determined using EdgeR [309] (v.3.12.0) based on both relative and absolute changes in expression. Gene counts were calculated using all reads mapping to exons of unique UCSC mm9 Known Genes [304], determined as described above. For relative differences, normalization factors were determined using the EdgeR function "calcNormFactors". For absolute differences, the normalization factors were determined as the sum of ERCC FPKM divided by the average ERCC FPKM across all samples. A minimum 2-fold change was imposed upon criteria for both relative and absolute differences. *P*-values calculated by EdgeR were corrected for multiple hypothesis testing using Benjamini-Hochberg FDR correction [306]. In total, DEGs had an FDR ≤0.01 with a fold-change ≥2 using both relative and absolute criteria to be considered significant. R code is available upon request.

Bioinformatic Analyses of Expression Data

Heatmaps and hierarchical clustering of gene expression data used an 'average' or unweighted pair group method with arithmetic mean agglomeration method applied to the Z-score normalized probe signal (microarray) or average number of mRNAs / cell

(RNA-seq) using the R/Bioconductor functions 'hclust' and 'image' in a method very similar to that employed by the 'heatmap' function [305] (R code available upon request). Principle components analysis was done using the R/Bioconductor function 'prcomp' also applied to Z-score normalized expression data.

Gene ontology analysis was conducted on differentially expressed genes (DEGs) using the R/Bioconductor package GOstats (v2.32.0) [310]. For microarray data, all genes with probes that passed QC were used as background. For RNA-seq data, all UCSC mm9 Known Genes [304] were used as background. Gene Set Enrichment Analysis (GSEA v2.1.0) was performed using the pre-ranked list option. For microarray data the rank was determined by the average *t*-statistic for all probes mapping to a given gene. Only probes that were detected in at least one sample were used. For RNA-seq data the rank was determined by the $-\log_{10}$(FDR) x sign(fold-change). Here, the FDR is the average FDR determined using both absolute and criteria described above.

DNA Methylation Assays

DNA isolation was obtained from cells digested overnight with Proteinase K and RNase at 67°C and extracted using phenol-chloroform-isoamylalcohol and ethanol precipitation. RRBS libraries were made from 10 to 500 ng of DNA and were digested overnight with 20 U *Msp*I (New England Biolabs) following the manufacturer's protocol. RRBS libraries for B cell divisions were also digested with *TaqI* in separate reactions (New England Biolabs). Digested DNA was purified using a 1.8X Solid Phase Reversible Immobilization (SPRI) clean-up with Agencourt AMPure XP beads (Beckman Coulter). Illumina compatible sequencing adaptors were used and contained fully methylated cytosine residues and were either NEXTflex Bisulfite-Seq Barcodes (BIOO Scientific)

or were designed similar to those previously described [311] and synthesized by Integrated DNA Technologies. DNA was end-repaired and A-tailed and sequencing adaptors were ligated using the Hyper Prep Kit (KAPA Biosystems) following the manufacturer's protocol. Adaptor-ligated DNA was bisulfite treated using the EpiTect Bisulfite Kit (Qiagen), modifying the manufacturer's protocol by extending the denaturation thermocycler step from 5 to 10 minutes at 99˚C. Adaptor-ligated bisulfite treated libraries were amplified 10-15 times using HiFi Uracil+ Polymerase (KAPA Biosystems) and library concentration was estimated using the KAPA quantification kit (KAPA Biosystems). Size was estimated using a high sensitivity DNA chip (Agilent Technologies). Libraries were sequenced using 50 bp single-end or paired-end reads on a HiSeq2500 by the Genome Technology Center at New York University (NYU).

Combined Bisulfite Restriction Analysis (COBRA) was performed similar to that previously described [288]. Briefly, high molecular weight DNA was bisulfite treated (see above) and 1 ng of bisulfite converted DNA was amplified 35-40 times using JumpStart *Taq* polymerase (Sigma) and bisulfite primers (Supplementary Table 6). Half of the amplified product was digested with *Taq*I (NEB) or *BstU*I (NEB) and the other half was mock digested as a control, prior to visualization on 1.5-2% agarose gels. DNA methylation was also quantitated by a qPCR approach where genomic DNA was aliquoted into three equal portions where one was mock digested to quantitate the total amount of DNA, one was digested with the methyl-sensitive restriction enzyme *Hpa*II to quantitate unmethylated DNA, and the final aliquot was digested with the methyl-insensitive isoschizomer *Msp*I as a negative control. Equal portions of each aliquot were subjected to qPCR and DNA methylation levels were quantitated as the ratio of *Hpa*II-

digested material to mock-digested material. Quantitation was based upon a standard curve of genomic DNA and all primers (**Supplementary Table 7**) were between 90% and 100% efficient.

DNA Methylation Bioinformatic Analyses

RRBS data were aligned to the *in silico* bisulfite converted genome (mm9) using Bismark [179] (v.0.13.1) with the Bowtie2 [312] option. Binary alignment map (BAM) files were parsed to derive DNA methylation calls that were collapsed to the CpG level using custom R scripts that made use of the 'Rsamtools' (v.1.22.0) and 'data.table' (v.1.9.6) packages (code available upon request). Data were compiled into datasets that included a sample specific coverage (minimum 10X coverage per sample) and a group specific coverage (minimum 10X coverage per group). Hierarchical clustering and principle component analyses were performed on sample specific data sets in a manner analogous to that described for the gene expression analysis except no normalization was performed and heatmaps were ordered by increasing DNA methylation from top to bottom. The distribution of DNA methylation values were assessed using the 'density' function in R/Bioconductor and represents the probability density function across sample specific coverage. Average methylation was determined based on sample coverage and differences were determined by Welch's *t*-test. B cell, plasmablast and plasma cell differentially methylated CpG loci (DML) were identified using Dispersion Shrinkage for Sequencing (DSS) [284] and division specific DML were determined using the general experimental design version of DSS [313]. DSS was applied to group level data and CpG loci that had an FDR $\leq$0.01 with a minimum change of 20% in DNA methylation were considered significant. Contiguous DML were defined as two or more DML that were

located adjacent to each other on the genome relative to assay group coverage. Significance of DML contiguity was assessed by permutation analysis. This involved randomly permuting DML 1,000 times and calculating the percent of permuted DML that occurred in contiguous regions for each permutation. The *P*-value was determined by the number of times that the permuted value was equal to or greater than the actual value.

Overlap of DML with enhancer elements (methods described in Meta-analysis) was assessed using Fisher's exact test [314] implemented in R/Bioconductor. Transcription factor motifs enriched within 50 bp of DML were determined with HOMER software [259] (v.4.7.2) relative to RRBS assay coverage. Results with an FDR ≤ 0.05 were considered significant. Motif position weight matrices were clustered using the "PWMSimilarity" function of the "TFBSTools" package [315] (v.1.8.2) with a minimum overlap of 6 nucleotides. Data were clustered using hierarchical clustering as described above. Code is available upon request.

Correlation of Gene Expression and DNA Methylation

To analyze DNA methylation and gene expression correlation, CpG loci were annotated to the closest UCSC mm9 Known Gene [304] transcript using custom R / Bioconductor code (available upon request). CpG loci that were within 100 kb of a transcript were assigned to the closest gene. Subsequently, the expression fold-change for each gene was plotted by the change in absolute DNA methylation. Significance of inverse correlation was assessed using Fisher's exact test [314] (implemented in R) to determine if more DML-DEG correlations were negatively associated than expected by chance.

Meta-Analysis

Analysis of histone modifications utilized previously described chromatin immunoprecipitation sequencing (ChIP-seq) experiments performed on primary B cells [227,259,292,316,317]. ChIP-seq data were obtained from Gene Expression Omnibus experiments GSE30859 [259], GSE38046 [317], GSE51336 [227], GSE42706 [316], and GSE51011 [292]; specific data sets are listed in Supplementary Table 4. Data were uniformly aligned and processed to the mouse genome (mm9) using Bowtie2 [312] (v.2.1.0). ChIP-seq fragment size for each data set was calculated using the "chip-seq" (v.1.20.0) package in R/Bioconductor based on the SISSER method previously described [318]. Enriched regions for those published by the ENCODE project [227] were downloaded from the UCSC genome browser (http://genome.ucsc.edu/cgi-bin/hgFileSearch?db=mm9). Enriched regions for other studies were determined using MACS software (v1.4) [319]. Enhancers were determined for spleen, CH12 cells, thymus, and whole brain H3K4me1, H3K27ac, and H3K4me3 data by taking overlapping H3K4me1 and H3K27ac regions that did not overlap a region enriched for H3K4me3. Odds ratio and significance of overlap with DML were determined using Fisher's exact test.

Chapter 3 : *De novo* DNA methylation limits B cell proliferation and differentiation

Abstract

B cells provide humoral immunity through the differentiation of antibody secreting plasma cells, a process that requires cell division and is linked to DNA hypomethylation that can augment plasma cell formation. Conversely, accumulation of DNA methylation in B cell differentiation is less apparent. To determine the role of *de novo* DNA methylation in B cell differentiation, the *de novo* DNA methyltransferases, Dnmt3a and Dnmt3b, were deleted in B cells resulting in phenotypically normal B cell development in the bone marrow, spleen and lymph nodes. However, upon immunologic challenge mice deficient for Dnmt3a and Dnmt3b (Dnmt3ab-deficient) accumulated more antigen-specific B cells and bone marrow chimeras showed this was cell-autonomous. Additionally, a fivefold increase in splenic and bone marrow plasma cells was observed. Molecular analysis revealed that Dnmt3ab-deficient bone marrow plasma cells failed to repress gene expression to the same level as their Dnmt3ab-sufficient counterparts. This was coupled with a failure of Dnmt3ab-deficient germinal center B cells and plasma cells to gain and/or maintain DNA methylation at several thousand loci that were clustered in enhancers of genes that function in B cell activation and homing. Analysis of chromatin accessibility showed Dnmt3ab-deficient plasma cells had increased accessibility at several genes involved in both hematopoiesis and B cell differentiation. These data show that *de novo* DNA methylation limits B cell activation, proliferation and differentiation, and support a model whereby DNA methylation represses the aberrant transcription of genes silenced in B cell differentiation to maintain plasma cell homeostasis.

Introduction

Appropriate regulation of B cell function is essential to humoral immunity. Transcriptional control of B cell differentiation determines both the fate [129,138,145] and magnitude of response [272,320], in part, through epigenetic mechanisms [142,272] including DNA methylation [273]. Maintenance of DNA methylation through cell division is mediated by *Dnmt1*, a process essential for mammalian survival [170], hematopoiesis [237,238] and lymphocyte maturation [321,322] and differentiation [273,321,323]. Deposition of DNA methylation by the *de novo* DNA methyltransferases Dnmt3a and Dnmt3b, is also required for mammalian development [172], controls hematopoietic self-renewal [184,243], and Dnmt3a mutations are associated with hematopoietic malignancies [246,248]. Recent evidence has indicated that during hematopoiesis, common lymphoid progenitors undergo a global gain in DNA methylation [241] that is propagated to mature B lymphocytes [279], but that upon activation and differentiation, B cells undergo dramatic DNA hypomethylation [273,278,279]. These data question the importance of *de novo* DNA methylation after B cell maturation, and may suggest a model whereby accumulation of DNA methylation is no longer required for B cell terminal differentiation.

To test the hypothesis that *de novo* DNA methylation is important for mature B cell function we studied B cell development, activation and differentiation in the absence of the *de novo* DNA methyltransferases: *Dnmt3a* and *Dnmt3b*. B cell development was phenotypically normal in the bone marrow, spleen and lymph nodes, and mature follicular B cells showed very few molecular defects. Upon antigenic stimulation, mice deficient for both Dnmt3a and Dnmt3b, had an increased frequency of germinal center B cells as well as antigen-specific B cell responses. This phenomenon was cell-

autonomous as bone marrow chimeras recapitulated this result and revealed a 5-fold increase in splenic and bone marrow plasma cells. Molecular characterization of antigen-specific germinal center B cells and bone marrow plasma cells revealed a failure to repress gene expression at proliferation and migration genes coupled with a failure to gain DNA methylation at several thousand CpG loci. These genomic regions were associated with enhancers and displayed increased accessibility in mice that lack both Dnmt3a and Dnmt3b. These data reveal the function of *de novo* DNA methylation in B cell responses and identify the molecular targets that maintain B cell homeostasis.

B cell development proceeds normally in the absence of the *de novo* DNA methyltransferases

To conditionally delete both *de novo* DNA methyltransferases in B cells, mice that contain conserved exons of *Dnmt3a* [324] and *Dnmt3b* [325] flanked by *loxP* sites (floxed), which have previously been shown to ablate Dnmt3a and Dnmt3b function, were bred to mice expressing the B cell specific *Cd19* *cre*-recombinase [39] (*Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$). Resultant mice showed genomic rearrangement of the floxed regions in *Dnmt3a* and *Dnmt3b* specifically in B cells whereas Dnmt3ab-sufficient (*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$) littermate controls did not (**Fig. 3-1**). Developing B cells in Dnmt3ab-deficient (*Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$) mice were analyzed by flow cytometry revealing similar numbers of bone marrow B220^{+}CD43^{-} B cells (**Fig. 3-2a)**. Closer examination also revealed similar frequency of bone marrow B220^{mid}CD43^{-}IgM^{-} pre-B cells, B220^{mid}CD43^{-}IgM^{+} immature B cells and B220^{hi}CD43^{-}IgM$^{+/lo}$ mature B cells [34] (**Fig. 3-2b**). Likewise, Dnmt3ab-deficient and Dnmt3ab-sufficient mice had similar numbers of splenic B220^{+}CD43^{-} B cells as well as B220^{+}CD43^{-}CD23^{+}CD21$^{int/-}$ follicular

B cells and B220^{+}CD43^{-}CD23^{-}CD21^{hi} marginal zone B cells [326] (**Fig. 3-2c-d**). We conclude that there are no major defects in B cell development when *Dnmt3a* and *Dnmt3b* are abrogated.

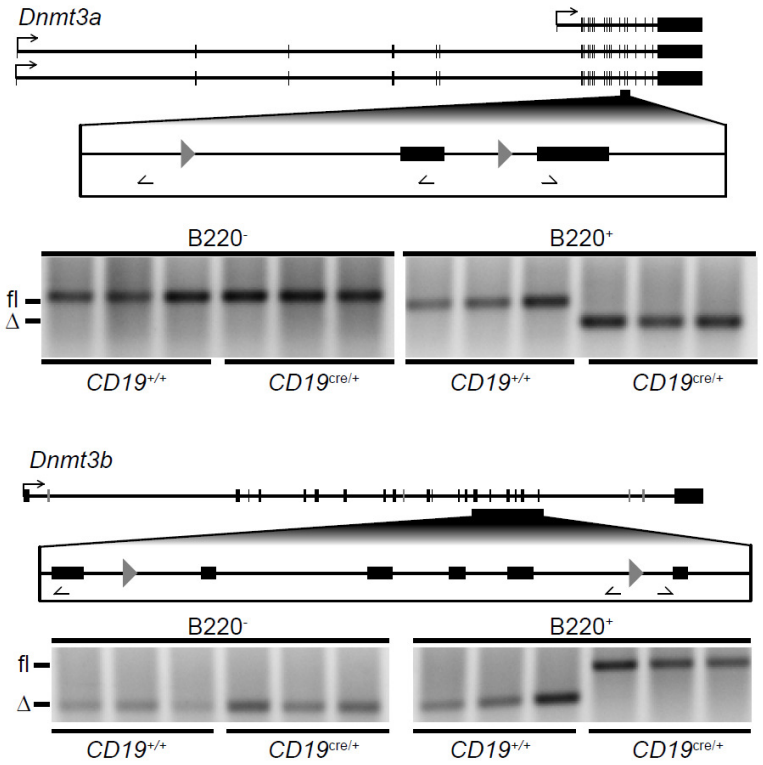


Figure 3-1. Conditional deletion of *Dnmt3a* and *Dnmt3b* using *Cd19*$^{cre/+}$. Schematic of *Dnmt3a* (top) and *Dnmt3b* (bottom) with region flanked by *loxP* sites (gray triangles) enlarged. PCR primers (black arrows) were used to amplify the region and measure genome rearrangement in bone marrow B cells and non B cells. Data are from one experiment with 3 *Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ mice and 3 *Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ littermate controls.

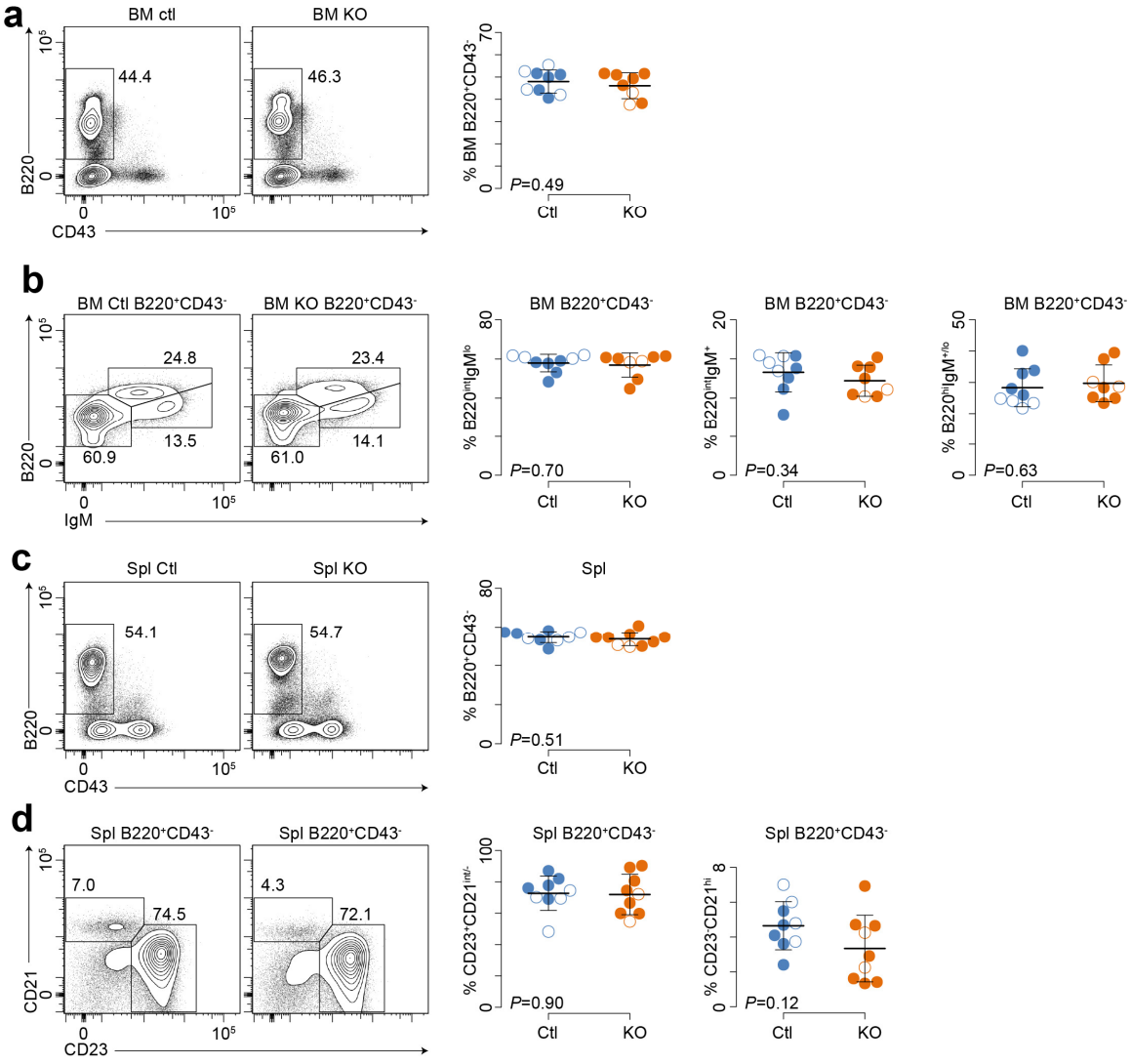


Figure 3-2. Normal proportions of bone marrow and splenic B cell subsets in *Cd19*^{cre}*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ (KO) mice. **(a)** B220 and CD43 expression in the bone marrow (BM) from representative control (Ctl) and KO mice. Data are summarized showing the frequency of B220^{+}CD43^{-} B cells in the BM (right). **(b)** IgM and B220 expression on B220^{+}CD43^{-} gated bone marrow B cells shown in part **a**. The three gates represent B220^{int}IgM^{-} pre-B cells, immature B220^{int}IgM^{+} B cells and B220^{hi}IgM$^{+/lo}$ mature B cells. **(c)** B220 and CD43 expression in the spleen (Spl) from in control (Ctl) and knock-out (KO) mice. **(d)** CD23 and CD21 expression on splenic B220^{+}CD43^{-} B cells. Quantitation of CD23^{+}CD21$^{int/-}$follicular and CD23^{-}CD21^{hi} marginal zone B cells are shown on the right. All data are lymphocyte size gated and CD11b^{-}. *P*-values were calculated using a two-sided *t*-test. Data are from two experiments with 8 and 10 mice per experiment, female mice are denoted by open circles and male by closed circles. (mean and s.d.).

Dnmt3a and *Dnmt3b* control germinal center B cell expansion and plasma cell differentiation

Given that B cell development appeared phenotypically normal in Dnmt3ab-deficient mice, these mice were suitable models to test the role of de novo DNA methylation in B cell differentiation. Challenge of Dnmt3ab-deficient and Dnmt3ab-sufficient mice with 50 μg of lipopolysaccharide (LPS) *i.v.* resulted in a similar splenic $CD138^+$ plasmablast responses 3 days post-challenge (**data not shown**). This was not surprising given recent data indicating more than 99% of DNA methylation changes in response to LPS are hypomethylation events (see Chapter 2). Studies of human steady-state plasma cells are consistent with a global loss of DNA methylation but reported a greater fraction of hypermethylation events [278,279]. To determine if Dnmt3a and Dnmt3b might function over longer time periods in response to more complex antigens, Dnmt3ab-deficient and Dnmt3ab-sufficient mice were immunized subcutaneously with phycoerythrin emulsified in complete Freud's adjuvant (PE-CFA). PE-CFA immunization allows tracking of the PE-specific B cell response [327,328], which predominately occurs in the draining inguinal and periaortic lymph nodes [329]. Analysis of these lymph nodes in unimmunized mice revealed that ~0.8% of $B220^+$ B cells were $GL7^+Fas^+$ germinal center B cells in both Dnmt3ab-deficient and Dnmt3ab-sufficient mice (**Fig. 3-3a**), however 30 days after PE-CFA immunization this number increased twofold in control animals but fourfold in Dnmt3ab-deficient mice (**Fig. 3-3a**). PE-binding of lymph node $B220^+$ B cells revealed that unimmunized Dnmt3ab-deficient and Dnmt3ab-sufficient mice had very low and equivalent levels of $B220^+$PE-specific B cells (**Fig. 3-3b**). However, 30 days after PE-CFA immunization Dnmt3ab-sufficient mice had a 9-fold expansion of $B220^+$PE-binding

B cells to 0.37% compared Dnmt3ab-deficient mice that had more than 25-fold increase to 0.99% (**Fig. 3b**), indicating that Dnmt3a and Dnmt3b help limit B cell activation.

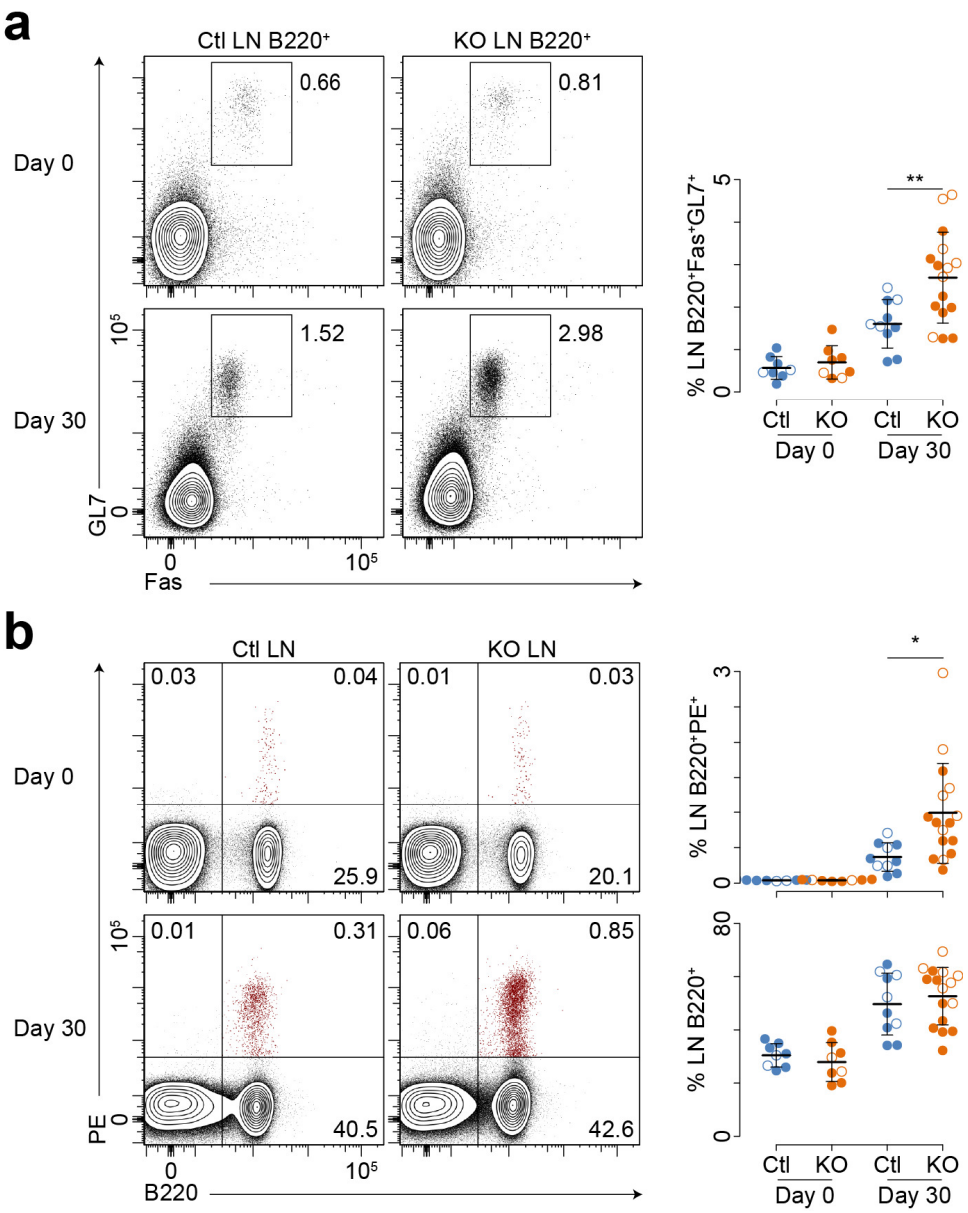


Figure 3-3: Increased germinal center B cell expansion in *Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ mice. **(a)** Fas and GL7 expression on inguinal and periaortic lymph node (LN) B220^{+} B cells on day 0 30 of PE-CFA immunization in *Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ mice (KO) and littermate controls (Ctl). **(b)** B220 expression and PE-binding of LN lymphocytes on Day 0 and Day 30 after PE-CFA challenge. All data are lymphocyte size gated and CD11b^{-}. *P*-values are calculated using a two-sided *t*-test (**P* <0.05, ***P* <0.01). Data are from two day 0 experiments with 8 mice each or

two day 30 experiments with 9 mice and 17 mice where open circles represent females and closed circles males.

To test if *Dnmt3a* and *Dnmt3b* controlled germinal center B cell responses through a cell-autonomous mechanism, bone marrow chimeras were made by mixing at a 1:1 ratio $CD45.2^+$ Dnmt3ab-deficient and $CD45.2^+CD45.1^+$ Dnmt3ab-sufficient bone marrow cells and transferring into lethally irradiated $CD45.1^+$ C57B/6J hosts. Six to eight weeks after bone marrow reconstitution mice were challenged with PE-CFA and analyzed 30 days later. Chimeric mice had approximately equal numbers of $CD45.2^+$ Dnmt3ab-deficient and $CD45.1^+CD45.2^+$ Dnmt3ab-sufficient cells in the lymph node (**Fig. 3-4a**). Analysis of B220 expression and PE-binding in the lymph node showed that $CD45.2^+$ cells from Dnmt3ab-deficient had significantly more PE-specific $B220^+$ B cells than those from the $CD45.1^+CD45.2^+$ Dnmt3ab-sufficient mice (**Fig. 3-4b**). More of the $CD45.2^+$ Dnmt3ab-deficient PE-specific $B220^+$ B cells were $GL7^+Fas^+$ germinal center B cells than their Dnmt3ab-sufficient counterparts (**Fig. 3-4c**). Analysis of $CD138^+$ differentiating plasma cells in the spleen revealed a remarkable increase in $CD138^+$ plasma cell frequency in Dnmt3ab-deficient cells as compared to control cells (**Fig. 3-4d,e**). These data show that *Dnmt3a* and *Dnmt3b* control B cell proliferation and differentiation through a cell-autonomous mechanism.

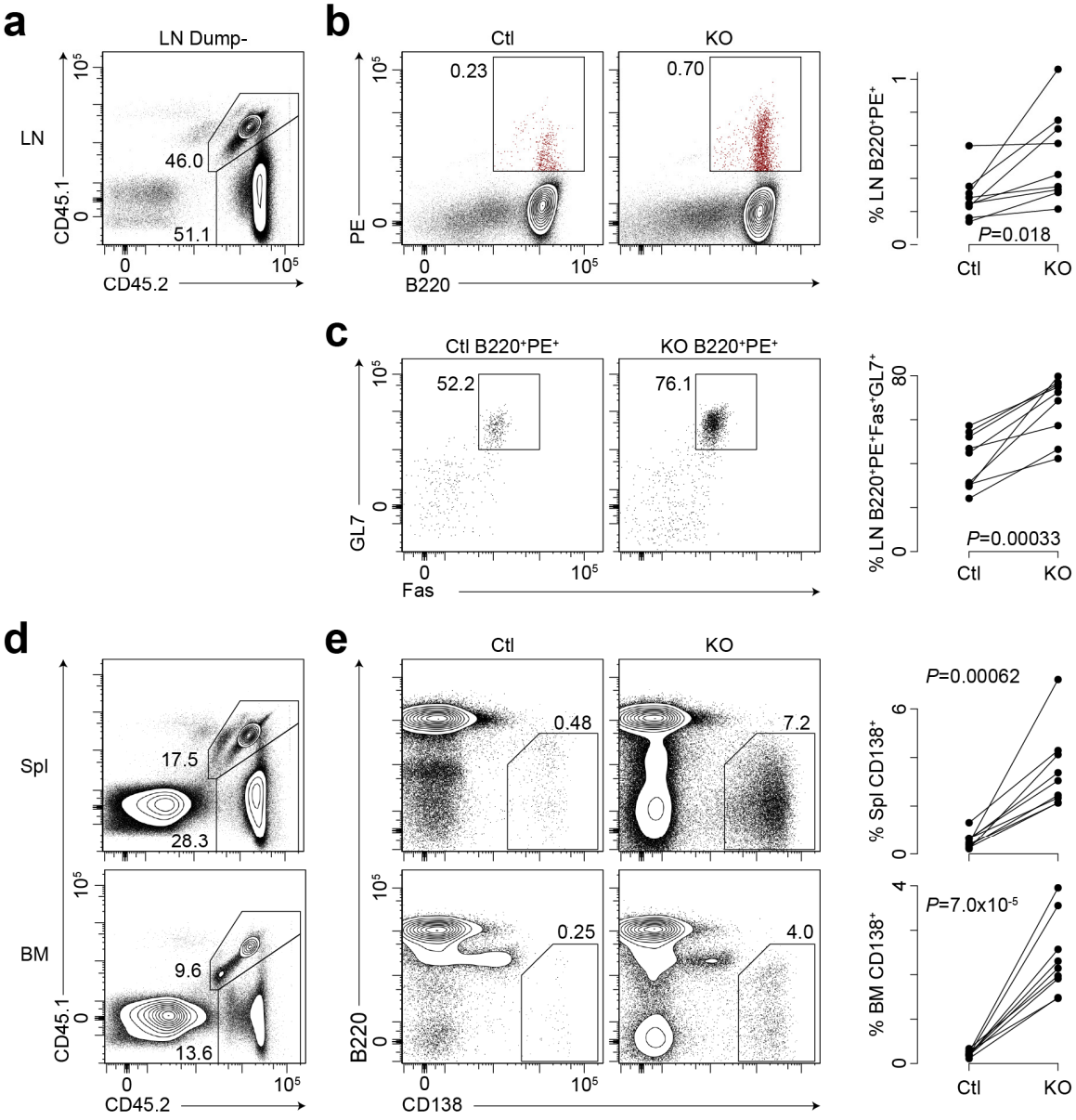


Figure 3-4: Cell autonomous control of B cell proliferation and differentiation. **(a)** CD45.1 and CD45.2 expression in the inguinal and periaortic lymph nodes (LN) of BM chimera mice made from CD45.2 *Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ (KO) and CD45.2 / CD45.1 *Dnmt3a*$^{fl/+}$*Dnmt3b*$^{fl/+}$ (Ctl) bone marrow cells **(b)** B220 expression and PE binding of CD45.1^{+}CD45.2^{+} Ctl and CD45.2^{+} KO cells in the LN 30 days after PE-CFA challenge. **(c)** Fas and GL7 expression on B220^{+}PE^{+} LN cells in **b**. **(d)** CD45.1 and CD45.2 expression in spleen (Spl) and bone marrow (BM) chimeric mice. (**e)** CD138 and B220 expression on Ctl and KO cells in part **d**. All data are lymphocyte size gated and CD11b^{-}CD11c^{-}F4/80^{-}Thy1.2^{-}. *P*-values are calculated using a paired *t*-test. Data represent 2 experiments with 4 and 5 mice per experiment.

***Dnmt3a* and *Dnmt3b* control repression of proliferation and homing genes.**

To determine the molecular targets of Dnmt3a and Dnmt3b in B cell differentiation, naïve B cells ($B220^+GL7^-Fas^-$) from the lymph node were purified from unchallenged mice and PE-specific germinal center B cells ($B220^+PE^+GL7^+Fas^+$) from mice 30 days post-challenge with PE-CFA (**Fig. 3-5a**). $CD138^+$ bone marrow plasma cells (BMPCs) were also isolated (**Fig. 3-5b**). RNA and DNA were extracted from the above populations and used for analysis by RNA-seq and reduced representation bisulfite sequencing (RRBS) [180]. Hierarchical clustering of RNA-seq data grouped genes into cell type specific expression patterns (**Fig. 3-5c**). Gene expression changes between naïve B cells and PE-specific germinal center B cells revealed similar changes in expression in both Dnmt3ab-sufficient and Dnmt3ab-deficient mice (**Fig. 3-5d**). Comparison of gene expression changes between naïve B cells and BMPCs showed similar sets of genes upregulated in BMPCs whereas more genes were downregulated in BMPCs of Dnmt3ab-sufficient than Dnmt3ab-deficient mice (**Fig. 3-5d**). Gene set enrichment analysis indicated that PE-specific germinal center B cells upregulated cell cycle and downregulated cell adhesion genes, while BMPCs upregulated ribosome genes and downregulated B cell specific genes (**Fig. 3-5e**). Direct comparison of differentially expressed genes (DEGs) between Dnmt3ab-sufficient and Dnmt3ab-deficient mice revealed very few differences in naïve B cells and PE-specific germinal center B cells but indicated 345 genes upregulated and 104 downregulated genes in Dnmt3ab-deficient BMPCs compared to those from Dnmt3ab-sufficient mice (**Fig. 3-5f**). A heatmap of the above DEGs suggested that the majority were downregulated in Dnmt3ab-sufficient but not Dnmt3ab-deficient BMPCs (**Fig. 3-5g**). This was confirmed by plotting the average

of Dnmt3ab-specific differentially expressed genes, which showed the highest expression in naïve B cells with progressive downregulation in PE-specific germinal center B cells, and further repressed in BMPCs of control animals as compared to KO (**Fig 3-5h**). Examples of these genes include the homing receptor *S1pr4* and the cell cycle regulator *Bmyc*, which is a member of the Myc/Mad/Max family of transcription factors (**Fig. 3-5i**). These data suggest that Dnmt3a and Dnmt3b help to repress genes as naïve B cells differentiate into bone marrow plasma cells.

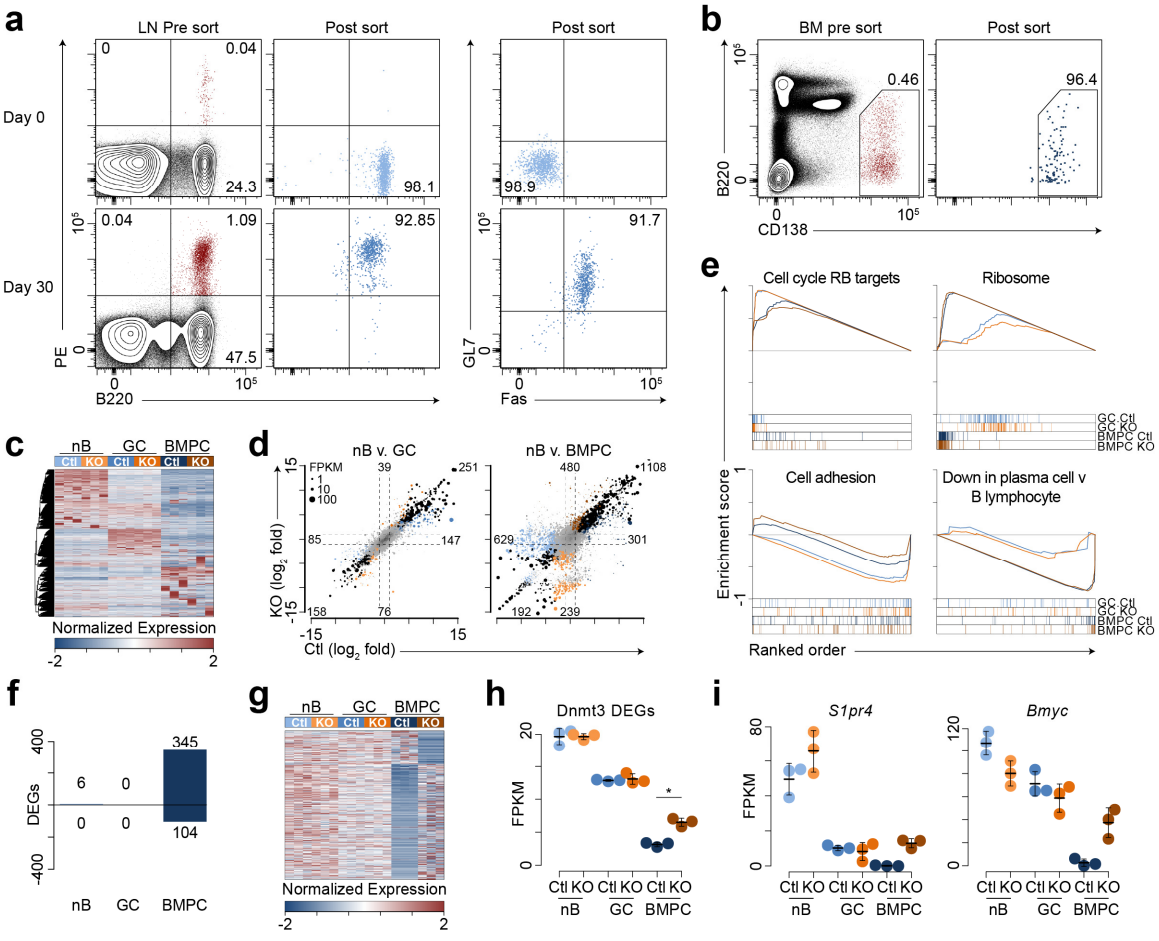


Figure 3-5: Failure to repress gene expression without Dnmt3a and Dnmt3b. **(a)** B220 expression and PE-binding in lymph nodes of representative mice (left) and post sort populations (right). **(b)** B220 and CD138 expression in bone marrow before (left) and post sort. **(c)** Hierarchical clustering of RNA-seq expression data in lymph node $B220^{+}GL7^{-}Fas^{-}$ naïve B cells (nB), $B220^{+}PE^{+}GL7^{+}Fas^{+}$ germinal center B cells (GC), and bone marrow plasma cells (BMPC)

for both *Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ (KO) and littermate control (Ctl) mice. **(d)** Gene expression fold changes between nB and GC cells (left) and nB and BMPC (right) in both Ctl and KO mice. **(e)** Gene set enrichment analysis of cell cycle, ribosome, cell adhesion and plasma cell genes. **(f)** Barplot of differentially expressed genes (DEGs) between KO and Ctl cell types. **(g)** Heatmap of Dnmt3a and Dnmt3b specific DEGs. **(h)** Average expression of Dnmt3a and Dnmt3b specific DEGs. **(i)** Examples of DEGs.

De novo DNA methylation at genes silenced in B cell differentiation

DNA methylation state at 1,736,270 CpGs was derived by RRBS. Data were plotted in a heatmap (**Fig. 3-6a**, left), which suggested that PE-specific germinal center B cells, and BMPCs had reduced levels of DNA methylation as compared to naïve B cells. Comparison of the average DNA methylation level between samples showed this was indeed the case (**Fig. 3-6a**). Principle components analysis (PCA) of DNA methylation data showed principle components one through three separated samples by cell type, but the fourth principle component separated Dnmt3ab-deficient from Dnmt3ab-sufficient samples (**Fig. 3-6b**), indicating that while the majority of variation was cell type specific, there was a reproducible component attributed to genetic ablation of *Dnmt3a* and *Dnmt3b*.

Differentially methylated loci (DML) between the naïve B cells, PE-specific germinal center B cells and BMPCs were determined, as well as those between Dnmt3ab-deficient and Dnmt3ab-sufficient samples within a given cell type. Organization of DML using K-means clustering revealed that most changes occurred at highly methylated CpGs in naïve B cells with commensurate reductions in germinal center B cells and BMPCs (**Fig. 3-6c**; clusters 2 and 3). There was also a subset that underwent greater demethylation in PE-specific germinal center B cells (**Fig. 3-6c**, cluster 4), perhaps a result of the high levels of proliferation. There was a minority subset of differentially

methylated loci that had low levels of DNA methylation in all cell types, but tended to contain more DNA methylation in samples from Dnmt3ab-sufficient than Dnmt3ab-deficient mice (**Fig. 3-6c**; cluster 1). This was most pronounced in BMPCs (**Fig. 3-6c**, right). This class of differentially methylated loci contained a CpG content equivalent to assay coverage and more than that of other differentially methylated loci (**Fig. 3-6d**). Moreover, cluster 1 was more likely to overlap $H3K4^{me1+}H3K27^{ac+}H3K4^{me3-}$ active enhancers found in B cells, splenocytes, CH12 cells, as well as whole tissue homogenates from thymus, testis and brain of C57B/6J mice (**Fig. 3-6e**).

To directly compare DNA methylation differences attributed to *Dnmt3a* and *Dnmt3b*, differentially methylated loci between Dnmt3ab-sufficient than Dnmt3ab-deficient cells were determined for naïve B cells, PE-specific germinal center B cells and BMPCs. This revealed that there were only 150 differentially methylated loci in naïve B cells but that both PE-specific germinal center B cells and BMPCs had more than 2000 and 1400, respectively (**Fig 3-5f**). As expected, the majority (94%) of Dnmt3ab-specific differentially methylated loci were less methylated in Dnmt3ab-deficient cells (**Fig. 3-5f**). A heatmap of Dnmt3ab-specific differences (**Fig. 3-5g**) suggested that these were regions that gained DNA methylation in Dnmt3ab-sufficient PE-specific germinal center B cells and BMPCs but failed to do so in Dnmt3ab-deficient cells. Inspection of the average DNA methylation at these Dnmt3ab-specific loci showed that indeed, they gained DNA methylation in PE-specific germinal center B cells and BMPCs for Dnmt3ab-sufficient mice, but lost DNA methylation in the same cell types of Dnmt3ab-deficient mice (**Fig. 3-5h**). Analysis of DNA methylation at differentially expressed genes between Dnmt3ab-deficient and Dnmt3ab-sufficient mice showed that differentially expressed

genes in BMPCs had significantly more DNA methylation in Dnmt3ab-sufficient mice than those of Dnmt3ab-deficient mice (**Fig. 3-6i**). This corresponds with the majority of differentially expressed genes occurring in BMPCs (**Fig. 3-5f**). Examples of differentially methylated loci between naïve B cells, germinal center B cells and bone marrow plasma cells were found at genes downstream of B cell receptor signaling such as *Nfkbia* and *Pik3r2*, which encode IκB-α and p85-β, respectively whereas differentially methylated loci between Dnmt3ab-deficient and Dnmt3ab-sufficient mice could be found at regulators of growth and B cell identity such as the early growth response gene *Egr3* and the B cell identity factor Pax5 (**Fig. 3-6j**). These data showed that the majority of DNA methylation changes were losses in response to B cell activation and differentiation, but nonetheless there were regions targeted for *de novo* DNA methylation in germinal center B cells and bone marrow plasma cells that help control the gene expression program.

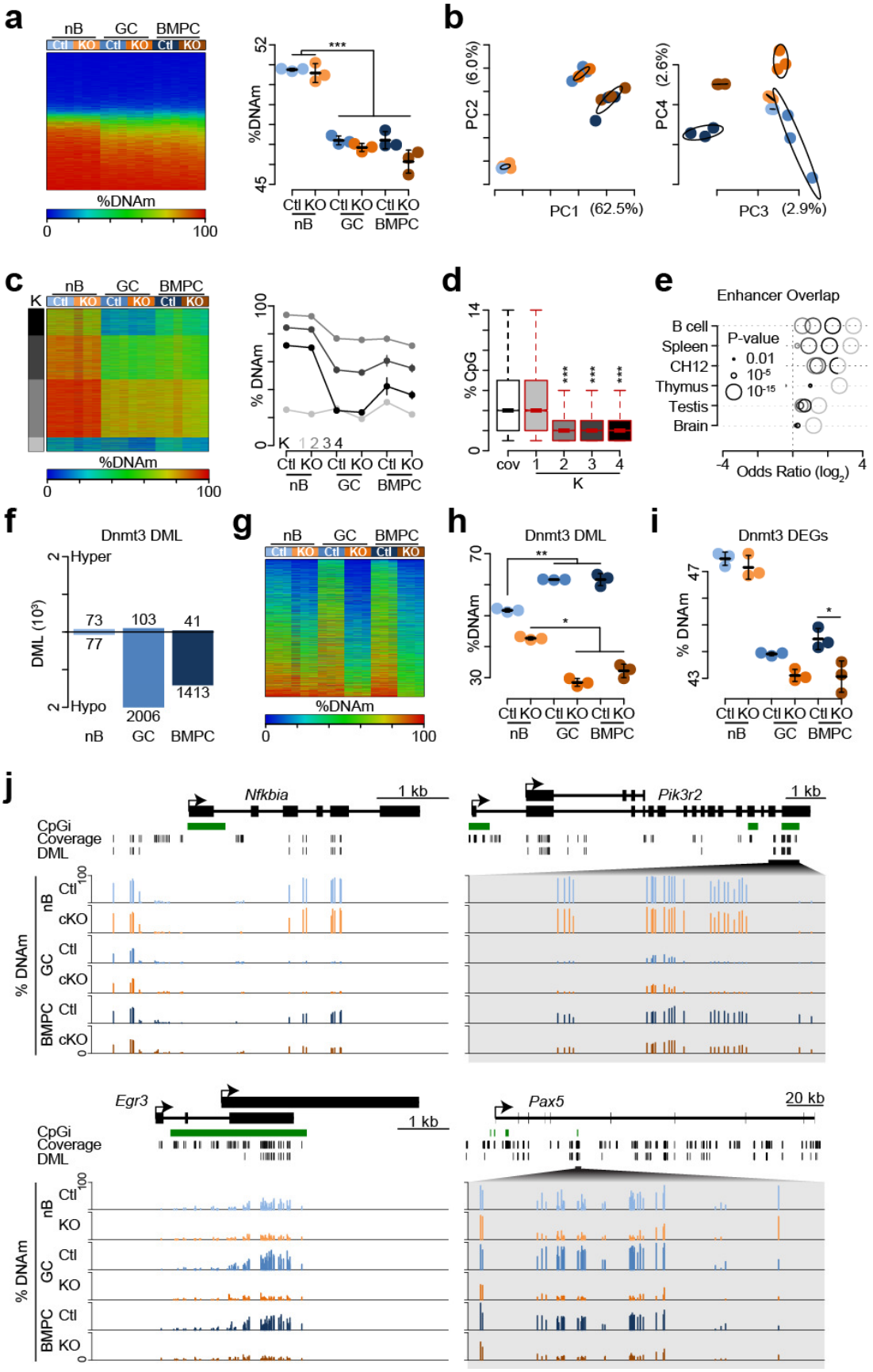

a
nB GC BMPC
Ctl KO Ctl KO Ctl KO
%DNAm
0 100
%DNAm
52 45
***
Ctl KO Ctl KO Ctl KO
nB GC BMPC
b
PC2 (6.0%)
PC1 (62.5%)
PC4 (2.6%)
PC3 (2.9%)
c
K
nB GC BMPC
Ctl KO Ctl KO Ctl KO
%DNAm
0 100
% DNAm
100 0
K 1 2 3 4
Ctl KO Ctl KO Ctl KO
nB GC BMPC
d
% CpG
14 0
***
cov 1 2 3 4
K
e
Enhancer Overlap
B cell
Spleen
CH12
Thymus
Testis
Brain
P-value
0.01
10-5
10-15
-4 4
Odds Ratio (log2)
f
Dnmt3 DML
Hyper
Hypo
DML (10³)
2 2
73 103 41
77 2006 1413
nB GC BMPC
g
nB GC BMPC
Ctl KO Ctl KO Ctl KO
%DNAm
0 100
h
Dnmt3 DML
%DNAm
70 30
**
*
Ctl KO Ctl KO Ctl KO
nB GC BMPC
i
Dnmt3 DEGs
% DNAm
47 43
*
Ctl KO Ctl KO Ctl KO
nB GC BMPC
j
Nfkbia
1 kb
Pik3r2
1 kb
CpGi
Coverage
DML
% DNAm
nB Ctl cKO
GC Ctl cKO
BMPC Ctl cKO
100 0
Egr3
1 kb
Pax5
20 kb
CpGi
Coverage
DML
% DNAm
nB Ctl KO
GC Ctl KO
BMPC Ctl KO
100 0

Figure 3-6. *De novo* DNA methylation at repressed genes in B cell differentiation. **(a)** Heatmap of DNA methylation in lymph node $B220^{+}GL7^{-}Fas^{-}$ naïve B cells (nB), $B220^{+}PE^{+}GL7^{+}Fas^{+}$ germinal center B cells (GC) and $CD138^{+}$ bone marrow plasma cells (BMPC) in $Cd19^{cre/+}Dnmt3a^{fl/fl}Dnmt3b^{fl/fl}$ (KO) and littermate control (Ctl) mice (left) and average DNA methylation determined by reduced representation bisulfite sequencing (RRBS; right). **(b)** principle components analysis of RRBS data. The percentage in parenthesis is the proportion of variation explained by each component. **(c)** K-means clustering of all differentially methylated loci (DML; left) and average DNAm level for each K-cluster (right). **(d)** Average CpG content of different K-clusters in **c**. **(e)** Overlap of DML in K-clusters with enhancers of the designated cell type. **(f)** Number of DML between KO and control mice. **(g)**. Heatmap of DNA methylation for DML in part **f**. **(h)** Average DNA methylation for DML in part **f**. **(i)** Average DNA methylation for differentially expressed genes in **Fig. 3-5f**. **(j)** Examples of differentially methylated loci between naïve B cells, germinal center B cells and BMPCs (top) as well as between KO and littermate controls (bottom).

Dnmt3a and *Dnmt3b* repress enhancer accessibility in B cell differentiation

Analysis of DNA methylation changes in Dnmt3ab-deficient mice indicated that these changes were preferentially occurring at enhancers of transcription. To determine if Dnmt3a and Dnmt3b help silence enhancers of transcription and more generally to understand the epigenetic programming defect that leads to increased plasma cell accumulation, ATAC-seq was performed on naïve B cells, PE-specific germinal center B cells and BMPCs for both Dnmt3ab-deficient and Dnmt3ab-sufficient mice. Hierarchical clustering of accessible regions segregated samples by cell type indicating chromatin accessibility reflected cell type specific programming (**Fig. 3-7a**). Accessibility changes between naïve B cells and PE-specific germinal center B cells or bone marrow plasma cells indicated that both Dnmt3ab-deficient and Dnmt3ab-sufficient cells had similar accessible genomes (**Fig 3-7b.**). Stratification of differential accessible regions by K-means clustering resulted in regions uniquely regulated between the cell types (**Fig. 3-7c**). Examples of these differences could be found at lineage markers and transcription factors such as *Cxcr5*, *Ciita*, *Aicda*, *Il10*, *Cd28*, and *Ccr9* (**Fig. 3-7d**). Accessibility

differences between Dnmt3ab-deficient and Dnmt3ab-sufficient naïve B cells, PE-specific germinal center B cells and BMPCs were determined (**Fig. 3-7e,f**), revealing no significant differences in naïve B cells or PE-specific germinal center B cells, but that BMPCs had 338 more accessible regions in Dnmt3ab-deficient BMPCs and 129 less accessible regions. Accessible increases in Dnmt3ab-deficient BMPCs included regions near genes involved in hematopoietic development, signal transduction, and leukocyte differentiation and activation. Examples of genes included that had regions with increased accessibility included *Cxcr4*, *Prdm1*, and *Card11* (**Fig. 3-7g**). Cumulatively, these data show exquisite epigenetic regulation between naïve B cells, PE-specific germinal center B cells and BMPCs, and suggest that without Dnmt3a and Dnmt3b regions gain aberrant accessibility in BMPCs.

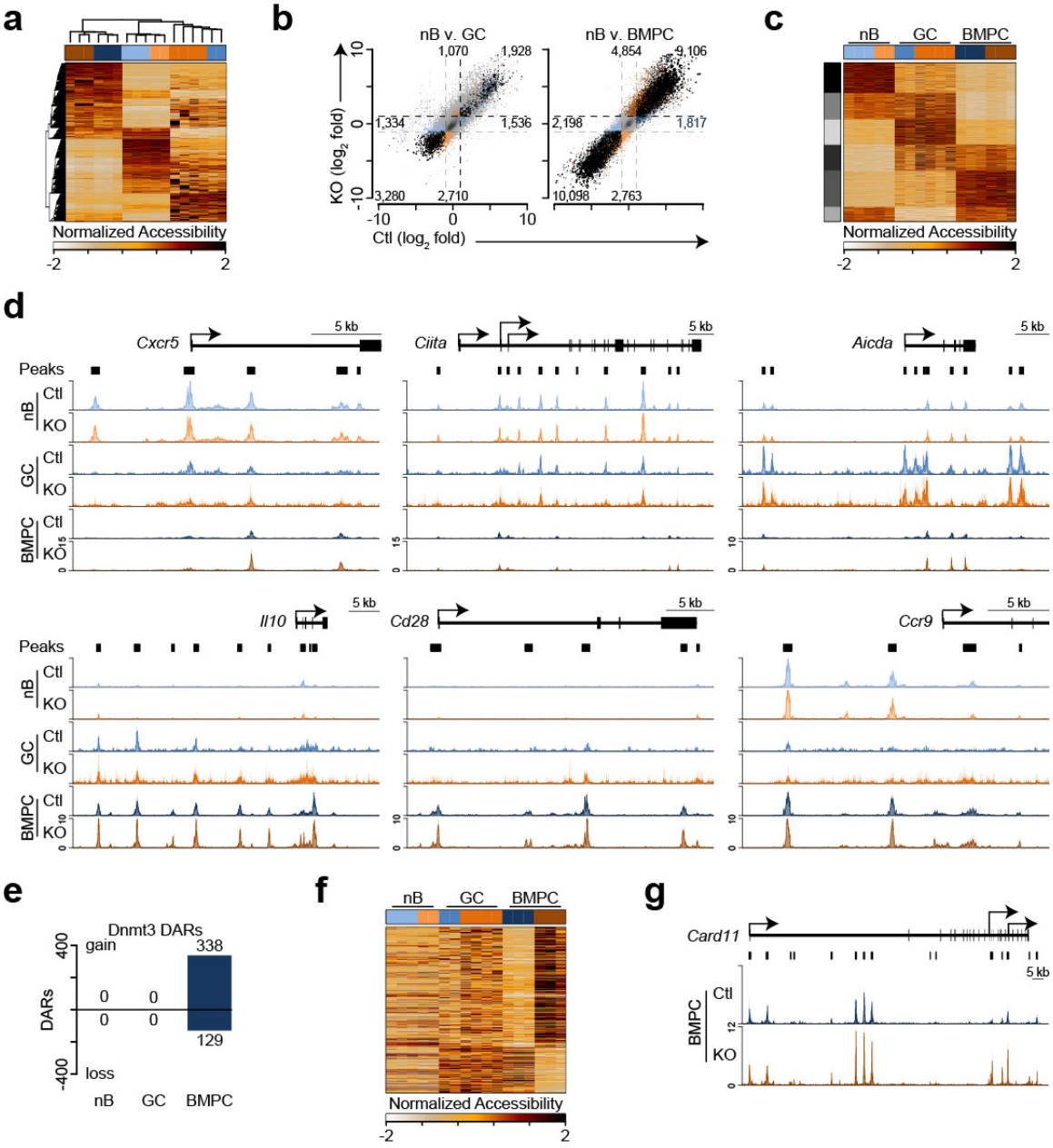


Figure 3-7. Increased chromatin accessibility in *Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ CD138^{+} bone marrow plasma cells (BMPCs). **(a)** Differentially accessible regions (DARs) determined by ATAC-seq between lymph node B220^{+}GL7^{-}Fas^{-} naïve B cells (nB) and B220^{+}PE^{+}GL7^{+}Fas^{+} germinal center B cells (GC) and between naïve B cells and BMPCs in *Cd19*$^{cre/+}$*Dnmt3a*$^{fl/fl}$*Dnmt3b*$^{fl/fl}$ (KO) and littermate control (Ctl) mice. **(b)** Chromatin accessibility changes between naïve B cells and germinal center B cells (left) and naïve B cells and BMPCs (right) for both KO and littermate mice. **(c)** Heatmap of K-means clustering of DARs. **(d)** Examples of cell types specific accessible regions. **(e)** Barplot of DARs between KO and control mice in the cell types specified. **(f)** Heatmap of Dnmt3 DARs in **e**. **(g)** Example of Dnmt3 specific DAR near the gene *Card11*.

Discussion

Deletion of Dnmt3a and Dnmt3b during B cell development results in mature B cells that are primarily normal with regards to phenotype and molecular programming. This was somewhat surprising given the importance of these enzymes for mammalian development [172], but may be explained by the fact that common lymphoid precursors gain copious amounts of DNA methylation [241], a developmental step that occurs just before Cd19 expression, which was used to delete Dnmt3a and Dnmt3b here. Supporting the notion that B cell maturation is primarily undisturbed as a result of Cd19 mediated deletion Dnmt3a and Dnmt3b, molecular analyses of Dnmt3ab-deficient mature naïve follicular B cells from the lymph nodes showed very few molecular changes in terms of RNA expression, DNA methylation and chromatin accessibility as compared to their Dnmt3ab-sufficient counterparts. Subsequently, we conclude that it is reasonable to use this genetic model to study the role of *de novo* DNA methylation in B cell activation and differentiation.

Immunization of Dnmt3ab-deficient mice resulted in increased B cell activation and differentiation, a phenomenon that was readily available over extended antigenic challenge (*i.e.* 30 days), but harder to discern over shorter time periods. This is consistent with Dnmt3a and Dnmt3b functioning to slowly accumulate DNA methylation and thereby changing the course of cell fate. Similar results were observed in the hematopoietic stem cell compartment, where deletion of Dnmt3a [243] or Dnmt3a and Dnmt3b [184] resulted in increased hematopoietic self-renewal and progressive defects in hematopoiesis that were most apparent after several weeks and multiple serial transplantations in competition with wild-type control cells. Likewise, analysis of bone

marrow chimeras here, showed that Dnmt3ab-deficient B cells had increased activation and accumulation of splenic and bone marrow plasma cells as compared to controls, a result not readily discernable without competing the Dnmt3ab-deficient cells against their wild-type counterparts.

Detailed molecular analysis of Dnmt3ab-deficient animals identified contiguous regions that failed to gain and/or maintain DNA methylation in both PE-specific germinal center B cells and bone marrow plasma cells. These regions were clustered around genes normally repressed during B cell differentiation such as *Pax5* and *Egr3*. Although in most cases repression appears to be largely independent of *de novo* DNA methylation, analysis of gene expression changes indicated bone marrow plasma cells from Dnmt3ab-deficient mice aberrantly overexpressed genes that contained more DNA methylation in Dnmt3ab-sufficient than Dnmt3ab-deficient mice. These Dnmt3ab-specific DNA methylation differences had a remarkable overlap with enhancers identified from publically available data sets as well as those determined using ATAC-seq in the same cell types. Cumulatively, this data show that Dnmt3a and Dnmt3b function to repress B cell activation and differentiation and identify the molecular targets that maintain plasma cell homeostasis.

Chapter 4 : Discussion

B cells are required for humoral immunity and vaccination and dysregulation of B cell differentiation can lead to hematological malignancies such as lymphoma and multiple myeloma, or contribute to autoimmune diseases such as systemic lupus erythematosus where auto-antibodies are pathologic. Understanding the cellular and molecular programming of how B cells mount normal immune responses as well as how these processes are disrupted in disease has been under intense study for the past 50 years. Invaluable to many discoveries of B cell biology has been the genetic mapping of B cell deficiencies [144,287] and the use of mouse model systems [297,330,331]. Resultantly, the transcription factor networks required for B cell development [19] and differentiation [20,123] are relatively well understood. What is less well understood is how transcription factor function is restricted by chromatin accessibility. For many years it has been known that epigenetic marks that control the accessibility of DNA to transcription factors, such as DNA methylation, are dysregulated in B cell malignancies such as multiple myeloma [332,333]. However, this is correlative and thus does not reveal whether these epigenetic modifications are a cause or consequence of general B cell dysregulation. More recently, several studies have conditionally removed epigenetic enzymes in hematopoiesis [184,243–245,334] or in B cells [272,273,335] revealing functional consequences for the specific epigenetic modifications that are deposited by such enzymes. Combining these genetic models with phenotypic and molecular analyses has allowed inferences about both the cellular stage and molecular mechanism at which these epigenetic modifications are most important for hematopoiesis, B cell differentiation, and humoral immunity.

Here, the role of DNA methylation was studied during B cell differentiation using cellular and molecular analyses and genetic models. These data showed that during acute differentiation in response to the model antigen LPS, B cells rapidly divide, differentiate, and undergo dramatic transcriptional changes coupled to DNA hypomethylation. These findings are congruent with other studies of B cell differentiation in that this process has consistently been reported to be tied to an overall loss in DNA methylation [273,278,279]. Here, 99.7% of LPS-induced DNA methylation changes were hypomethylation events which is more striking than previous reports that found 5-30% of total DNA methylation changes were gains in DNA methylation [273,278,279]. One likely explanation for this discrepancy is that the aforementioned studies measured steady-state levels of differentiated B cell subsets and are thus a reflection of homeostatic populations whereas the plasmablast and plasma cell populations studied here were actively differentiating and formed within the previous three days prior to LPS challenge. This may suggest that DNA methylation is dramatically lost during the proliferative burst known to be required for B cell differentiation[280], but that accumulation of some DNA methylation occurs as differentiated B cell subsets mature.

Investigation of cellular division during B cell differentiation using a dilution dye combined with flow cytometry revealed that B cells underwent many divisions prior to differentiation. This phenomenon has been consistently reported by many groups using *ex vivo* culture systems that mimic B cell differentiation [124,127,280]. Here, addition of 5-azacytidine, a biochemical inhibitor of DNA methylation, resulted in augmented plasma cell differentiation, suggesting that the presence of DNA methylation at certain loci inhibits plasma cell formation. Additionally, use of an *in vivo* model system where

dilution dye labelled B cells were transferred into a B cell deficient host, showed more proliferation occurred prior to plasma cell differentiation *in vivo* than what occurs *ex vivo*. Isolation and characterization of distinct cellular divisions during this process recapitulated the DNA hypomethylation events observed in wild-type mice and provided critical insight into the dynamics of the cellular division coupled DNA methylation and gene expression changes that occur during differentiation. These findings showed that very few DNA methylation changes occurred in cells that underwent only a few divisions, suggesting that DNA hypomethylation changes are replication linked and specific to the B cell populations that were rapidly dividing. Analysis of mRNA expression in the same cell populations revealed that 95% of genes were upregulated during cellular division and that the number of mRNAs increased from ~30,000 mRNAs per cell to over 160,000 mRNAs per cell. This is similar to a phenomenon that occurs in oncogenic cell types, known as transcriptional amplification [336]. Subsequent detailed analysis of the gene expression and DNA methylation changes showed remarkable correlation both in terms of the average DNA methylation and gene expression per cell type, as well as at specific genes where losses in DNA methylation corresponded with induction of expression in a division-specific manner. These data clearly explain that DNA methylation losses during B cell differentiation correspond with transcriptional activation, a relationship that was not realized by other studies that only measured relative gene expression and thus were not able to observe the global upregulation of almost all mRNAs [273,278,279]. Furthermore, analysis of the distinct cellular divisions captured transient expression of many mRNAs that was associated with division-specific losses in DNA methylation, and thus explains the perceived lack of correlation between

gene expression and DNA methylation changes at these loci in steady state plasma cells. These data therefore suggest that DNA hypomethylation represents a cis-regulatory history of transcriptional events during B cell differentiation.

DNA methylation is known to be removed by both cell replication-coupled passive demethylation and active removal through the oxidation of 5mC by Tet enzymes and subsequent removal by the DNA repair enzyme TDG [189–192]. Given the large amount of replication that occurs during B cell differentiation, it is likely that passive demethylation at least partially accounts for some of the observed DNA hypomethylation. However, DNA methylation losses were clearly specific to ~10% CpG loci assayed and still retained at the majority of highly methylated loci. This is somewhat counterintuitive to a strictly passive model that might predict reduced levels of DNA methylation throughout the genome. Additionally, regions that lost DNA methylation tended to co-localize with motifs of transcription factors such as Oct2 [148,149], Irf4 [145,147], Fra1 [293], and NF-κB [294] that are required for B cell differentiation. A similar phenomenon has been reported where DNA methylation was lost at regions occupied by transcription factors [193,194]. Mechanistically, it is still unclear whether this occurs because the transcription factors block maintenance methylation by Dnmt1 through replication or because these transcription factors recruit the Tet enzymes to facilitate oxidation of 5mC and subsequent removal and/or dilution by replication. In either case, it is likely that B cell demethylation occurs through replication-coupled passive demethylation and through Tet-mediated oxidation either through the failure to maintain oxidized forms of 5mC or through active removal. Supporting this, recent studies have found that Dnmt1 is

essential for germinal center B cell proliferation [273] and that deletion of Tet1 results in B cell malignancies [255].

It is now clear that the majority of DNA methylation reprogramming during B cell differentiation is hypomethylation associated with B cell activation and differentiation; however, the function of *de novo* DNA methylation during this process is less well understood. Recent studies have suggested regions that gain DNA methylation during B cell differentiation are also prone to H3K27me3 inhibition [279] and that B cell enhancers are aberrantly methylated in multiple myeloma [337], but the functional consequences of *de novo* DNA methylation during normal B cell differentiation are unclear. Here, deletion of both *de novo* DNA methyltransferases: *Dnmt3a* and *Dnmt3b* showed that B cell development was largely unaffected as a result of deletion, but that activated and differentiated B cell subsets were expanded in Dnmt3ab-deficient mice. This was most apparent upon extended immunologic challenge with T-cell dependent antigen. While this may appear incongruent with reports of these same mice having mostly normal B cell responses [338], it is important to note that the authors of the aforementioned study looked exclusively at earlier time points and focused mostly on the function Dnmt3a and Dnmt3b on murine gamma herpesvirus. Nonetheless, the author's data did suggest that Dnmt3ab-deficient mice had increased germinal center B cells, albeit not significantly so, in response to a series of immunogens. Had the authors followed the mice for a longer duration, it is possible that this defect would have become more pronounced. Supporting this notion, deletion of Dnmt3a and/or Dnmt3b in the hematopoietic stem cell compartment resulted in increased B cell expansion, an observation that was only made after serial transplantation of Dnmt3ab-deficient

hematopoietic stem cells [243,334]. Such a defect may exist in the mice analyzed here, which contain a B cell conditional deletion of Dnmt3a and Dnmt3b, but this defect is not readily observable. Furthermore, molecular analyses of mRNA expression, DNA methylation, and chromatin accessibility of naïve B cells from both Dnmt3ab-sufficient and Dnmt3ab-deficient mice suggested that B cell development was very similar in the absence of the *de novo* DNA methyltransferases.

In contrast to the minimal molecular changes found in Dnmt3ab-deficient naïve B cells, analysis of antigen-specific germinal center B cells 30 days after immunization indicated that several thousand CpG loci failed to gain DNA methylation in Dnmt3ab-deficient mice. These DNA methylation changes were clustered in contiguous regions and overlapped enhancers of many cell types indicating that Dnmt3a and Dnmt3b function to repress enhancers of B cell identity genes such as *Pax5* that would antagonize plasma cell function and homeostasis [133]. This is congruent with global gene expression changes that indicated Dnmt3ab-deficient plasma cells failed to repress genes during B cell differentiation as compared to their Dnmt3ab-sufficient counterparts. Finally, analysis of chromatin accessibility confirmed that Dnmt3ab-deficient plasma cells contained increased accessibility of several hundred regions showing that *de novo* DNA methylation controls repression of enhancers.

In summary, the data in this dissertation provides critical insight into the dynamics of DNA methylation reprogramming in B cell differentiation. The data support a model where upon B cell activation DNA methylation is dynamically lost in a cell division-dependent manner and occurs with transcriptional amplification. DNA hypomethylation was lost at regions that reflected the binding sites of active transcription factors, and thus

DNA hypomethylation represents a cis-regulatory history of B cell differentiation. However, over extended immunologic challenge, slow accumulation of DNA methylation was important for limiting B cell differentiation and plasma cell homeostasis by repressing enhancers of B cell genes. Cumulatively, these data show that the balance of DNA methylation helps control B cell differentiation versus homeostasis and provide insight into the functional consequences of aberrant DNA methylation in B cell malignancies such as cell lymphoma and multiple myeloma.

Bibliography

1. I. An Inquiry Into the Causes and Effects of the Variolae Vaccinae, Or Cow-Pox. 1798. Jenner, Edward. 1909-14. The Three Original Publications on Vaccination Against Smallpox. The Harvard Classics. Available at: http://www.bartleby.com/38/4/1.html. (Accessed: 22nd March 2016)
2. van Bebring, E. & Kitasato, S. The mechanism of immunity in animals to diphtheria and tetanus. (1890).
3. Paul Ehrlich - Nobel Lecture: Partial Cell Functions. Available at: http://www.nobelprize.org/nobel_prizes/medicine/laureates/1908/ehrlich-lecture.html. (Accessed: 13th March 2016)
4. Plasma Cellular Reaction and its Relation to the Formation of Antibodies *in vitro* : Abstract : Nature. Available at: http://www.nature.com.proxy.library.emory.edu/nature/journal/v159/n4041/abs/159499a0.html. (Accessed: 13th March 2016)
5. The Plasma Cellular Reaction and Its Relation to the Formation of Antibodies in Vitro. *J. Immunol.* **58,** 1–13 (1948).
6. Bruton, O. C. Agammaglobulinemia. *Pediatrics* **9,** 722–728 (1952).
7. Good, R. A. & Varco, R. L. A clinical and experimental study of agammaglobulinemia. *J. Lancet* **75,** 245–271 (1955).
8. Glick, B., Chang, T. S. & Jaap, R. G. The Bursa of Fabricius and Antibody Production. *Poult. Sci.* **35,** 224–225 (1956).
9. Cooper, M. D., Peterson, R. D. A. & Good, R. A. Delineation of the Thymic and Bursal Lymphoid Systems in the Chicken. *Nature* **205,** 143–146 (1965).

10. Köhler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256,** 495–497 (1975).
11. Research, C. for D. E. and. Therapeutic Biologic Applications (BLA) - Adalimumab Product Approval Information - Licensing Action 12/31/02. Available at: http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm080610.htm. (Accessed: 15th March 2016)
12. Research, C. for D. E. and. Therapeutic Biologic Applications (BLA) - Rituximab Product Approval Information - Licensing Action. Available at: http://www.fda.gov/drugs/developmentapprovalprocess/howdrugsaredevelopedandapproved/approvalapplications/therapeuticbiologicapplications/ucm093345.htm. (Accessed: 15th March 2016)
13. Research, C. for D. E. and. Therapeutic Biologic Applications (BLA) - Trastuzumab Product Approval Information - Licensing Action 9/25/98. Available at: http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/ApprovalApplications/TherapeuticBiologicApplications/ucm080591.htm. (Accessed: 15th March 2016)
14. Drug Approval Package: Avastin (Bevacizum) NDA #125085. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2004/STN-125085_Avastin.cfm. (Accessed: 15th March 2016)
15. Drug Approval Package: Brand Name (Generic Name) NDA #. Available at: http://www.accessdata.fda.gov/drugsatfda_docs/nda/2014/125554Orig1s000TOC.cfm. (Accessed: 15th March 2016)

16. Hozumi, N. & Tonegawa, S. Evidence for Somatic Rearrangement of Immunoglobulin Genes Coding for Variable and Constant Regions. *Proc. Natl. Acad. Sci. U. S. A.* **73,** 3628–3632 (1976).
17. Tonegawa, S. Somatic generation of antibody diversity. *Nature* **302,** 575–581 (1983).
18. Jacob, J., Kelsoe, G., Rajewsky, K. & Weiss, U. Intraclonal generation of antibody mutants in germinal centres. *Nature* **354,** 389–392 (1991).
19. Nutt, S. L. & Kee, B. L. The Transcriptional Regulation of B Cell Lineage Commitment. *Immunity* **26,** 715–725 (2007).
20. Nutt, S. L., Hodgkin, P. D., Tarlinton, D. M. & Corcoran, L. M. The generation of antibody-secreting plasma cells. *Nat. Rev. Immunol.* **15,** 160–171 (2015).
21. Kincade, P. W., Lawton, A. R., Bockman, D. E. & Cooper, M. D. Suppression of Immunoglobulin G Synthesis as a Result of Antibody-Mediated Suppression of Immunoglobulin M Synthesis in Chickens. *Proc. Natl. Acad. Sci.* **67,** 1918–1925 (1970).
22. Tipton, C. M. *et al.* Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat. Immunol.* **advance online publication,** (2015).
23. Alt, F. W., Zhang, Y., Meng, F.-L., Guo, C. & Schwer, B. Mechanisms of Programmed DNA Lesions and Genomic Instability in the Immune System. *Cell* **152,** 417–429 (2013).
24. Osmond, D. G. & Nossal, G. J. V. Differentiation of lymphocytes in mouse bone marrow: II. Kinetics of maturation and renewal of antiglobulin-binding cells studied by double labeling. *Cell. Immunol.* **13,** 132–145 (1974).

25. Ryser, J.-E. & Vassalli, P. Mouse Bone Marrow Lymphocytes and Their Differentiation. *J. Immunol.* **113,** 719–728 (1974).

26. Owen, J. J. T., Cooper, M. D. & Raff, M. C. In vitro generation of B lymphocytes in mouse foetal liver, a mammalian 'bursa equivalent'. *Nature* **249,** 361–363 (1974).

27. Scott, E. W., Simon, M. C., Anastasi, J. & Singh, H. Requirement of transcription factor PU.1 in the development of multiple hematopoietic lineages. *Science* **265,** 1573–1577 (1994).

28. Bain, G. *et al.* E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. *Cell* **79,** 885–892 (1994).

29. Sudo, T. *et al.* Expression and function of the interleukin 7 receptor in murine lymphocytes. *Proc. Natl. Acad. Sci.* **90,** 9125–9129 (1993).

30. Igarashi, H., Gregory, S. C., Yokota, T., Sakaguchi, N. & Kincade, P. W. Transcription from the RAG1 Locus Marks the Earliest Lymphocyte Progenitors in Bone Marrow. *Immunity* **17,** 117–130 (2002).

31. Oettinger, M. A., Schatz, D. G., Gorka, C. & Baltimore, D. RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. *Science* **248,** 1517–1523 (1990).

32. Schatz, D. G., Oettinger, M. A. & Baltimore, D. The V(D)J recombination activating gene, RAG-1. *Cell* **59,** 1035–1048 (1989).

33. Wesemann, D. R. *et al.* Microbial colonization influences early B-lineage development in the gut lamina propria. *Nature* **advance online publication,** (2013).

34. Hardy, R. R., Carmack, C. E., Shinton, S. A., Kemp, J. D. & Hayakawa, K. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J. Exp. Med.* **173,** 1213–1225 (1991).

35. Hagman, J., Belanger, C., Travis, A., Turck, C. W. & Grosschedl, R. Cloning and functional characterization of early B-cell factor, a regulator of lymphocyte-specific gene expression. *Genes Dev.* **7,** 760–773 (1993).

36. Ramsden, D. A., Baetz, K. & Wu, G. E. Conservation of sequence in recombination signal sequence spacers. *Nucleic Acids Res.* **22,** 1785–1796 (1994).

37. Nutt, S. L., Heavey, B., Rolink, A. G. & Busslinger, M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature* **401,** 556–562 (1999).

38. Kozmik, Z., Wang, S., Dörfler, P., Adams, B. & Busslinger, M. The promoter of the CD19 gene is a target for the B-cell-specific transcription factor BSAP. *Mol. Cell. Biol.* **12,** 2662–2672 (1992).

39. Rickert, R. C., Rajewsky, K. & Roes, J. Impairment of T-cell-dependent B-cell responses and B-l cell development in CD19-deficient mice. *Nature* **376,** 352–355 (1995).

40. Hardy, R. R., Kincade, P. W. & Dorshkind, K. The Protean Nature of Cells in the B Lymphocyte Lineage. *Immunity* **26,** 703–714 (2007).

41. Macallan, D. C. *et al.* B-cell kinetics in humans: rapid turnover of peripheral blood memory cells. *Blood* **105,** 3633–3640 (2005).

42. Kouzine, F. *et al.* Global Regulation of Promoter Melting in Naive Lymphocytes. *Cell* **153,** 988–999 (2013).

43. Allman, D. & Pillai, S. Peripheral B cell subsets. *Curr. Opin. Immunol.* **20,** 149–157 (2008).

44. Martin, F., Oliver, A. M. & Kearney, J. F. Marginal zone and B1 B cells unite in the early response against T-independent blood-borne particulate antigens. *Immunity* **14,** 617–629 (2001).

45. Cinamon, G., Zachariah, M. A., Lam, O. M., Foss, F. W. & Cyster, J. G. Follicular shuttling of marginal zone B cells facilitates antigen transport. *Nat. Immunol.* **9,** 54–62 (2008).

46. Cariappa, A., Chase, C., Liu, H., Russell, P. & Pillai, S. Naive recirculating B cells mature simultaneously in the spleen and bone marrow. *Blood* **109,** 2339–2345 (2007).

47. Kindred, B. & Shreffler, D. C. H-2 Dependence of Co-Operation Between T and B Cells in Vivo. *J. Immunol.* **109,** 940–943 (1972).

48. Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.-M. & Hoffmann, J. A. The Dorsoventral Regulatory Gene Cassette spätzle/Toll/cactus Controls the Potent Antifungal Response in Drosophila Adults. *Cell* **86,** 973–983 (1996).

49. Poltorak, A. *et al.* Defective LPS Signaling in C3H/HeJ and C57BL/10ScCr Mice: Mutations in Tlr4 Gene. *Science* **282,** 2085–2088 (1998).

50. Rock, F. L., Hardiman, G., Timans, J. C., Kastelein, R. A. & Bazan, J. F. A family of human receptors structurally related to Drosophila Toll. *Proc. Natl. Acad. Sci.* **95,** 588–593 (1998).

51. Medzhitov, R., Preston-Hurlburt, P. & Janeway, C. A. A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature* **388,** 394–397 (1997).

52. Hornung, V. *et al.* Quantitative Expression of Toll-Like Receptor 1–10 mRNA in Cellular Subsets of Human Peripheral Blood Mononuclear Cells and Sensitivity to CpG Oligodeoxynucleotides. *J. Immunol.* **168,** 4531–4537 (2002).

53. Dasari, P., Nicholson, I. C., Hodge, G., Dandie, G. W. & Zola, H. Expression of toll-like receptors on B lymphocytes. *Cell. Immunol.* **236,** 140–145 (2005).

54. Schwandner, R., Dziarski, R., Wesche, H., Rothe, M. & Kirschning, C. J. Peptidoglycan- and Lipoteichoic Acid-induced Cell Activation Is Mediated by Toll-like Receptor 2. *J. Biol. Chem.* **274,** 17406–17409 (1999).

55. Takeuchi, O. *et al.* Differential Roles of TLR2 and TLR4 in Recognition of Gram-Negative and Gram-Positive Bacterial Cell Wall Components. *Immunity* **11,** 443–451 (1999).

56. Hayashi, F. *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410,** 1099–1103 (2001).

57. Alexopoulou, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. *Nature* **413,** 732–738 (2001).

58. Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S. & Sousa, C. R. e. Innate Antiviral Responses by Means of TLR7-Mediated Recognition of Single-Stranded RNA. *Science* **303,** 1529–1531 (2004).

59. Heil, F. *et al.* Species-Specific Recognition of Single-Stranded RNA via Toll-like Receptor 7 and 8. *Science* **303,** 1526–1529 (2004).

60. Hemmi, H. *et al.* Small anti-viral compounds activate immune cells via the TLR7 MyD88–dependent signaling pathway. *Nat. Immunol.* **3,** 196–200 (2002).

61. Hemmi, H. *et al.* A Toll-like receptor recognizes bacterial DNA. *Nature* **408,** 740–745 (2000).

62. Krieg, A. M. *et al.* CpG motifs in bacterial DNA trigger direct B-cell activation. *Nature* **374,** 546–549 (1995).

63. Charles A. Janeway, J. & Medzhitov, and R. Innate Immune Recognition. *Annu. Rev. Immunol.* **20,** 197–216 (2002).

64. Medzhitov, R. *et al.* MyD88 Is an Adaptor Protein in the hToll/IL-1 Receptor Family Signaling Pathways. *Mol. Cell* **2,** 253–258 (1998).

65. Yamamoto, M. *et al.* Role of Adaptor TRIF in the MyD88-Independent Toll-Like Receptor Signaling Pathway. *Science* **301,** 640–643 (2003).

66. Yamamoto, M. *et al.* TRAM is specifically involved in the Toll-like receptor 4–mediated MyD88-independent signaling pathway. *Nat. Immunol.* **4,** 1144–1150 (2003).

67. Sen, R. & Baltimore, D. Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* **46,** 705–716 (1986).

68. Schnare, M. *et al.* Toll-like receptors control activation of adaptive immune responses. *Nat. Immunol.* **2,** 947–950 (2001).

69. Pasare, C. & Medzhitov, R. Control of B-cell responses by Toll-like receptors. *Nature* **438,** 364–368 (2005).

70. Hombach, J., Leclercq, L., Radbruch, A., Rajewsky, K. & Reth, M. A novel 34-kd protein co-isolated with the IgM molecule in surface IgM-expressing cells. *EMBO J.* **7,** 3451–3456 (1988).

71. Campbell, K. S. & Cambier, J. C. B lymphocyte antigen receptors (mIg) are non-covalently associated with a disulfide linked, inducibly phosphorylated glycoprotein complex. *EMBO J.* **9,** 441–448 (1990).

72. Burkhardt, A. L., Brunswick, M., Bolen, J. B. & Mond, J. J. Anti-immunoglobulin stimulation of B lymphocytes activates src-related protein-tyrosine kinases. *Proc. Natl. Acad. Sci.* **88,** 7410–7414 (1991).

73. J C Cambier, C M Pleiman & Clark, and M. R. Signal Transduction by the B Cell Antigen Receptor and its Coreceptors. *Annu. Rev. Immunol.* **12,** 457–486 (1994).

74. Takata, M. *et al.* Tyrosine kinases Lyn and Syk regulate B cell receptor-coupled Ca2+ mobilization through distinct pathways. *EMBO J.* **13,** 1341–1349 (1994).

75. Khan, W. N. *et al.* Defective B cell development and function in Btk-deficient mice. *Immunity* **3,** 283–299 (1995).

76. Rawlings, D. J. *et al.* Mutation of unique region of Bruton's tyrosine kinase in immunodeficient XID mice. *Science* **261,** 358–361 (1993).

77. Fu, C., Turck, C. W., Kurosaki, T. & Chan, A. C. BLNK: a Central Linker Protein in B Cell Activation. *Immunity* **9,** 93–103 (1998).

78. Ishiai, M. *et al.* BLNK Required for Coupling Syk to PLCγ2 and Rac1-JNK in B Cells. *Immunity* **10,** 117–125 (1999).

79. Minegishi, Y. *et al.* An Essential Role for BLNK in Human B Cell Development. *Science* **286,** 1954–1957 (1999).

80. Pappu, R. *et al.* Requirement for B Cell Linker Protein (BLNK) in B Cell Development. *Science* **286,** 1949–1954 (1999).

81. Wang, D. *et al.* Phospholipase Cγ2 Is Essential in the Functions of B Cell and Several Fc Receptors. *Immunity* **13,** 25–35 (2000).

82. Tarakhovsky, A. *et al.* Defective antigen receptor-mediated proliferation of B and T cells in the absence of Vav. *Nature* **374,** 467–470 (1995).

83. Zhang, R., Alt, F. W., Davidson, L., Orkin, S. H. & Swat, W. Defective signalling through the T- and B-cell antigen receptors in lymphoid cells lacking the vav proto-oncogene. *Nature* **374,** 470–473 (1995).

84. Tuveson, D. A., Carter, R. H., Soltoff, S. P. & Fearon, D. T. CD19 of B cells as a surrogate kinase insert region to bind phosphatidylinositol 3-kinase. *Science* **260,** 986–989 (1993).

85. Fruman, D. A. *et al.* Impaired B Cell Development and Proliferation in Absence of Phosphoinositide 3-Kinase p85α. *Science* **283,** 393–397 (1999).

86. Niiro, H. & Clark, E. A. Regulation of B-cell fate by antigen-receptor signals. *Nat. Rev. Immunol.* **2,** 945–956 (2002).

87. Regulation of B-cell activation and differentiation by the phosphatidylinositol 3-kinase and phospholipase Cγ pathways. *Immunol. Rev.* **176,** 30–46 (2000).

88. MacLennan, I. C. M., Liu, Y.-J. & Johnson, G. D. Maturation and Dispersal of B-Cell Clones during T Cell-Dependent Antibody Responses. *Immunol. Rev.* **126,** 143–161 (1992).

89. Obukhanych, T. V. T-independent type II immune responses generate memory B cells. *J. Exp. Med.* **203,** 305–310 (2006).

90. Bortnick, A. *et al.* Long-Lived Bone Marrow Plasma Cells Are Induced Early in Response to T Cell-Independent or T Cell-Dependent Antigens. *J. Immunol.* **188,** 5389–5396 (2012).

91. Research, C. for B. E. and. Approved Products - Pneumovax 23 - Pneumococcal Vaccine, Polyvalent. Available at: http://www.fda.gov/biologicsbloodvaccines/vaccines/approvedproducts/ucm179996.htm. (Accessed: 24th March 2016)

92. Parker, D. C. T Cell-Dependent B Cell Activation. *Annu. Rev. Immunol.* **11,** 331–360 (1993).

93. Reif, K. *et al.* Balanced responsiveness to chemoattractants from adjacent zones determines B-cell position. *Nature* **416,** 94–99 (2002).

94. Legler, D. F. *et al.* B Cell–attracting Chemokine 1, a Human CXC Chemokine Expressed in Lymphoid Tissues, Selectively Attracts B Lymphocytes via BLR1/CXCR5. *J. Exp. Med.* **187,** 655–660 (1998).

95. Matloubian, M. *et al.* Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. *Nature* **427,** 355–360 (2004).

96. Cinamon, G. *et al.* Sphingosine 1-phosphate receptor 1 promotes B cell localization in the splenic marginal zone. *Nat. Immunol.* **5,** 713–720 (2004).

97. Armitage, R. J. *et al.* Molecular and biological characterization of a murine ligand for CD40. *Nature* **357,** 80–82 (1992).

98. MacLennan, I. C. Germinal centers. *Annu. Rev. Immunol.* **12,** 117–139 (1994).

99. Noelle, R. J. *et al.* A 39-kDa protein on activated helper T cells binds CD40 and transduces the signal for cognate activation of B cells. *Proc. Natl. Acad. Sci.* **89,** 6550–6554 (1992).

100. Zhu, J., Yamane, H. & Paul, W. E. Differentiation of Effector CD4 T Cell Populations. *Annu. Rev. Immunol.* **28,** 445–489 (2010).

101. Paul, W. E. & Seder, R. A. Lymphocyte responses and cytokines. *Cell* **76,** 241–251 (1994).

102. R A Seder & Paul, and W. E. Acquisition of Lymphokine-Producing Phenotype by CD4+ T Cells. *Annu. Rev. Immunol.* **12,** 635–673 (1994).

103. Mosmann, T. R., Cherwinski, H., Bond, M. W., Giedlin, M. A. & Coffman, R. L. Two types of murine helper T cell clone. I. Definition according to profiles of lymphokine activities and secreted proteins. *J. Immunol.* **136,** 2348–2357 (1986).

104. Snapper, C. M. & Paul, W. E. Interferon-gamma and B cell stimulatory factor-1 reciprocally regulate Ig isotype production. *Science* **236,** 944–947 (1987).

105. Kühn, R., Löhler, J., Rennick, D., Rajewsky, K. & Müller, W. Interleukin-10-deficient mice develop chronic enterocolitis. *Cell* **75,** 263–274 (1993).

106. Groux, H. *et al.* A CD4+T-cell subset inhibits antigen-specific T-cell responses and prevents colitis. *Nature* **389,** 737–742 (1997).

107. Ohara, J. & Paul, W. E. Production of a monoclonal antibody to and molecular characterization of B-cell stimulatory factor-1. *Nature* **315,** 333–336 (1985).

108. Vitetta, E. S. *et al.* Serological, biochemical, and functional identity of B cell-stimulatory factor 1 and B cell differentiation factor for IgG1. *J. Exp. Med.* **162,** 1726–1731 (1985).

109. Kuhn, R., Rajewsky, K. & Muller, W. Generation and analysis of interleukin-4 deficient mice. *Science* **254,** 707–710 (1991).

110. Takeda, K. *et al.* Essential role of Stat6 in IL-4 signalling. *Nature* **380,** 627–630 (1996).

111. Cua, D. J. *et al.* Interleukin-23 rather than interleukin-12 is the critical cytokine for autoimmune inflammation of the brain. *Nature* **421,** 744–748 (2003).

112. Aggarwal, S., Ghilardi, N., Xie, M.-H., Sauvage, F. J. de & Gurney, A. L. Interleukin-23 Promotes a Distinct CD4 T Cell Activation State Characterized by the Production of Interleukin-17. *J. Biol. Chem.* **278,** 1910–1914 (2003).

113. Murphy, C. A. *et al.* Divergent Pro- and Antiinflammatory Roles for IL-23 and IL-12 in Joint Autoimmune Inflammation. *J. Exp. Med.* **198,** 1951–1957 (2003).

114. Park, H. *et al.* A distinct lineage of CD4 T cells regulates tissue inflammation by producing interleukin 17. *Nat. Immunol.* **6,** 1133–1141 (2005).

115. Harrington, L. E. *et al.* Interleukin 17–producing CD4+ effector T cells develop via a lineage distinct from the T helper type 1 and 2 lineages. *Nat. Immunol.* **6,** 1123–1132 (2005).

116. Mitsdoerffer, M. *et al.* Proinflammatory T helper type 17 cells are effective B-cell helpers. *Proc. Natl. Acad. Sci.* **107,** 14292–14297 (2010).

117. Sakaguchi, S., Sakaguchi, N., Asano, M., Itoh, M. & Toda, M. Immunologic self-tolerance maintained by activated T cells expressing IL-2 receptor alpha-chains (CD25). Breakdown of a single mechanism of self-tolerance causes various autoimmune diseases. *J. Immunol.* **155,** 1151–1164 (1995).

118. Hori, S., Nomura, T. & Sakaguchi, S. Control of Regulatory T Cell Development by the Transcription Factor Foxp3. *Science* **299,** 1057–1061 (2003).

119. Fontenot, J. D., Gavin, M. A. & Rudensky, A. Y. Foxp3 programs the development and function of CD4+CD25+ regulatory T cells. *Nat. Immunol.* **4,** 330–336 (2003).

120. Lim, H. W., Hillsamer, P., Banham, A. H. & Kim, C. H. Cutting Edge: Direct Suppression of B Cells by CD4+CD25+ Regulatory T Cells. *J. Immunol.* **175,** 4180–4183 (2005).

121. MacLennan, I. C. M. Germinal Centers. *Annu. Rev. Immunol.* **12,** 117–139 (1994).

122. Muramatsu, M. *et al.* Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. *Cell* **102,** 553–563 (2000).

123. Nutt, S. L., Taubenheim, N., Hasbold, J., Corcoran, L. M. & Hodgkin, P. D. The genetic network controlling plasma cell differentiation. in (2011).

124. Hasbold, J., Corcoran, L. M., Tarlinton, D. M., Tangye, S. G. & Hodgkin, P. D. Evidence from the generation of immunoglobulin G–secreting cells that stochastic mechanisms regulate lymphocyte differentiation. *Nat. Immunol.* **5,** 55–63 (2004).

125. Slifka, M. K., Antia, R., Whitmire, J. K. & Ahmed, R. Humoral immunity due to long-lived plasma cells. *Immunity* **8,** 363–372 (1998).

126. Amanna, I. J., Carlson, N. E. & Slifka, M. K. Duration of humoral immunity to common viral and vaccine antigens. *N. Engl. J. Med.* **357,** 1903–1915 (2007).

127. Carotta, S. *et al.* The transcription factors IRF8 and PU.1 negatively regulate plasma cell differentiation. *J. Exp. Med.* jem.20140425 (2014). doi:10.1084/jem.20140425

128. Polli, M. *et al.* The development of functional B lymphocytes in conditional PU.1 knock-out mice. *Blood* **106,** 2083–2090 (2005).

129. Kometani, K. *et al.* Repression of the Transcription Factor Bach2 Contributes to Predisposition of IgG1 Memory B Cells toward Plasma Cell Differentiation. *Immunity* doi:10.1016/j.immuni.2013.06.011

130. Dent, A. L., Shaffer, A. L., Yu, X., Allman, D. & Staudt, L. M. Control of inflammation, cytokine expression, and germinal center formation by BCL-6. *Science* **276,** 589–592 (1997).

131. Urbánek, P., Wang, Z.-Q., Fetka, I., Wagner, E. F. & Busslinger, M. Complete block of early B cell differentiation and altered patterning of the posterior midbrain in mice lacking Pax5BSAP. *Cell* **79,** 901–912 (1994).

132. Nera, K.-P. *et al.* Loss of Pax5 Promotes Plasma Cell Differentiation. *Immunity* **24,** 283–293 (2006).

133. Delogu, A. *et al.* Gene Repression by Pax5 in B Cells Is Essential for Blood Cell Homeostasis and Is Reversed in Plasma Cells. *Immunity* **24,** 269–281 (2006).

134. Reimold, A. M. *et al.* Transcription factor B cell lineage-specific activator protein regulates the gene for human X-box binding protein 1. *J. Exp. Med.* **183,** 393–401 (1996).

135. Klemsz, M. J., McKercher, S. R., Celada, A., Van Beveren, C. & Maki, R. A. The macrophage and B cell-specific transcription factor PU.1 is related to the ets oncogene. *Cell* **61,** 113–124 (1990).

136. Muto, A. *et al.* Identification of Bach2 as a B-cell-specific partner for small Maf proteins that negatively regulate the immunoglobulin heavy chain gene 3′ enhancer. *EMBO J.* **17,** 5734–5743 (1998).

137. Muto, A. *et al.* The transcriptional programme of antibody class switching involves the repressor Bach2. *Nature* **429,** 566–571 (2004).

138. Turner Jr., C. A., Mack, D. H. & Davis, M. M. Blimp-1, a novel zinc finger-containing protein that can drive the maturation of B lymphocytes into immunoglobulin-secreting cells. *Cell* **77,** 297–306 (1994).

139. Lin, K.-I., Angelin-Duclos, C., Kuo, T. C. & Calame, K. Blimp-1-Dependent Repression of Pax-5 Is Required for Differentiation of B Cells to Immunoglobulin M-Secreting Plasma Cells. *Mol. Cell. Biol.* **22,** 4771–4780 (2002).

140. Piskurich, J. F. *et al.* BLIMP-1 mediates extinction of major histocompatibility class II transactivator expression in plasma cells. *Nat. Immunol.* **1,** 526–532 (2000).

141. Shaffer, A. L. *et al.* Blimp-1 Orchestrates Plasma Cell Differentiation by Extinguishing the Mature B Cell Gene Expression Program. *Immunity* **17,** 51–62 (2002).

142. Minnich, M. *et al.* Multifunctional role of the transcription factor Blimp-1 in coordinating plasma cell differentiation. *Nat. Immunol.* **advance online publication,** (2016).

143. Tellier, J. *et al.* Blimp-1 controls plasma cell function through the regulation of immunoglobulin secretion and the unfolded protein response. *Nat. Immunol.* **advance online publication,** (2016).

144. Iida, S. *et al.* Deregulation of MUM1/IRF4 by chromosomal translocation in multiple myeloma. *Nat. Genet.* **17,** 226–230 (1997).

145. Sciammas, R. *et al.* Graded Expression of Interferon Regulatory Factor-4 Coordinates Isotype Switching with Plasma Cell Differentiation. *Immunity* **25,** 225–236 (2006).

146. Klein, U. *et al.* Transcription factor IRF4 controls plasma cell differentiation and class-switch recombination. *Nat. Immunol.* **7,** 773–782 (2006).

147. Ochiai, K. *et al.* Transcriptional Regulation of Germinal Center B and Plasma Cell Fates by Dynamical Control of IRF4. *Immunity* doi:10.1016/j.immuni.2013.04.009

148. Emslie, D. *et al.* Oct2 enhances antibody-secreting cell differentiation through regulation of IL-5 receptor α chain expression on activated B cells. *J. Exp. Med.* **205,** 409–421 (2008).

149. Schubart, K. *et al.* B cell development and immunoglobulin gene transcription in the absence of Oct-2 and OBF-1. *Nat. Immunol.* **2,** 69–74 (2001).

150. Reimold, A. M. *et al.* Plasma cell differentiation requires the transcription factor XBP-1. *Nature* **412,** 300–307 (2001).

151. Shaffer, A. L. *et al.* XBP1, downstream of Blimp-1, expands the secretory apparatus and other organelles, and increases protein synthesis in plasma cell differentiation. *Immunity* **21,** 81–93 (2004).

152. Duffy, K. R. *et al.* Activation-Induced B Cell Fates Are Selected by Intracellular Stochastic Competition. *Science* **335,** 338–341 (2012).

153. Taylor, J. J., Pape, K. A., Steach, H. R. & Jenkins, M. K. Apoptosis and antigen affinity limit effector cell differentiation of a single naïve B cell. *Science* **347,** 784–787 (2015).

154. Jones, P. A. & Takai, D. The Role of DNA Methylation in Mammalian Epigenetics. *Science* **293,** 1068–1070 (2001).

155. Egger, G., Liang, G., Aparicio, A. & Jones, P. A. Epigenetics in human disease and prospects for epigenetic therapy. *Nature* **429,** 457–463 (2004).

156. Waddington, C. H. The Epigenotype. *Int. J. Epidemiol.* **41,** 10–13 (2012).

157. Waddington, C. H. Genetic Assimilation of the Bithorax Phenotype. *Evolution* **10,** 1 (1956).

158. Lee, R. C., Feinbaum, R. L. & Ambros, V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell* **75,** 843–854 (1993).

159. Berger, S. L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes Dev.* **23,** 781–783 (2009).

160. Scarano, E., Talarico, M., Bonaduce, L. & De Petrocellis, B. Enzymatic Deamination of 5-Deoxycytidylic Acid and of 5-Methyl-5′-deoxycytidylic Acid in Growing and in Non-growing Tissues. *Nature* **186,** 237–238 (1960).

161. Biswas, B. B. & Myers, J. A Methyl Cytidine from the Ribonucleic Acid of Anacystis nidulans. *Nature* **186,** 238–239 (1960).

162. Fleissner, E. & Borek, E. A NEW ENZYME OF RNA SYNTHESIS: RNA METHYLASE*. *Proc. Natl. Acad. Sci. U. S. A.* **48,** 1199–1203 (1962).

163. Scarano, E., Iaccarino, M., Grippo, P. & Winckelmans, D. On methylation of DNA during development of the sea urchin embryo. *J. Mol. Biol.* **14,** 603–607 (1965).

164. Grippo, P., Iaccarino, M., Parisi, E. & Scarano, E. Methylation of DNA in developing sea urchin embryos. *J. Mol. Biol.* **36,** 195–208 (1968).

165. Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187,** 226–232 (1975).

166. Bestor, T., Laudano, A., Mattaliano, R. & Ingram, V. 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. *J. Mol. Biol.* **203,** 971–983 (1988).

167. Yen, R.-W. C. *et al.* Isolation and characterization of the cDNA encoding human DNA methyltransferase. *Nucleic Acids Res.* **20,** 2287–2291 (1992).

168. Pradhan, S. *et al.* Baculovirus-mediated expression and characterization of the full-length murine DNA methyltransferase. *Nucleic Acids Res.* **25,** 4666–4673 (1997).

169. Weier, H.-U. & Bestor, T. H. A Targeting Sequence Directs DNA Methyltransferase to Sites of DNA Replication in Mammalian Nuclei.

170. Li, E., Bestor, T. H. & Jaenisch, R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69,** 915–926 (1992).

171. Okano, M., Xie, S. & Li, E. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nat. Genet.* **19,** 219–220 (1998).

172. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA Methyltransferases Dnmt3a and Dnmt3b Are Essential for De Novo Methylation and Mammalian Development. *Cell* **99,** 247–257 (1999).

173. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921 (2001).

174. Chinwalla, A. T. *et al.* Initial sequencing and comparative analysis of the mouse genome. *Nature* **420,** 520–562 (2002).

175. Bird, A., Taggart, M., Frommer, M., Miller, O. J. & Macleod, D. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* **40,** 91–99 (1985).

176. Bird, A. P. CpG-rich islands and the function of DNA methylation. *Nature* **321,** 209–213 (1986).

177. Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14,** 204–220 (2013).

178. Bentley, D. R. *et al.* Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456,** 53–59 (2008).

179. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27,** 1571–1572 (2011).

180. Meissner, A. *et al.* Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res.* **33,** 5868–5877 (2005).

181. Meissner, A. *et al.* Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454,** 766–770 (2008).

182. Lister, R. *et al.* Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462,** 315–322 (2009).

183. Liao, J. *et al.* Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. *Nat. Genet.* **advance online publication,** (2015).

184. Challen, G. A. *et al.* Dnmt3a and Dnmt3b Have Overlapping and Distinct Functions in Hematopoietic Stem Cells. *Cell Stem Cell* doi:10.1016/j.stem.2014.06.018

185. Baubec, T. *et al.* Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520,** 243–247 (2015).

186. Monk, M., Boubelik, M. & Lehnert, S. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99,** 371–382 (1987).

187. Oswald, J. *et al.* Active demethylation of the paternal genome in the mouse zygote. *Curr. Biol.* **10,** 475–478 (2000).

188. Ooi, S. K. T. & Bestor, T. H. The Colorful History of Active DNA Demethylation. *Cell* **133,** 1145–1148 (2008).

189. Tahiliani, M. *et al.* Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* **324,** 930–935 (2009).

190. Ito, S. *et al.* Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science* **333,** 1300–1303 (2011).

191. He, Y.-F. *et al.* Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. *Science* **333,** 1303–1307 (2011).

192. Cortellino, S. *et al.* Thymine DNA Glycosylase Is Essential for Active DNA Demethylation by Linked Deamination-Base Excision Repair. *Cell* **146,** 67–79 (2011).

193. Stadler, M. B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480,** 490–495 (2011).

194. Feldmann, A. *et al.* Transcription Factor Occupancy Can Mediate Active Turnover of DNA Methylation at Regulatory Regions. *PLOS Genet* **9,** e1003994 (2013).

195. Hashimoto, H. *et al.* Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res.* **40,** 4841–4849 (2012).

196. Bernstein, B. E., Meissner, A. & Lander, E. S. The Mammalian Epigenome. *Cell* **128,** 669–681 (2007).

197. Allfrey, V. G., Faulkner, R. & Mirsky, A. E. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci.* **51,** 786–794 (1964).

198. Hebbes, T. R., Clayton, A. L., Thorne, A. W. & Crane-Robinson, C. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *EMBO J.* **13,** 1823–1830 (1994).

199. Heintzman, N. D. *et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **459,** 108–112 (2009).

200. Creyghton, M. P. *et al.* Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci.* **107,** 21931–21936 (2010).

201. Santos-Rosa, H. *et al.* Active genes are tri-methylated at K4 of histone H3. *Nature* **419,** 407–411 (2002).

202. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406,** 593–599 (2000).

203. Tie, F. *et al.* CBP-mediated acetylation of histone H3 lysine 27 antagonizes Drosophila Polycomb silencing. *Development* **136,** 3131–3141 (2009).

204. Bernstein, B. E. *et al.* A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. *Cell* **125,** 315–326 (2006).

205. Li, B., Howe, L., Anderson, S., Yates, J. R. & Workman, J. L. The Set2 Histone Methyltransferase Functions through the Phosphorylated Carboxyl-terminal Domain of RNA Polymerase II. *J. Biol. Chem.* **278,** 8897–8903 (2003).

206. Nishioka, K. *et al.* PR-Set7 Is a Nucleosome-Specific Methyltransferase that Modifies Lysine 20 of Histone H4 and Is Associated with Silent Chromatin. *Mol. Cell* **9,** 1201–1213 (2002).

207. Kapoor-Vazirani, P., Kagey, J. D. & Vertino, P. M. SUV420H2-Mediated H4K20 Trimethylation Enforces RNA Polymerase II Promoter-Proximal Pausing by Blocking hMOF-Dependent H4K16 Acetylation. *Mol. Cell. Biol.* **31,** 1594–1609 (2011).

208. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403,** 41–45 (2000).

209. Jenuwein, T. & Allis, C. D. Translating the Histone Code. *Science* **293,** 1074–1080 (2001).

210. Hashimoto, H., Vertino, P. M. & Cheng, X. Molecular coupling of DNA methylation and histone methylation. *Epigenomics* **2,** 657–669 (2010).

211. Bourc'his, D., Xu, G.-L., Lin, C.-S., Bollman, B. & Bestor, T. H. Dnmt3L and the Establishment of Maternal Genomic Imprints. *Science* **294,** 2536–2539 (2001).

212. Bourc'his, D. & Bestor, T. H. Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431,** 96–99 (2004).

213. Ooi, S. K. T. *et al.* DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448,** 714–717 (2007).

214. Jia, D., Jurkowska, R. Z., Zhang, X., Jeltsch, A. & Cheng, X. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* **449,** 248–251 (2007).

215. Chang, Y. *et al.* MPP8 mediates the interactions between DNA methyltransferase Dnmt3a and H3K9 methyltransferase GLP/G9a. *Nat. Commun.* **2,** 533 (2011).

216. Dunham, I. *et al.* An integrated encyclopedia of DNA elements in the human genome. *Nature* **489,** 57–74 (2012).

217. Cheng, Y. *et al.* Principles of regulatory information conservation between mouse and human. *Nature* **515,** 371–375 (2014).

218. Gerstein, M. B. *et al.* Architecture of the human regulatory network derived from ENCODE data. *Nature* **489,** 91–100 (2012).

219. Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473,** 43–49 (2011).

220. Ho, J. W. K. *et al.* Comparative analysis of metazoan chromatin organization. *Nature* **512,** 449–452 (2014).

221. Maurano, M. T. *et al.* Large-scale identification of sequence variants influencing human transcription factor occupancy in vivo. *Nat. Genet.* **advance online publication,** (2015).

222. Neph, S. *et al.* An expansive human regulatory lexicon encoded in transcription factor footprints. *Nature* **489,** 83–90 (2012).

223. Stergachis, A. B. *et al.* Exonic Transcription Factor Binding Directs Codon Choice and Affects Protein Evolution. *Science* **342,** 1367–1372 (2013).

224. Stergachis, A. B. *et al.* Developmental Fate and Cellular Maturity Encoded in Human Regulatory DNA Landscapes. *Cell* **154,** 888–903 (2013).

225. Stergachis, A. B. *et al.* Conservation of trans-acting circuitry during mammalian regulatory evolution. *Nature* **515,** 365–370 (2014).

226. Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489,** 75–82 (2012).

227. Yue, F. *et al.* A comparative encyclopedia of DNA elements in the mouse genome. *Nature* **515,** 355–364 (2014).

228. Dixon, J. R. *et al.* Chromatin architecture reorganization during stem cell differentiation. *Nature* **518,** 331–336 (2015).

229. Elliott, G. *et al.* Intermediate DNA methylation is a conserved signature of genome regulation. *Nat. Commun.* **6,** (2015).

230. Farh, K. K.-H. *et al.* Genetic and epigenetic fine mapping of causal autoimmune disease variants. *Nature* **518,** 337–343 (2015).

231. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518,** 317–330 (2015).

232. Tsankov, A. M. *et al.* Transcription factor binding dynamics during human ES cell differentiation. *Nature* **518,** 344–349 (2015).

233. Ziller, M. J. *et al.* Dissecting neural differentiation regulatory networks through epigenetic footprinting. *Nature* **518,** 355–359 (2015).

234. Mizuno, S. *et al.* Expression of DNA methyltransferases DNMT1,3A, and 3B in normal hematopoiesis and in acute and chronic myelogenous leukemia. *Blood* **97,** 1172–1179 (2001).

235. Tadokoro, Y., Ema, H., Okano, M., Li, E. & Nakauchi, H. De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. *J. Exp. Med.* **204,** 715–722 (2007).

236. Uchida, T. *et al.* Hypermethylation of the p15INK4B Gene in Myelodysplastic Syndromes. *Blood* **90,** 1403–1409 (1997).

237. Bröske, A.-M. *et al.* DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction. *Nat. Genet.* **41,** 1207–1215 (2009).

238. Trowbridge, J. J., Snow, J. W., Kim, J. & Orkin, S. H. DNA Methyltransferase 1 Is Essential for and Uniquely Regulates Hematopoietic Stem and Progenitor Cells. *Cell Stem Cell* **5,** 442–449 (2009).

239. Jackson-Grusby, L. *et al.* Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat. Genet.* **27,** 31–39 (2001).

240. Kuhn, R., Schwenk, F., Aguet, M. & Rajewsky, K. Inducible gene targeting in mice. *Science* **269,** 1427–1429 (1995).

241. Ji, H. *et al.* Comprehensive methylome map of lineage commitment from haematopoietic progenitors. *Nature* **467,** 338–342 (2010).

242. Tadokoro, Y., Ema, H., Okano, M., Li, E. & Nakauchi, H. De novo DNA methyltransferase is essential for self-renewal, but not for differentiation, in hematopoietic stem cells. *J. Exp. Med.* **204,** 715–722 (2007).

243. Challen, G. A. *et al.* Dnmt3a is essential for hematopoietic stem cell differentiation. *Nat. Genet.* **44,** 23–31 (2012).

244. Jeong, M. *et al.* Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat. Genet.* **advance online publication,** (2013).

245. Mayle, A. *et al.* Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation. *Blood* **125,** 629–638 (2015).

246. Ley, T. J. *et al.* DNMT3A Mutations in Acute Myeloid Leukemia. *N. Engl. J. Med.* **363,** 2424–2433 (2010).

247. Yang, L., Rau, R. & Goodell, M. A. DNMT3A in haematological malignancies. *Nat. Rev. Cancer* **15,** 152–165 (2015).

248. Couronné, L., Bastard, C. & Bernard, O. A. TET2 and DNMT3A Mutations in Human T-Cell Lymphoma. *N. Engl. J. Med.* **366,** 95–96 (2012).

249. Lorsbach, R. B. *et al.* TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia* **17,** 637–641 (2003).

250. Delhommeau, F. *et al.* Mutation in TET2 in Myeloid Cancers. *N. Engl. J. Med.* **360,** 2289–2301 (2009).

251. Abdel-Wahab, O. *et al.* Genetic characterization of TET1, TET2, and TET3 alterations in myeloid malignancies. *Blood* **114,** 144–147 (2009).

252. Network, T. C. G. A. R. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. *N. Engl. J. Med.* **368,** 2059–2074 (2013).

253. Jankowska, A. M. *et al.* Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms. *Blood* **113,** 6403–6410 (2009).

254. Quivoron, C. *et al.* TET2 Inactivation Results in Pleiotropic Hematopoietic Abnormalities in Mouse and Is a Recurrent Event during Human Lymphomagenesis. *Cancer Cell* **20,** 25–38 (2011).

255. Cimmino, L. *et al.* TET1 is a tumor suppressor of hematopoietic malignancy. *Nat. Immunol.* **advance online publication,** (2015).

256. Moran-Crusio, K. *et al.* Tet2 Loss Leads to Increased Hematopoietic Stem Cell Self-Renewal and Myeloid Transformation. *Cancer Cell* **20,** 11–24 (2011).

257. Ko, M. *et al.* Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc. Natl. Acad. Sci.* **108,** 14566–14571 (2011).

258. An, J. *et al.* Acute loss of TET function results in aggressive myeloid cancer in mice. *Nat. Commun.* **6,** 10071 (2015).

259. Heinz, S. *et al.* Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38,** 576–589 (2010).

260. Lin, Y. C. *et al.* A global network of transcription factors, involving E2A, EBF1 and Foxo1, that orchestrates B cell fate. *Nat. Immunol.* **11,** 635–643 (2010).

261. Kee, B. L. & Murre, C. Induction of Early B Cell Factor (EBF) and Multiple B Lineage Genes by the Basic Helix-Loop-Helix Transcription Factor E12. *J. Exp. Med.* **188,** 699–713 (1998).

262. Treiber, T. *et al.* Early B Cell Factor 1 Regulates B Cell Gene Networks by Activation, Repression, and Transcription- Independent Poising of Chromatin. *Immunity* **32,** 714–725 (2010).

263. Decker, T. *et al.* Stepwise Activation of Enhancer and Promoter Regions of the B Cell Commitment Gene Pax5 in Early Lymphopoiesis. *Immunity* **30,** 508–520 (2009).

264. Yancopoulos, G. D., Blackwell, T. K., Suh, H., Hood, L. & Alt, F. W. Introduced T cell receptor variable region gene segments recombine in pre-B cells: Evidence that B and T cells use a common recombinase. *Cell* **44,** 251–259 (1986).

265. Lichtenstein, M., Keini, G., Cedar, H. & Bergman, Y. B cell-specific demethylation: A novel role for the intronic κ chain enhancer sequence. *Cell* **76,** 913–923 (1994).

266. Kirillov, A. *et al.* A role for nuclear NF–κB in B–cell–specific demethylation of the Igκ locus. *Nat. Genet.* **13,** 435–441 (1996).

267. Manoharan, A., Roure, C. D., Rolink, A. G. & Matthias, P. De novo DNA Methyltransferases Dnmt3a and Dnmt3b regulate the onset of Igκ light chain rearrangement during early B-cell development. *Eur. J. Immunol.* **45,** 2343–2355 (2015).

268. Goldmit, M. *et al.* Epigenetic ontogeny of the Igk locus during B cell development. *Nat. Immunol.* **6,** 198–203 (2005).

269. Liu, Y., Subrahmanyam, R., Chakraborty, T., Sen, R. & Desiderio, S. A Plant Homeodomain in Rag-2 that Binds Hypermethylated Lysine 4 of Histone H3 Is Necessary for Efficient Antigen-Receptor-Gene Rearrangement. *Immunity* **27,** 561–571 (2007).

270. Matthews, A. G. W. *et al.* RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* **450,** 1106–1110 (2007).

271. Morin, R. D. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* **42,** 181–185 (2010).

272. Béguelin, W. *et al.* EZH2 Is Required for Germinal Center Formation and Somatic EZH2 Mutations Promote Lymphoid Transformation. *Cancer Cell* **23,** 677–692 (2013).

273. Shaknovich, R. *et al.* DNA methyltransferase 1 and DNA methylation patterning contribute to germinal center B-cell differentiation. *Blood* **118,** 3559–3569 (2011).

274. Dominguez, P. M. *et al.* DNA Methylation Dynamics of Germinal Center B Cells Are Mediated by AID. *Cell Rep.* **12,** 2086–2098 (2015).

275. Muramatsu, M. *et al.* Class Switch Recombination and Hypermutation Require Activation-Induced Cytidine Deaminase (AID), a Potential RNA Editing Enzyme. *Cell* **102,** 553–563 (2000).

276. Revy, P. *et al.* Activation-Induced Cytidine Deaminase (AID) Deficiency Causes the Autosomal Recessive Form of the Hyper-IgM Syndrome (HIGM2). *Cell* **102,** 565–575 (2000).

277. Fritz, E. L. *et al.* A comprehensive analysis of the effects of the deaminase AID on the transcriptome and methylome of activated B cells. *Nat. Immunol.* **14,** 749–755 (2013).

278. Lai, A. Y. *et al.* DNA methylation profiling in human B cells reveals immune regulatory elements and epigenetic plasticity at Alu elements during B-cell activation. *Genome Res.* **23,** 2030–2041 (2013).

279. Kulis, M. *et al.* Whole-genome fingerprint of the DNA methylome during human B cell differentiation. *Nat. Genet.* **advance online publication,** (2015).

280. Hodgkin, P. D., Lee, J. H. & Lyons, A. B. B cell differentiation and isotype switching is related to division cycle number. *J. Exp. Med.* **184,** 277–281 (1996).

281. Kallies, A. *et al.* Plasma Cell Ontogeny Defined by Quantitative Changes in Blimp-1 Expression. *J. Exp. Med.* **200,** 967–977 (2004).

282. Chernova, I. *et al.* Lasting Antibody Responses Are Mediated by a Combination of Newly Formed and Established Bone Marrow Plasma Cells Drawn from Clonally Distinct Precursors. *J. Immunol.* **193,** 4971–4979 (2014).

283. Tarte, K., Zhan, F., Vos, J. D., Klein, B. & Shaughnessy, J. Gene expression profiling of plasma cells and plasmablasts: toward a better understanding of the late stages of B-cell differentiation. *Blood* **102,** 592–600 (2003).

284. Feng, H., Conneely, K. N. & Wu, H. A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucleic Acids Res.* **42,** e69–e69 (2014).

285. Matsumoto, M. *et al.* Interleukin-10-Producing Plasmablasts Exert Regulatory Function in Autoimmune Inflammation. *Immunity* **0,**

286. Shi, W. *et al.* Transcriptional profiling of mouse B cell terminal differentiation defines a signature for antibody-secreting plasma cells. *Nat. Immunol.* **advance online publication,** (2015).

287. Salzer, U. *et al.* Mutations in TNFRSF13B encoding TACI are associated with common variable immunodeficiency in humans. *Nat. Genet.* **37,** 820–828 (2005).

288. Xiong, Z. & Laird, P. W. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* **25,** 2532–2534 (1997).

289. Mittrücker, H.-W. *et al.* Requirement for the Transcription Factor LSIRF/IRF4 for Mature B and T Lymphocyte Function. *Science* **275,** 540–543 (1997).

290. Herrscher, R. F. *et al.* The immunoglobulin heavy-chain matrix-associating regions are bound by Bright: a B cell-specific trans-activator that describes a new DNA-binding protein family. *Genes Dev.* **9,** 3067–3082 (1995).

291. Creyghton, M. P. *et al.* From the Cover: Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc. Natl. Acad. Sci.* **107,** 21931–21936 (2010).

292. Sabò, A. *et al.* Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. *Nature* **511,** 488–492 (2014).

293. Grötsch, B. *et al.* The AP-1 transcription factor Fra1 inhibits follicular B cell differentiation into plasma cells. *J. Exp. Med.* jem.20130795 (2014). doi:10.1084/jem.20130795

294. Sasaki, Y. *et al.* Canonical NF-κB Activity, Dispensable for B Cell Development, Replaces BAFF-Receptor Signals and Promotes B Cell Proliferation upon Activation. *Immunity* **24,** 729–739 (2006).

295. Peperzak, V. *et al.* Mcl-1 is essential for the survival of plasma cells. *Nat. Immunol.* **14,** 290–297 (2013).

296. Yoon, H. S. *et al.* ZBTB32 Is an Early Repressor of the CIITA and MHC Class II Gene Expression during B Cell Differentiation to Plasma Cells. *J. Immunol.* **189,** 2393–2403 (2012).

297. Kitamura, D., Roes, J., Kühn, R. & Rajewsky, K. A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin μ chain gene. *Nature* **350,** 423–426 (1991).

298. Hu, M. *et al.* p32 protein levels are integral to mitochondrial and endoplasmic reticulum morphology, cell metabolism and survival. *Biochem. J.* **453,** 381–391 (2013).

299. Bruni, F., Gramegna, P., Oliveira, J. M. A., Lightowlers, R. N. & Chrzanowska-Lightowlers, Z. M. A. REXO2 Is an Oligoribonuclease Active in Human Mitochondria. *PLoS ONE* **8,** e64670 (2013).

300. Manz, R. A., Thiel, A. & Radbruch, A. Lifetime of plasma cells in the bone marrow. *Nature* **388,** 133–134 (1997).

301. Chang, J. T. *et al.* Asymmetric T Lymphocyte Division in the Initiation of Adaptive Immune Responses. *Science* **315,** 1687–1691 (2007).

302. Hahne, F. *et al.* flowCore: a Bioconductor package for high throughput flow cytometry. *BMC Bioinformatics* **10,** 106 (2009).

303. Langmead, B., Trapnell, C., Pop, M., Salzberg, S. L. & others. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10,** R25 (2009).

304. Hsu, F. *et al.* The UCSC Known Genes. *Bioinformatics* **22,** 1036–1046 (2006).

305. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5,** R80 (2004).

306. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Methodol.* **57,** 289–300 (1995).

307. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36 (2013).

308. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges. *PLoS Comput Biol* **9,** e1003118 (2013).

309. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. 139–140 (2009).

310. Falcon, S. & Gentleman, R. Using GOstats to test gene lists for GO term association. *Bioinformatics* **23,** 257–258 (2007).

311. Bowman, S. K. *et al.* Multiplexed Illumina sequencing libraries from picogram quantities of DNA. *BMC Genomics* **14,** 466 (2013).

312. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9,** 357–359 (2012).

313. Park, Y. & Wu, H. Differential methylation analysis for BS-seq data under general experimental design. *Bioinformatics* btw026 (2016).

314. Fisher, R. A. On the interpretation of x(2) from contingency tables, and the calculation of P. *J. R. Stat. Soc.* **85,** 87–94 (1922).

315. Tan, G. & Lenhard, B. TFBSTools: an R/Bioconductor package for transcription factor binding site analysis. *Bioinformatics* btw024 (2016).

316. Frangini, A. *et al.* The Aurora B Kinase and the Polycomb Protein Ring1B Combine to Regulate Active Promoters in Quiescent Lymphocytes. *Mol. Cell* **51,** 647–661 (2013).

317. Revilla-i-Domingo, R. *et al.* The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J.* **31,** 3130–3146 (2012).

318. Jothi, R., Cuddapah, S., Barski, A., Cui, K. & Zhao, K. Genome-wide identification of in vivo protein–DNA binding sites from ChIP-Seq data. *Nucleic Acids Res.* **36,** 5221–5231 (2008).

319. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9,** R137 (2008).

320. Njau, M. N. *et al.* CD28–B7 Interaction Modulates Short- and Long-Lived Plasma Cell Function. *J. Immunol.* **189,** 2758–2767 (2012).

321. Lee, P. P. *et al.* A critical role for Dnmt1 and DNA methylation in T cell development, function, and survival. *Immunity* **15,** 763–774 (2001).

322. Sellars, M. *et al.* Regulation of DNA methylation dictates Cd4 expression during the development of helper and cytotoxic T cell lineages. *Nat. Immunol.* **16,** 746–754 (2015).

323. Chappell, C., Beard, C., Altman, J., Jaenisch, R. & Jacob, J. DNA methylation by DNA methyltransferase 1 is critical for effector CD8 T cell expansion. *J. Immunol.* **176,** 4562–4572 (2006).

324. Kaneda, M. *et al.* Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. 900–903 (2004).

325. Dodge, J. E. *et al.* Inactivation of Dnmt3b in Mouse Embryonic Fibroblasts Results in DNA Hypomethylation, Chromosomal Instability, and Spontaneous Immortalization. *J. Biol. Chem.* **280,** 17986–17991 (2005).

326. Oliver, A. M., Martin, F., Gartland, G. L., Carter, R. H. & Kearney, J. F. Marginal zone B cells exhibit unique activation, proliferative and immunoglobulin secretory responses. *Eur. J. Immunol.* **27,** 2366–2374 (1997).

327. Schittek, B. & Rajewsky, K. Maintenance of B-cell memory by long-lived cells generated from proliferating precursors. *Nature* **346,** 749–751 (1990).

328. Pape, K. A., Taylor, J. J., Maul, R. W., Gearhart, P. J. & Jenkins, M. K. Different B Cell Populations Mediate Early and Late Memory During an Endogenous Immune Response. *Science* **331,** 1203–1207 (2011).

329. Taylor, J. J., Pape, K. A. & Jenkins, M. K. A germinal center–independent pathway generates unswitched memory B cells early in the primary response. *J. Exp. Med.* **209,** 597–606 (2012).

330. Shinkai, Y. *et al.* RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell* **68,** 855–867 (1992).

331. Rickert, R. C., Roes, J. & Rajewsky, K. B Lymphocyte-Specific, Cre-mediated Mutagenesis in Mice. *Nucleic Acids Res.* **25,** 1317–1318 (1997).

332. Ng, M. H. L. *et al.* Frequent Hypermethylation of p16 and p15 Genes in Multiple Myeloma. *Blood* **89,** 2500–2506 (1997).

333. Chim, C.-S., Fung, T.-K., Cheung, W.-C., Liang, R. & Kwong, Y.-L. SOCS1 and SHP1 hypermethylation in multiple myeloma: implications for epigenetic activation of the Jak/STAT pathway. *Blood* **103,** 4630–4635 (2004).

334. Challen, G. A. *et al.* Dnmt3a and Dnmt3b Have Overlapping and Distinct Functions in Hematopoietic Stem Cells. *Cell Stem Cell* **15,** 350–364 (2014).

335. Su, I. *et al.* Ezh2 controls B cell development through histone H3 methylation and Igh rearrangement. *Nat. Immunol.* **4,** 124–131 (2002).

336. Lin, C. Y. *et al.* Transcriptional Amplification in Tumor Cells with Elevated c-Myc. *Cell* **151,** 56–67 (2012).

337. Agirre, X. *et al.* Whole-epigenome analysis in multiple myeloma reveals DNA hypermethylation of B cell-specific enhancers. *Genome Res.* gr.180240.114 (2015). doi:10.1101/gr.180240.114

338. Gray, K. S., Forrest, J. C. & Speck, S. H. The De Novo Methyltransferases DNMT3a and DNMT3b Target the Murine Gammaherpesvirus Immediate-Early Gene 50 Promoter during Establishment of Latency. *J. Virol.* **84,** 4946–4959 (2010).