Summer 8-18-2017

DNMT3A Haploinsufficiency Provokes Hematologic Malignancy of B-Lymphoid, T-Lymphoid, and Myeloid Lineage in Mice

Garland Michael Upchurch
University of Nebraska Medical Center

DNMT3A HAPLOINSUFFICIENCY PROVOKES HEMATOLOGIC MALIGNANCY OF B-LYMPHOID, T-LYMPHOID, AND MYELOID LINEAGE IN MICE

By

Garland Michael Upchurch

A DISSERTATION

Presented to the Faculty of the University of Nebraska Graduate College in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Cancer Research Graduate Program

Eppley Institute for Cancer Research and Allied Diseases

Under the Supervision of Professors Rene Opavsky, Ph.D. & Angie Rizzino, Ph.D.

University of Nebraska Medical Center

Omaha, Nebraska

May, 2017

Supervisory Committee:

Adam R. Karpf, Ph.D. David L. Klinkebiel, Ph.D. Samuel J. Pirruccello, M.D.

ACKNOWLEDGMENTS

Firstly, I would like to gratefully acknowledge my primary mentor, Dr. Rene Opavsky, for his contributions to my development as a scientific researcher during my tenure as a doctoral candidate in his laboratory. I wish to thank Rene and my fellow laboratory members, Staci and Jana, for fostering in me empiricism, experimentalism, and productivity. Likewise, I am eternally grateful to Angie, Wanguo, and Bobby as sources of additional mentorship during my graduate career. Thank you for encouraging critical thinking, challenging dogma, and always supporting academic thinking.

Secondly, I wish to acknowledge the support, advice, and expert opinion bestowed upon me by my committee members Dave, Adam, and Sam. Thank you for your sage wisdom, scientific discussion, and pleasant conversations.

Lastly, to my dear wife Jessica, daughters Megan, Emily, Laura, and those children who one day may yet still arrive; you have my eternal gratitude for your past, present, and future love, patience, and support of my lifelong academic training.

~ Garland Michael Upchurch

ABSTRACT

DNMT3A HAPLOINSUFFICY PROVOKES HEMATOLOGIC MALIGNANCY OF B-LYMPHOID, T-LYMPHOID, AND MYELOID LINEAGE IN MICE

Garland Michael Upchurch, Ph.D.

University of Nebraska, 2017

Supervisors: Rene Opavsky, Ph.D. & Angie Rizzino, Ph.D

DNA methyltransferase 3A (DNMT3A) is a master epigenetic regulator of benign and malignant hematopoiesis. To dissect the biological consequences of homozygous and heterozygous *Dnmt3a* inactivation in malignant hematopoiesis, we generated *Dnmt3a* homozygous null (*Dnmt3a*$^{\Delta/\Delta}$) and *Dnmt3a* heterozygous (*Dnmt3a*$^{+/-}$) mice and compared the presentations of hematologic malignancies between cohorts. Bi-allelic inactivation of *Dnmt3a* results in the presentation of mature lymphoid neoplasms resembling chronic lymphocytic leukemia (CLL; B220^{+}CD19^{+}CD5^{+}; 88% penetrance (37/42)) and CD8+ peripheral T-cell lymphoma (PTCL; TCRβ^{+}CD3^{+}CD8^{+}CD4^{-}; 40% penetrance (17/42)). In contrast, mono-allelic inactivation of *Dnmt3a* results in the presentation of CLL and PTCL at reduced penetrance (47% (14/30) & 10% (3/30), respectively) and, rarely, a mature myeloproliferative neoplasm (MPN; CD11b^{+}Gr-1^{+}; 10% penetrance (3/30)).

Molecular interrogation of PTCLs revealed genome-wide deregulation of DNA methylation, characterized by 10-fold greater hypomethylation than hypermethylation of promoters and enhancers. Transcription factor binding sites for AML1, NF-κB, and OCT1 were enriched in hypomethylated promoters, implicating these transcription factors in tumor pathogenesis or

DNMT3A-associated DNA methylation. Whereas 71 hypomethylated genes showed an increased expression in PTCL, only 3 hypermethylated genes were silenced, suggesting cancer-specific hypomethylation more frequently affects the transcriptome than hypermethylation in lymphoma. Importantly, we observed in *Dnmt3a*-deficient PTCLs the downregulation of p53 protein by western blot and p53 target genes by gene set enrichment analysis. Decreased p53 protein expression occurred in pre-tumor thymocytes of 9 months old, but not 6 weeks old, $Dnmt3a^{+/-}$ disease-free mice, demonstrating that p53 downregulation occurs an intermediate event in tumorigenesis. Analysis of $Dnmt3a^{+/-}$ tumors revealed that PTCL develops without mutation or silencing of the remaining wild-type *Dnmt3a* allele. These data demonstrate that *Dnmt3a* is a haploinsufficient tumor suppressor in murine mature $CD8^+$ PTCL and loss of p53 protein occurs as an intermediate event in tumorigenesis.

To better understand the dysregulated epigenetic events favoring the development of CLL and PTCL from B-1a and CD8+ cells in $Dnmt3a^{\Delta/\Delta}$ mice, we compared the methylomes and transcriptomes of normal B-1a and CD8+ T-cells in addition to comparison of malignant CLL and PTCL cells. We observe that whereas patterns of methylation and transcription in normal B-1a cells and CD8+ T cells are similar, methylomes and transcriptomes in malignant B-1a and CD8+ T cells are remarkably distinct, suggesting a cell-type specific function for *Dnmt3a* in cellular transformation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AITL	Angioimmunoblastic T-Cell Lymphoma
ALL	Acute Lymphoblastic Leukemia
AML	Acute Myeloid Leukemia
B-ALL	B-Cell Acute Lymphoblastic Leukemia
Bp	Base Pair
CLL	Chronic Lymphocytic Leukemia
COBRA	Combined Bisulfite Restriction Analysis
DMC	Differentially Methylated Cytosine
DMR	Differentially Methylated Region
DNMTs	DNA Methyltransferases
EGFP	Enhanced Green Fluorescent Protein
HAT	Histone Acetyltransferase
HDAC	Histone Deacetylases
HDM	Histone Demethylase
HMT	Histone Methyltransferase
HSC	Hematopoietic Stem Cell
HOC	Hypomethylated and Overexpressed in *Dnmt3a*$^{\Delta/\Delta}$ CLL
HOT	Hypomethylated and Overexpressed in *Dnmt3a*$^{\Delta/\Delta}$ PTCL
IPA	Ingenuity Pathway Analysis
MBL	Monoclonal B-Cell Lymphocytosis
MDS	Myelodysplastic Syndrome
MPD	Myeloproliferative Disease
MPN	Myeloproliferative Neoplasm
NHL	Non-Hodgkin Lymphoma

PTCL	Peripheral T-Cell Lymphoma
PTCL-NOS	Peripheral T-Cell Lymphoma – Not Otherwise Specified
RRBS	Reduced Representative Bisulfite Sequencing
T-ALL	T-Cell Acute Lymphoblastic Leukemia
TCL	T-Cell Lymphoma
T-PLL	T-Cell Prolymphocytic Leukemia
TSS	Transcriptional Start Site
WGBS	Whole Genome Bisulfite Sequencing

CHAPTER 1:

INTRODUCTION

Preface

Within the past decade, scientific exploration of epigenetics has developed from budding nascence into full bloom, particularly with focus in oncology research. Cancer epigenetics – the field of study born from the marriage of oncology and molecular epigenetics – has blossomed in large part due to the discovery of mutations in epigenetic modifier genes present in hematologic malignancies. Mutations of certain epigenetic modifiers are commonly detected in varieties of neoplasms such as *IDH1/2* in gliomas [1] or *SMARCB1* ("*hSNF5"*, a member of the human SWI/SNF nucleosome remodeling complex) in pediatric malignant rhabdoid tumors [2], however none more so than in the hematopoietic compartment, where myriad epigenetic modifier genes are detected to be mutated or dysregulated in malignancy. Under normal physiologic conditions, these genes operate to maintain the fidelity of fundamental cellular processes crucial to hematopoietic cells such as regulation of DNA methylation (*DNMT3A, IDH1/2, TET2*), chromatin remodeling (*MLL1/2, ASXL-1, EZH2*), and pre-mRNA splicing (*SF3B1, SRSF2, U2AF1, ZRSR2*) [3, 4]. The aberrant activity (or inactivity) of proteins encoded by these genes begets a state of epigenetic dysregulation and pre-leukemic clonal hematopoiesis that eventually culminates in the provocation of frank hematologic malignancy [5, 6]. Incredibly, *DNMT3A* is the most frequently inactivated gene among epigenetic modifiers and is one of the most commonly mutated genes among all hematologic malignancy. The precise mechanism as to how mutation or inactivation of *DNMT3A* provokes cellular transformation is a subject of current investigation and is a focus of this dissertation. As epigenetic modifiers are classically associated with regulation of gene expression during development, one prevailing theory in the field of cancer epigenetics is that lesion of epigenetic modifiers, such as *DNMT3A*, deregulates gene transcription via loss of epigenetic regulation, which in turn favors the acquisition of additional driver events and leads to cellular transformation. The work presented here examines the cellular and molecular consequences of *Dnmt3a* inactivation in mice with particular emphasis on how loss of *Dnmt3a* results in gene

promoter hypomethylation and transcriptional over-expression in murine hematologic malignancies.

Myeloid Malignancy & the Two-Hit Hypothesis of Leukemogenesis

Hematologic malignancies are neoplasms arising from the cellular components of the blood, bone marrow, and lymphatic systems. These neoplasms derive from cellular transformation of hematopoietic stem and progenitor cells (HSPCs) or their downstream lineage-committed progeny and can be generalized into two broad classifications bases on their developmental lineage: myeloid-erythroid ("myeloid") and lymphoid. Cells of the myeloid lineage constitute an ancient evolutionary branch of the hematopoietic system and give rise to cells that primarily function to oxygenate the body, generate blood clots, and form the cellular innate immune system. The lymphoid lineage is a more recent evolutionary branch of the hematopoietic system that arose in jawed vertebrate fish 500 million years ago due to the immunological advantage of an adaptive immune system on host defense [7]. (However, exceptions to the canonical roles of myeloid cells and lymphoid cells in immunological function exist, as antigen presentation by macrophage and dendritic cells allow these cells to transcend both the innate and adaptive immune systems. Meanwhile, lymphoid NK cells only participate in innate immunity.) Knowledge of the developmental ontogeny of the cells that constitute the hematopoietic system is crucial to understanding the pathological mechanisms underlying both the various blocks in cellular differentiation observed in acute leukemia and the excess abundance of select mature lineages observed in chronic myeloproliferative and lymphoproliferative neoplasms. This morphological classification by ontogeny and state of differentiation was in fact the basis of the FAB (French-American-British) system for classification of hematologic disease and, with incorporation of molecular data, has now evolved into the current WHO Classification of Hematological Malignancies, most recently updated in 2016 [8, 9].

Mature cells of the myeloid lineage present in peripheral blood or tissue include platelets (megakaryocytic lineage), erythrocytes (erythroid lineage), segmented neutrophils, eosinophils, basophils (granulocytic lineage), monocytes/macrophage, myeloid dendritic cells (monocytic lineage), and mast cells. Under the classical hierarchical model of hematopoietic development, multipotent HSPC blasts (LT-HSCs and ST-HSCs) present in the bone marrow may self-renew or differentiate into one of two lineage-restricted oligopotent progenitor cells: common myeloid progenitors (CMP) or common lymphoid progenitors (CLP) (also known as myeloid and lymphoid stem cells, respectively). Under this model, all differentiated myeloid cells derive from the CMP cell. Following cytokine stimulation of cognate cellular receptors and appropriately induced transcription factor expression, the developmental fate of CMPs may be further restricted to one of two committed precursor cells: megakaryocyte-erythroid progenitors (MEP) or granulocyte-macrophage progenitors (GMP). MEPs subsequently differentiate into unipotent megakaryoblasts or proerythroblasts (pronormoblasts) in order to give rise to the terminally differentiated cells of the megakaryocytic or erythroid lineages, respectively. Though megakaryocytes constitute a population of terminally differentiated cells, under normal physiological conditions megakaryocytes do not leave the confines of the bone marrow, but shed cellular fragments of themselves to form platelets that enter blood circulation. Similar to MEPs, GMPs may give rise to myeloblasts or monoblasts. Differentiation of myeloblasts proceeds through the promyelocyte stage and other intermediate cellular stages to form the granulocytic lineage of leukocytes. Though it is known that mast cells arise from the myeloid lineage, the precise ontogeny of these cells is equivocal. It has been postulated that mast cells directly arise from CMPs, GMPs, or a theoretical shared basophil/mast cell precursor [10]. Monoblasts are precursors of the monocytic lineage that give rise to cells of the mononuclear phagocytic system (i.e. monocyte / macrophage) and professional antigen presenting cells (i.e. myeloid dendritic cells and macrophage). Monocytes travel through blood where they eventually home to organ systems throughout the body and further differentiate into a tissue-resident macrophage (e.g. a Kupffer cell of the liver).

Importantly, blockade of any one of these steps in differentiation leads to impaired cellular maturation and gradual accumulation of progenitor cells at the blocked developmental branch point [11]. A prototypical example of such an event occurring in myeloid malignancy is the mutation of the gene *CEBPA*, which encodes for a lineage-restricted transcription factor whose functionally intact expression is required for the production of GMPs from CMPs [11, 12]. Monoallelic mutations in *CEBPA* result in an N-terminal truncated protein product that exerts a dominant-negative effect on wild-type C/EBPα protein encoded by the complementary allele. As *CEBPA* expression is restricted to cells of the myelomonocytic lineage, C/EBPα-deficiency blocks granulocytic differentiation and results in the accumulation of upstream blasts [11-13]. Impaired differentiation of a clonal population of hematopoietic cells is one hallmark of MDS and acute leukemia.

Reciprocally, activating mutations in cytokine receptors and downstream signal transduction effector molecules confer dual pro-proliferation and pro-differentiation effects on cells leading to excessive production of fully functional terminally differentiated cell types [14]. Though these mutations occur in LT-HSCs, myeloproliferation of select differentiated lineages is observed due to the fact that the normal function of the mutated genes is to favor differentiation towards a particular developmental branch point or cell type (e.g. MEPs or megakaryocytes) [6, 14, 15]. When expressed in developing progenitor cells, hyperactive signaling of these mutated cytokine receptors / effectors biases hematopoietic differentiation and leads to an abundance of mature cells in the marrow or peripheral blood. Classical myeloproliferative neoplasms (MPN), such as polycythemia vera (PV), essential thrombocythemia (ET), and primary myelofibrosis (PMF), are associated with gain of function mutations in a trio of genes: *JAK2*, *MPL*, and *CALR*. Mutations in these three genes constitute the so-called "MPN-restricted driver mutations" and result in the hyper-activation of cytokine receptor/JAK2/STAT signaling pathways. PV is a neoplasm characterized by an absolute increase in red blood cell mass (erythrocytosis, often accompanied by

an increased number of leukocytes) and bone marrow hypercellularity with trilineage growth (panmyelosis) that includes erythroid, granulocytic, and megakaryocytic proliferation [8]. ET and PMF are proliferative disorders of megakaryocytes that result in either excessive production of platelets (ET) or fibrosis of the bone marrow due to release of pro-fibrotic cytokines such as TGFβ and PDGF (PMF). Interestingly, *JAK2* mutations are found in PV, ET, and PMF, while thrombopoietin receptor (*MPL*) and calreticulin (*CALR*) mutations are only found in ET and PMF. The explanation for this mutational spectrum can be explained by the biology of the encoded signaling proteins. Whereas mutation of the *MPL* and *CALR* genes converge on the selective activation of thrombopoietin receptor signaling, the V617F mutation of the receptor associated kinase *JAK2* results in activation of all JAK2-associated cytokine receptor signaling pathways – of which the erythropoietin receptor (*EPOR*), granulocyte colony-stimulating factor receptor (*CSF3R*), and thrombopoietin receptor (*MPL*) are the 3 main cytokine receptors of myelopoiesis responsible for the differentiation of erythrocytes, granulocytes, and megakaryocytes [14]. Consequently, activation of these associated cytokine receptor signaling pathways explain why the $JAK2^{V617F}$ mutation is found in all three classical MPNs and explain the ancillary features of this disease such as concomitant neutrophilic leukocytosis and eventual progression to myelofibrosis [16]. Likewise, convergence on the thrombopoietin receptor signaling pathway explains the select proliferation of megakaryocytes due to *MPL* or *CALR* mutations in ET and PMF. Increased proliferation or acquisition of a selective growth advantage is a hallmark of MPN and acute leukemia.

The existence of both myeloid neoplasms characterized by either impaired differentiation (MDS) or increased cellular proliferation (MPN) became the foundation of the two-hit model of leukemogenesis [17]. In this model mutant leukemogenic genes fall into one of two classes: class I mutant genes are typically involved in signal transduction and function to increase cellular proliferation whereas class II mutant genes are typically transcription factors / tumor suppressors

and function to block cellular differentiation. It is believed that the coexistence of both class I and class II mutant genes in a clonal population is sufficient to induce leukemic transformation to AML. This is supported by evidence that both MDS and MPN are able to transform to secondary AML. In the case of MDS, this model predicts that lesion first occurs in class II genes such as *RUNX1, P53, CEBPA, RARA,* and *RPS14,* resulting in a differentiation block that is responsible for the peripheral cytopenia(s) observed in this disease. Next, only after the secondary acquisition of mutation in one or more class I genes such as *FLT3, NRAS,* or *KIT* does MDS gain the proliferative advantage needed to transform to full blown AML [18]. Likewise, the converse appears to be true for MPNs. This model predicts that MPN derive from lesion of class I genes and subsequently transform to AML upon loss of function mutation of class II genes. Among MPNs, the best example of this transformation is seen in *BCR-ABL*$^{+}$ chronic myeloid leukemia (CML). CML begins as a chronic myeloproliferation of mature cells of the granulocytic lineage, but over time, the natural course of the disease is to progress to an accelerated phase and culminate in blast crisis in which blasts may resemble either those seen in AML or ALL. During transformation from chronic phase to blast crisis leukemic CML cells acquire mutations in the class II genes *P53* and *RUNX1* in myeloid blast crisis (AML-like) or in the genes *IKZF1* (*Ikaros*) and *CDKN2A/B* in lymphoid blast crisis (ALL-like) [19]. Additional lesions can occur to promote stem-ness such as over-expression of β-catenin – a crucial regulator of stem cell self-renewal – or a GOF mutation in *GATA2* – a negative regulator of hematopoietic stem/progenitor cell differentiation [19, 20]. Consequently, the two-hit model of leukemogenesis generally appeared to accurately describe the origins of AML and transformation of myeloid neoplasms. However, certain genes were discovered to be quite prevalent in myeloid neoplasms that did not fit neatly into either one of these to classifications. Moreover, there were gaps in this theory as approximately 50% of AML patients did not have a detectable class I mutation [21]. A revision to the classical model was needed.

Perturbation of the Epigenome in Hematologic Malignancy: Re-evaluation of the "two hit" model of leukemogenesis

As DNA sequencing and cell sorting technologies developed, it became ever increasingly apparent that a new class of genes was recurrently mutated in hematologic malignancies. These genes were predominately found to be mutated or rearranged in myeloid malignancies and acute leukemia [22], however later on almost all neoplasms of the hematopoietic and lymphoid systems harbored documented mutations, including mature neoplasms [23]. This new class of leukemogenic genes functioned to regulate the dark matter of the biology: epigenetic phenomena. These epigenetic modifier genes functioned to regulate fundamental cellular processes such as DNA methylation (*DNMT3A, TET1/2, IDH1/2*), chromatin modification (*ASXL1, EZH2, MLL*), and RNA splicing (*SF3B1, SRSF2, ZRSR2, U2AF1, U2AF2*), but how dysregulation of these processes contributed to the provocation of hematologic malignancy was unclear (and is still a matter of ongoing research) [21, 24]. Pioneering work done in transgenic mice discovered that inactivation of *Dnmt3a* in LT-HSCs resulted in the cellular expansion of this population due to de-repression of genes necessary for stem cell self-renewal [25, 26]. Seminal work in humans by Shlush *et al*. discovered that *DNMT3A* mutations give rise to a pre-leukemic state in bone marrow termed clonal hematopoiesis of indeterminate potential (CHIP) that creates "fitter" stem cells that contribute to all branches of hematopoiesis [5, 23]. Followed up by Xie *et al*. revealed numerous other genes are mutated in those with CHIP, lesion of *DNMT3A* occurs the most frequently, and this simply occurs as a consequence of aging in otherwise healthy individuals [6]. Amazingly the genes found mutated in CHIP were also found as clonal initiating mutations in the complete spectrum of known human hematologic malignancy of myeloid, T-lymphoid, and B-lymphoid lineage that including *DNMT3A, TET2, JAK2, ASXL1, TP53, GNAS, BCORL1, SF3B1, CDKN2A, IDH2, CREBBP, SETBP1,* and *CBL* among others [6]. The findings that such a disproportionate

number of epigenetic modifier genes are found mutated in CHIP and as initiating lesions in acute leukemia has now redefined the working model of leukemogenic class genes. Consequently, it appears that perturbation of the epigenome in necessary in most if not all hematologic malignancies.

Lymphoid Malignancy

All mature lymphocytes stem from development along one of three branches of lymphopoiesis radiating from the CLP: the T-cell, NK-cell, and B-cell lymphoid lineages. Mature T-cells of the peripheral blood and secondary lymphoid tissues (e.g. spleen, lymph nodes, tonsils, etc…) are classified as either helper or cytotoxic T-cells based on respective CD4 or CD8 co-receptor expression and immunological function. Following polarization of naïve helper T-cells (Th_0) by antigen and combinatorial expression of lineage-restricting transcription factors, cells are further functionally classified based on their ability to secrete certain inflammatory cytokines or express certain CD makers and include Th1, Th2, Th9, Th17, Th22, Treg (T regulatory), and Tfh (T follicular helper) [27]. In contrast to helper T-cells, cytotoxic T-cells do not predominately produce cytokine but instead produce cytotoxins such as perforin, granzyme, and granulysin to destroy the cell membranes of infected or cancerous cells (thus the moniker, "Killer T-Cell"). Of similar function and developmental origin as cytotoxic T-cells, NK cells constitute a branch of the innate immune system that act as indiscriminate sentinels surveying the body to kill any tumor cell (or pathogen) lacking MHC I expression and to assist in antibody-dependent cell-mediated cytotoxicity. The vast majority of B lymphocytes reside in the secondary lymphoid organs and constitute a panoply of cell types. Three main B-lineage lymphocytes are naïve B-cells, memory B-cells, and plasma B cells. Naïve B-cells have not yet experienced antigen nor passed through the germinal center reaction, whereas memory B-cells have experienced both. Plasma B-cells are the terminally differentiated activated antibody-secreting cells that function to neutralize pathogens. Importantly, the advantage an adaptive immune system poses for host defense is ultimately a double-edged sword in aged individuals as the very mechanisms used to generate antibody or T-cell receptor diversity can culminate in lymphomas/leukemia initiated by chromosomal translocations. It is not surprising then that a large number of mature lymphoid

neoplasms are associated with chronic pathogen exposure (e.g. *H. pyloi*) or that lymphocytes are in fact primary reservoirs of the oncogenic virus they intend to seek out and destroy (e.g. HTLV-1, 2, EBV) [28]. Similarly, the cellular developmental checkpoints necessary to create physiologically competent mature lymphocytes capable of recognizing a diverse set of antigens add several additional areas of vulnerability where a maturing cell could become susceptible to transformation. Consequently, the ontogeny of these cells intimately describes the pathological basis of disease.

In mammals, T-cell development occurs in the thymus. For this to occur, immature hematopoietic progenitor cells immigrate from the bone marrow to the thymus where they form early thymic precursors (ETP) and subsequently differentiate through a series of stages demarcated by functional and immunophenotypic changes prior maturation and egress into peripheral blood or secondary lymphoid organs. Though thymocyte development in humans and mice differ by cell surface markers, the biological stages of development are quite similar in progression and only differ by the inclusion of one additional "double negative" (DN) in mice. In thymocyte development, "double negative" refers to the absence of either mature single positive T-cell maker CD4 or CD8, thus defining DN stage cells as immature progenitors of which ETPs mark the DN1 stage of thymocyte development. The DN1 stage (Mouse: CD44+CD25−cKit+; Human: CD34+CD38^{lo}CD7−CD1a−) constitutes the most immature population of cells in the thymus and are not yet completely committed to the T-lymphoid lineage [29, 30]. The DN2 (Mouse: CD44+CD25+cKit+; Human: CD34+CD38+CD7+CD1a−) and DN3 stages (Mouse: CD44−CD25+cKit−; Human: CD34+CD38+CD7+CD1a+) signify specification and commitment to the T-cell lineage as seen by surface expression of the alpha subunit of the IL-2R (CD25) in mice and the T-lineage marker CD7 in humans [30]. Additionally, initial T-cell receptor (TCR) rearrangement occurs during these two stages and the consequences of these rearrangements determine if thymocytes will develop into γδ T-cells, αβ T-cells, or undergo apoptosis. At the DN2 stage V(D)J somatic recombination begins in the *TRD* genes and then in the DN3 stage begins in

the *TRG* gene in order to rearrange TCRδ and TCRγ and commit to γδ T-cell fate [30, 31]. If rearrangements in the *TRD* gene are nonfunctional, cells will instead proceed with somatic recombination of the *TRB* locus beginning during the transition from the DN3 to DN4 stage in mice (Mouse DN4: CD44−CD25−cKit−) and during the transition from DN3 to immature single positive (ISP) stage in humans (Human ISP: CD34+CD38+CD7+CD1a+CD4+CD8−CD3−) (Note: humans do not have a DN4 stage like mice.) [30, 31]. This process is termed β-selection, and if successful in producing a functional TCRβ receptor, thymocytes will commit to the αβ t-cell lineage, however, if unsuccessful cells will undergo p53-dependent apoptosis. From this developmental point forward, engagement of TCRβ in conjunction with pre-Tα (pre-T-cell receptor α; i.e. non-recombined TCRα) (or later with TCRα) is now the primary stimulatory factor for thymocyte development and in fact stimulates the transition from human DN3 / murine DN4 to the ISP stage (Mouse: CD4−CD8+CD3−) to the early DP stage (Mouse & Human: CD4+CD8+CD3−). (Note: in humans CD4 is expressed in ISP cells whereas in mice the CD8 co-receptor is predominantly expressed in ISP cells in most strains of mice [31]). Thus, β-selection is a major checkpoint in thymopoiesis preceded by events governed by hematopoietic-like mechanisms and followed by events driven by microenvironmental interactions with the TCR [32]. Interestingly, inappropriate passage through the β-selection checkpoint is also a common feature of many T-cell leukemias and lymphomas due to the acquisition of mutations sufficient to activate downstream mediators of TCR signaling such as LCK [33]. Likewise, loss of p53-mediated apoptosis at the DN4/ISP stage allows murine cells to proceed to the DP stage of development [34, 35]. In both mice and humans, somatic recombination of the *TRA* gene occurs during the early DP stage to produce a fully rearranged TCRα so to pair with TCRβ so that the heterodimeric TCRαβ receptor may then undergo positive and negative antigenic selection. Over 85% of thymic cells are found in the DP stage and of these, more than 95% will undergo apoptosis after having failed criteria for selection [31, 36, 37]. Surviving DP thymocytes next gain expression of surface CD3 to enter the late DP stage (Mouse & Human: CD4+CD8+CD3+) just as they are posed with the cell fate decision of becoming either

a CD4 or CD8 single positive (SP) T-cell [31]. The CD4 vs CD8 lineage choice, much like that observed in T-cell polarization, is a balancing act between two (or more) opposing lineage-biasing transcription factor complexes. Simplistically, the fate of DP thymocytes is determined by Th-POK (*ZBTB7b*) and RUNX complex family members – namely the RUNX3 (which functions as a heterodimer with the beta subunit CBFβ common to all RUNX transcription factors) [38, 39]. Relative over-expression of Th-POK compared to RUNX3 lead to increased CD4 expression, CD8 downregulation, and helper T-cell lineage commitment. Inversely, relative over-expression of RUNX3 compared to Th-POK result in increased CD8 expression, CD4 downregulation, and commitment to the cytotoxic T-cell lineage. This balancing act is first accomplished through the reciprocal inhibitory binding each of Th-POK and RUNX3 in the other's promoter such that transcription of the opposing gene is repressed. This allows Th-POK to directly regulate *RUNX3* transcription and vice versa. Furthermore, RUNX3 binds both the CD8 enhancer and CD4 silencer to bias expression of these genes whereas Th-POK functions to directly repress the CD8 locus and interrupt the physical interaction of RUNX3 with the CD4 silencer [38]. Positive transcriptional regulators of *Th-POK/ZBTB7b* include TCF-1 (TCF7), LEF1, MYB, GATA3, and TOX, whereas negative regulators include BCL11B and MAZR (in addition to RUNX family members) [38, 39]. Positive transcriptional regulators of *RUNX3* include ETS1 and IL-7 signaling, whereas negative regulators include Th-POK and SOCS [38-40].

Malignancies that manifest as neoplasms of mature lymphoid cells are classified as lymphoma or lymphocytic leukemia depending on the anatomical compartment in which the neoplasms presents (i.e. lymph node vs. blood). Lymphoid neoplasms are by definition clonal proliferations of a specific subtype of B, T, or NK-cell. Two processes characterize lymphoid maturity: 1) degree of cellular differentiation and 2) the immigration of cells from primary lymphoid tissues (e.g. bone marrow and thymus) to take up new residence in peripheral / secondary lymphoid structures (e.g. lymph node, spleen, tonsil, etc…). Thus, most mature lymphoid

neoplasms can be demarcated from immature neoplasms through analysis of the morphology, immunophenotype, and anatomical location of the presenting lymphoproliferation. Approximately 85-90% of mature lymphoid neoplasms are of B-cell origin, whereas the remaining 10-15% derives from T-cell and NK-cell origins.

Mature lymphoproliferative disorders of T-cell origin are a heterogeneous group of rare neoplasms that display functional and immunophenotypic properties of post-thymic T-cells. As all post-thymic T-cells are technically peripheral T-cells, all mature T-cell neoplasms are classified under the general umbrella term peripheral T-cell lymphoma (PTCL). Despite the term PTCL, these neoplasms do not always present as a lymphoma, but may also present as a T-cell leukemia (e.g. T-cell large granular lymphocytic leukemia (T-LGL leukemia); T-cell prolymphocytic leukemia (T-PLL)) or a hybrid leukemia/lymphoma (e.g. adult T-cell leukemia/lymphoma (ATLL); Sézary Syndrome). Likewise, T-cell lymphomas do not always manifest in secondary lymphoid tissues (e.g. lymph nodes), but can present routinely in extranodal locations such as the liver, skin, small bowel, and nasal cavity (hepatosplenic T-cell lymphoma (HSTCL), mycosis fungoides, enteropathy-associated T-cell lymphoma (EATL), and extranodal NK/T-cell lymphoma, nasal type (NKTCL), respectively). Nodal PTCLs include angioimmunoblastic T-cell lymphoma (AITL), ALK-positive anaplastic large cell lymphoma (ALCL), ALK-negative ALCL, and peripheral T-cell lymphoma, not otherwise specified (PTCL-NOS). Of the rare mature T-cell neoplasms, these nodal PTCLs occur most frequently and AITL has been the best molecularly profiled neoplasm among this group of lymphoproliferative disorders.

AITL is an aggressive lymphoma detected in aged individuals that accounts for ~20% of all PTCL diagnoses [41]. This neoplasm derives from the transformation of follicular helper T-cells (T_{FH}) that are located in B-cell follicles of secondary lymphoid tissues [41]. T_{FH}-cells express a distinct immunophenotypic signature of CD279/PD1, CD10, BCL6, CXCL13, ICOS, SAP, and CCR5 that is also expressed by AITL, reflecting the ontogeny of this neoplasm [9]. AITL is

characterized by recurrent genetic abnormalities in epigenetic regulators (*TET2, DNMT3A, IDH2*^{R172}) and T-cell receptor signaling factors (*RHOA, CD28, PLCG1, FYN,* fusion *ITK-SYK,* fusion *CTLA4-CD28*), of which the small GTPase RHOA is most frequently mutated [9, 41-44]. Under the current working model of AITL pathogenesis, lesion of *TET2* and/or *DNMT3A* occur during very early hematopoiesis in HSCs leading to enhanced self-renewal, stem cell expansion, and clonal hematopoiesis [5, 6, 25, 45]. These lesions in *TET2* and *DNMT3A* are initiating, but, perhaps in and of themselves, do not impact the lineage of the neoplasm in humans as the mutant alleles can be readily detected in the neoplastic T-cell population as well as unaffected cells of the B-lymphoid lineages [23, 45]. As previously described, *DNMT3A* and *TET2* are the two most frequently mutated genes in CHIP – a pre-neoplastic condition associated with aging. Similar to mutations found in CHIP/MDS/AML, lesion of *TET2* is inactivating and often leads to loss of the gene product through a nonsense or frameshift mutation. Unlike the dominant negative *DNMT3A*^{R882} mutation found in in MDS/AML, the majority of *DNMT3A* mutations that occur in PTCL (including AITL) are inactivating due to a nonsense or frameshift mutation [23, 46]. *DNMT3A* mutations in PTCL have been found to be predominantly monoallelic, which is in stark contrast to early thymic precursor T-ALL in which *DNMT3A* mutations are exclusively biallelic [46].

Acquisition of mutations in *IDH2* and *RHOA* next occur in pre-neoplastic cells, and it is believed that this biases the hematopoietic development of cells towards that of the T-lymphoid lineage [47]. Heterozygous missense mutation of *IDH2* occurs exclusively at amino acid R172 in AITL to produce a dominant negative IDH2^{R172X} protein with neomorphic catalytic activity. The isocitrate dehydrogenase enzymes – IHD1 and IDH2 – are important members of the tricarboxylic acid cycle that normally function to convert isocitrate to α-ketoglutarate (α-KG). Distinct gain-of-function mutations (e.g. IDH1^{R132X}, IDH2^{R140X}, and IDH2^{R172X}) confer theses enzymes with the ability to produce the oncometabolite 2-hydroxyglutarate (2-HG) instead of α-KG. Importantly, 2-

HG is a competitive inhibitor of α-ketoglutarate-dependent dioxygenases such as TET family enzymes, including TET2, and histone demethylases [48]. Consequently, oncogenic IDH2^{R172} can induce repressive DNA hypermethylation (through inhibition of TET2 by 2-HG) and histone hypermethylation [42]. It is interesting to note that *IDH1/2* mutations also occur in myeloid neoplasms. However, several interesting caveats exist regarding the role of oncogenic IDH2 in AITL and myeloid neoplasms. Firstly, the position of the missense mutation resulting in an amino acid change is different between these two groups of diseases. In myeloid neoplasms, mutations occur at *IDH1*^{R132} or *IDH2*^{R140}, but rarely occur at *IDH2*^{R172}, whereas the *IDH2*^{R172} mutation appears to occur exclusively in AITL [41, 42, 45, 49]. Although each IDH1/2 mutant is capable of producing the 2-HG oncometabolite, these observations suggest the existence of an undiscovered, but functionally relevant difference between these mutations necessary for the propagation of myeloid or lymphoid neoplasms. Secondly, in contrast to myeloid neoplasms where mutations in *IDH1/2* are mutually exclusive of TET2 mutations, *IDH2*^{R172} mutations directly correlate with *TET2* mutations in AITL [42, 49, 50]. The co-occurrence of *TET2* and *IDH2* mutations in AITL emphasizes DNA hypermethylation as an apparently essential pathway in this disease. However, this interpretation is certainly confounded by the fact that loss of *DNMT3A* is also occurs in AITL and biochemically results in DNA hypomethylation. Despite the specific biochemical reactions catalyzed, on the larger biological scale, functional studies appear agree that loss of either TET2 and/or DNMT3A favor oncogenesis in the hematopoietic compartment. Speculatively, it is conceivable that TET2 and DNMT3A do not always act upon the same regions of the genome and thus in nonmalignant cells create topological boundaries marked by DNA hypomethylation and hypermethylation that erode upon loss (or biochemical inhibition) of one or both of these enzymes [51-53].

T-cell lineage-impactful mutations that favor increased TCR signaling constitute the final major group of genes found mutated in AITL (and other PTCLs). Among members of the TCR

signaling pathway, the *RHOA* gene is by far the most frequently mutated and is detected in 60-70% of all AITL and PTCLs [44, 54]. RHOA is a small GTPase functionally downstream of TCR-LAT-VAV1/2 signaling that vacillates between an active GTP-bound and inactive GDP-bound state to modulate cytoskeletal remodeling [55, 56]. Highly recurrent G17V mutations occur in *RHOA* and confer dominant negative activity to the $RHOA^{G17V}$ GTPase, resulting in transformation by a mechanism that is not yet fully elucidated [55]. Additionally, it is important to note that by targeted deep sequencing of AITL and lymphomas derived from follicular T-helper cells, Vallois *et al.* demonstrated that half of lymphomas harbored mutually exclusive activating mutations in other TCR pathway genes [55]. These signaling genes include *PLCG1* (14.1%), CD28 (9.4%), *PI3K* elements (7%), *CTNNB1* (6%), and *GTF2I* (6%) [55, 56]. Multiple gain of function mutations in TCR signaling pathway genes implicate this pathway as a major signaling nexus important in the pathogenesis of T-cell lymphoma. As the allelic frequencies of these mutations are low in tumor samples, it is at present unclear whether this is a true transformative event in T-cell lymphomagenesis or an adaptation of an existing neoplasm to circumvent the need for chronic antigenic signaling through engagement of the TCR by an exogenous antigen.

The epigenetic landscape of AITL and PTCL is a fascinating frontier that has not yet been explored adequately in human lymphoma samples due to numerous technical hurdles and issues with specimen purity. Future work profiling the DNA methylation, DNA hydroxymethylation, and chromatin modifications in PTCL (and the progenitor cells giving rise to this tumor) will be invaluable to the continued scientific understanding of how perturbation of the epigenome contributes to T-cell lymphoma.

Of mature B-cell neoplasms, chronic lymphocytic leukemia / small lymphocytic lymphoma (CLL/SLL) is extremely common and is the most frequently diagnosed leukemia in the United States [57]. CLL/SLL is a chronic, indolent, low-grade lymphoproliferation of mature B-cells that most frequently occurs in elderly patients, presenting as either a leukemia or lymphoma

[58]. Although one singular disease process, by convention CLL/SLL is referred to as CLL when presenting in the blood and bone marrow (leukemia) or SLL when presenting in the lymph nodes (lymphoma). CLL/SLL (hereafter referred to singularly as CLL) is universally preceded by monoclonal B-cell lymphomatosis (MBL) – a "pre-neoplastic" precursor lesion found in up to 12% of aged individuals and defined as the presence of a clonal B-cell population in the peripheral blood with fewer than 5×10^9 / L clonal B-cells in the absence of any other signs of a lymphoproliferative disorder such as lymphadenopathy or organomegaly [9, 59, 60]. Immunophenotypic analysis of MBL cells describes these clonal expansions as CLL-like phenotype ($CD5^+CD23^+$), atypical CLL phenotype ($CD5^+CD23^-CD20^{bright}$), or non-CLL phenotype ($CD5^-$) [61, 62]. Quantitation of the lymphocytosis further classifies the lymphoproliferation as "low-count" (clonal B-cells < 0.5 x 10^9 / L) and "high-count" MBL (0.5 x 10^9 / L < clonal B-cells < 5×10^9 / L). Importantly, a correlation with low-count MBL and the development of frank leukemia has not as of yet been established, whereas high-count MBL with CLL-like (or atypical CLL) phenotype is directly correlated with progression to CLL requiring treatment at a rate of 1-2% per year [61, 63]. (Note: High-count MBL with non-CLL phenotype is postulated to be a precursor of marginal-zone lymphoma and consequently the classification "clonal B lymphocytosis of marginal-zone origin" (CBL-MZ) has been proposed to better reflect the origin of this precursor lesion [64].) Thus, high-count MBL with CLL-like phenotype is a pre-malignant state that precedes the presentation of frank CLL.

CLL is a highly heterogeneous disease characterized by extensive inter- and intra-tumoral genetic heterogeneity, chromosomal aberration, epigenetic dysregulation (particularly DNA hypomethylation), aberrant transcriptional splicing, and clonal evolution [65-69]. Approximately 80% of these tumors harbor recurrent cytogenetic anomalies that also are detectable in MBL, indicating that cytogenetic lesions classically associated with CLL occur in the most insipient stages of disease [70]. Classical CLL cytogenetic anomalies include chromosomal deletions 13q− (Type I del13q14 = *DLEU2* / *miR-15a* / *miR-16-1*, *DLEU7*. Type II del13q14 = Type I + *RB1*),

11q− (*ATM, BIRC3*), 17p− (*TP53*), and chromosomal duplication +12 [71, 72]. Of which, the most common anomaly is a 1 Mb deletion at 13q14 (13q−) that is found in ~50% of CLL patients at diagnosis and results in loss of the *miR-15a* / *miR-16-1* mircoRNA cluster located in the intron of *DLEU2* as well as the *DLEU7* gene [71]. These genes behave as tumor suppressors whereby *miR-15a* / *miR-16-1* post-transcriptionally inhibit *BCL2* and *MCL1* and DLEU7 protein inhibits activation of NFκB and NFAT by blocking APRIL (a proliferation-inducing TNF ligand) signaling through TACI, a TNF superfamily receptor [73, 74]. Thus, this lesion imparts an anti-apoptotic and pro-survival property to the neoplasm. Interestingly, biallelic loss of the 13q14 locus occurs in ~30% of patients with previously diagnosed monoallelic 13q− karyotype CLL, indicating that the APRIL-TACI-NFκB/NFAT signaling axis and the anti-apoptotic effects of BCL-2 confer a significant selective advantage to neoplastic cells. Additionally, with less frequent incidence, a "Type II" del13q14 alteration can occur deleting ~3 Mb to also include loss of the tumor suppressor gene *RB1* [72]. Alterations affecting chromosome 11 (11q−) present in 5-20% of patients at diagnosis and can be of variable length, however these lesions result invariably in the monoallelic loss of the gatekeeper tumor suppressor gene *ATM* [71]. Deletion of 17p (17p−) is infrequent at diagnosis (3-8% of patients) and results in the loss of tumor suppressor *TP53* – a predictor of chemorefractoriness, poor overall survival, and an extremely unfavorable patient outcome [71]. Patients with 17p− CLL soon acquire mutations in the second *TP53* allele, develop aggressive chemorefractory CLL, or undergo Richter's Transformation to diffuse large B-cell lymphoma (DLBCL), and then succumb to their disease. Loss of *ATM* and *TP53* compromise the DNA damage response and implicate genomic instability in the pathogenesis of select CLL tumors. Similarly, loss of *RB1* and *TP53* implicate cell cycle deregulation as an additional means of acquiring a selective advantage in some CLL tumors.

Sequencing efforts to characterize the mutational landscape of the CLL genome have revealed few common mutations and numerous low frequency mutations, reflecting the highly

heterogeneous nature of this disease. The most frequently mutated genes in CLL include *SF3B1*, *NOTCH1*, *ATM*, *TP53*, and *MYD88*, however none of these genes are mutated at an incidence greater than 10-25% of patient cases [65, 66]. Of these, *SF3B1* – a member of the U2 spliceosome – is the most commonly mutated (50% of mutations occurring at K700E) [69]. The functional consequences of these mutations result in a neomorphic SF3B1 mutant protein that recognizes an alternative intronic branch point sequence and cryptic 3' splice site, resulting in aberrant posttranscriptional processing of pre-mRNA [75, 76]. Approximately 50% of aberrantly spliced mRNA transcripts undergo nonsense-mediated decay, resulting in functional inactivation of these gene products [75]. In those aberrant transcripts that do not undergo nonsense-mediated decay, gain of function, change of function, and loss of function can occur. In CLL, mutant SF3B1-induced aberrant splicing of *DVL2* results in a neomorphic protein that enhances intracellular NOTCH signaling [69]. *MYD88* is an adaptor protein necessary for Toll-like receptor signaling-induced activation of NFκB that, in turn, regulates the production of anti-apoptotic molecules. Oncogenic gain of function mutations in *MYD88* result in constitutive activation of NFκB pathway, conferring resistance to apoptosis [77]. In addition to recurrently mutated candidate driver genes, numerous non-recurrent mutations occur in CLL and cluster in the Wnt / β-catenin signaling pathway [78]. These mutations function to activate this signaling pathway and stabilize β-catenin whose direct transcriptional targets are *MYC* and *CCND1*. Consequently, deregulation of Wnt signaling contributes to leukemogenesis through proliferative signals. Notably, *MYC* is also a downstream transcriptional target of NOTCH signaling, implicating *MYC* over-expression as a convergent oncogenic mechanism shared between *SF3B1*, *NOTCH1*, or *WNT* signaling pathway mutated CLL.

High resolution examination of DNA methylation changes in human CLL demonstrates reproducibly that epigenetic lesions in human CLL are notable for DNA hypomethylation of functional genomic regulatory elements such as gene bodies, enhancers, promoters, and

transcription factor binding sites [67, 68]. Extensive characterization of the DNA methylomes of 139 patients with CLL with mutated or unmutated IGHV and of several mature B-cell subpopulations through of whole-genome bisulfite sequencing and high-density microarrays discovered that DNA hypomethylation in gene bodies, targeting mostly enhancer sites, was the most frequent difference between the two molecular subtypes of CLL and normal B cells [67]. Likewise, a recent study reported DNA methylation changes in transcription factor binding sites, leading to transcription factor dysregulation in CLL, and programming imbalances the normal B cell epigenetic program [68]. Although no mutations in the DNMT3A locus have been reported in CLL or other B-cell malignancies, global gene expression profiling of 448 human CLL samples revealed that DNMT3A and DNMT3B are among the most under-expressed genes (1^{st} percentile), implying a functionally deficient state [79]. Likewise, the TCL1A oncoprotein, which is highly over-expressed in MBL, CLL, and a panoply of other B-cell neoplasms, can biochemically inhibit the activity of DNMT3A and DNMT3B through direct protein-to-protein interaction [62, 68, 80]. These findings imply that epigenetic dysregulation of DNA methylation is fundamental to the pathogenesis of CLL.

Introduction to DNA Methylation and DNA Methyltransferases[1]

Cytosine methylation of DNA is an epigenetic modification affecting gene transcription and the integrity of the mammalian genome. The three catalytically active DNA methyltransferases in mammalian cells are DNMT1, DNMT3A, and DNMT3B. These enzymes are responsible for establishment and maintenance of DNA methylation during normal development and during mitotic cell division. Promoter methylation typically results in transcriptional repression of genes and plays a role in various normal physiologic processes, such as differentiation and hematopoiesis [82, 83].

One of the main observations that contributed to an interest in studying this phenomenon came from studies that discovered that in virtually all types of cancer aberrantly increased methylation in gene promoters was associated with transcriptional inhibition [84]. As a result, aberrant promoter hypermethylation resulting in silencing of tumor suppressor genes in cancer has been a major topic of numerous studies over the past 30 years. Such efforts not only resulted in identification of a number of epigenetically repressed tumor suppressor genes in hematologic malignancies, such as VHL, p16, and MLH1 [84], but also provided a conceptual approach to the treatment of cancer. Therefore, use of hypomethylating agents can be a valuable approach in anticancer therapy as reversal of DNA methylation may lead to reactivation of tumor suppressor genes and antagonize aberrant tumor proliferation and survival. Epigenetic studies in hematologic malignancies have consequently resulted in the approval of hypomethylating agents for the treatment of several malignancies, including myelodysplastic syndrome (MDS) and acute myeloid leukemia (AML). Numerous efforts are undergoing to refine hypomethylating agents and to test their efficacy in combination with other drugs targeting epigenetic changes.

[1] Material used in this section is in part from:

81. Upchurch, G.M., S.L. Haney, and R. Opavsky, *Aberrant Promoter Hypomethylation in CLL: Does It Matter for Disease Development?* Front Oncol, 2016. **6**: p. 182.

The recent emergence of new high-resolution methylation profiling techniques, such as whole-genome bisulfite sequencing (WGBS), have revealed that the methylome of cancer cells frequently contains regions of the genome that are hypomethylated relative to their normal cellular counterparts. As DNA hypermethylation correlates with gene silencing, it remains to be determined if the opposite is true: does DNA hypomethylation correlate with increased gene expression at differentially methylated loci? If so, whether or not such deregulated expression contributes to the initiation and progression of hematologic malignancies is currently unresolved and is actively under investigation.

Role of DNA Methyltransferases in Mouse Normal and Malignant Hematopoiesis

Two activities for DNA methyltransferases are important concerning genome-wide methylation patterns: de novo and maintenance. Early studies suggested that Dnmt3a and Dnmt3b are de novo enzymes capable of recognizing unmethylated DNA and catalyze the addition of a methyl group to cytosine [85]. By contrast, Dnmt1 plays a role in maintenance of methylation profiles by recognizing methylated DNA and adding a methyl mark to the newly produced DNA strand during replication [86]. In contrast to these early studies, the modern view is that all Dnmts possess both *de* novo and maintenance methylation activity. The role of Dnmts in generating and maintaining methylation profiles in cancer is, however, less clear. In regards to gene promoters, from the most simplistic conceptual point of view, de novo methylation activity could be considered oncogenic because aberrant activity can generate new methylation marks in promoter sequences and inactivate genes with tumor suppressor functions. By contrast, maintenance methylation activity could be considered to be tumor suppressive because it can safeguard the integrity of the methylome and prevent inappropriate gene activation. However, a number of studies in which Dnmts were genetically inactivated have painted a rather different picture.

The most studied Dnmt in mouse hematopoiesis is Dnmt3a. Interest in studying the role of Dnmt3a in normal hematopoiesis was fueled by findings that in a number of hematologic malignancies of myeloid and T-cell origin Dnmt3a was mutated primarily in the catalytic domain, suggesting that methyltransferase activity is critical to prevent tumor development [46]. Subsequently, several research groups have used mice to conditionally inactivate Dnmt3a in hematopoietic stem cells and early progenitors (HSPCs) by the Mx1-Cre transgene. Original reports suggested that inactivation of Dnmt3a results in a differentiation block and accumulation of progenitor cells; however, this phenotype was mainly observed upon serial bone marrow transplantation [25]. The first report demonstrating that Dnmt3a plays a tumor suppressive role in the prevention of hematologic malignancy was published by Peters et al. [79], which showed that conditional inactivation of Dnmt3a in HSPCs using EμSRα-tTA;TetoCre;Dnmt3a$^{fl/fl}$;Rosa26 LOXP$^{EGFP/EGFP}$ quadruple transgenic mice (designated as Dnmt3a$^{\Delta/\Delta}$ mice) results unexpectedly in the development of a CLL-like disease after 1 years' time. This finding was surprising in view of a lack of genetic alterations in the DNMT3A locus in B-cell malignancies. Since this time, other groups have reported that Dnmt3a loss in HSPCs results in myriad types of malignancies, such as myeloproliferative disorders, AML, T-cell acute lymphoblastic leukemia (T-ALL), and B-cell acute lymphoblastic leukemia (B-ALL) [87, 88]. Differences in phenotypes observed upon loss of Dnmt3a could stem from the different genetic background of mice used in these studies (**Table 1**) or the different properties of transgenes used to conditionally delete Dnmt3a alleles.

In addition, Dnmt3a loss can collaborate with gain of function mutant c-kitD814V to induce B-ALL, T-ALL, and mastocytosis with myeloid blasts [89], and with KrasG12D/+ to promote progression of juvenile and chronic myelomonocytic leukemia (CMML) [90]. However, under some circumstances Dnmt3a has acted as an oncogene by promoting the development of hematologic malignancies. For example, upregulation of Dnmt3a promoted AML/ETO-induced leukemia through de novo hypermethylation [91] and a methylation-independent repressor function

of Dnmt3a enhanced T-cell lymphomagenesis [92]. Despite these last two examples, the overwhelming body of work in the field suggests that Dnmt3a acts as a tumor suppressor in various types of hematologic malignancies.

Due to the lack of genetic alterations in Dnmt3b in human malignancies, the role of this enzyme has been less well studied in mouse hematopoiesis. Loss of Dnmt3b does not appear to affect normal hematopoiesis but still has both overlapping and specific functions in HSPCs [26, 93]. However, Dnmt3b has been shown to behave as a tumor suppressor gene in mouse models of Myc-induced T- and B-cell lymphomas [93, 94]. Loss of Dnmt3b also accelerated development of CLL in Dnmt3a$^{\Delta/\Delta}$ mice and also promoted the development of T-cell lymphomas [79].

Taken together, these data clearly show that Dnmt3a and Dnmt3b possess tumor suppressor functions in the prevention of the vast majority of hematologic malignancies (**Table 1**). These data further demonstrate that tumor development is accompanied by loss of methylation in promoter regions, suggesting that Dnmt3a and Dnmt3b may have a role in cancer-specific maintenance methylation of specific loci (**Table 1**). Given the large magnitude of promoter hypomethylation observed in various mouse hematologic malignancies [79, 87, 88, 95], it is possible that loss of Dnmt3a and Dnmt3b maintenance activity drives tumorigenesis due to the inappropriate expression of genes normally silenced.

In contrast to most studies on Dnmt3a and Dnmt3b, decreased levels of the bona fide maintenance enzyme Dnmt1 suppressed MYC-induced T-cell lymphomagenesis [96] (**Table 1**). In addition, Dnmt1 is critical for tumor maintenance in AML induced by MLL-AF9 overexpression or in B-cell leukemia induced by combined overexpression of Myc and Bcl2 [97, 98]. However, there is also evidence that Dnmt1 can play a tumor suppressor role in prevention of T-cell lymphomas. Dnmt1chip/− mice that carry a hypomorphic allele (Dnmt1^{chip}) almost exclusively (91%) develop T-cell lymphomas [99]. Thus, Dnmt1 seems to possess an oncogenic function by promoting cellular survival of both normal and tumor cells (**Table 1**).

Mouse Models of CLL, PTCL, & DNMTs in Hematologic Malignancy

Table 1 | Summary of studies performed in mouse models in which Dnmts were ablated in hematopoietic cells

Transgenic model	Background	Phenotype	ONC/TS	Promoter hypomethylation	Reference
Mx1-Cre; Dnmt1$^{fl/fl}$	Unknown	Bone marrow failure	N/A	Unknown	Bröske et al. (11)
Dnmt1$^{chip/-}$	Unknown	T-cell lymphoma	TS	Unknown	Gaudet et al. (12)
Dnmt1$^{chip/-}$; retroviral Myc-Bcl2	Unknown	Myeloid and T-lymphoid leukemia. Defects in B-cell development	N/A	Unknown	Bröske et al. (11)
MLL-AF9; Mx1-Cre; Dnmt1$^{fl/+}$	C57Bl/6	AML	ONC	Unknown	Trowbridge et al. (13)
EμSRα-tTA; Teto-Cre; Teto-MYC; Dnmt1$^{fl/fl}$; Rosa26 LOXP$^{EGFP/EGFP}$	FVB/N	T-cell lymphoma	ONC	Yes	Peters et al. (14)
Mx1-Cre; Dnmt3a$^{fl/fl}$	C57Bl/6	HSC expansion and impaired differentiation	N/A	Yes	Challen et al. (7)
Mx1-Cre; Dnmt3a$^{fl/fl}$; Dnmt3b$^{fl/fl}$	C57Bl/6	HSC expansion and impaired differentiation	N/A	Yes	Challen et al. (15)
EμSRα-tTA; Teto-Cre; Dnmt3a$^{fl/fl}$; Rosa26 LOXP$^{EGFP/EGFP}$	FVB/N	CLL, PTCL	TS	Yes	Peters et al. (8)
EμSRα-tTA; Teto-Cre; Teto-MYC; Dnmt3a$^{fl/fl}$; Rosa26 LOXP$^{EGFP/EGFP}$	FVB/N	T-cell lymphoma	ONC	Yes	Haney et al. (16)
Mx1-Cre; Dnmt3a$^{fl/fl}$	C57Bl/6	MDS, AML	TS	Unknown	Celik et al. (17)
Mx1-Cre; Dnmt3a$^{fl/fl}$	C57Bl/6	MDS, AML, PMF, CMML, T-ALL, B-ALL, ETP	TS	Yes	Mayle et al. (9)
Mx1-Cre; Dnmt3a$^{fl/fl}$	C57Bl/6	MDS/MPN	TS	Yes	Guryanova et al. (18)
Mx1-Cre; Dnmt3a$^{fl/fl}$; Kras$^{LSL\ G12D/+}$	C57Bl/6	AML	TS	Unknown	Chang et al. (19)
Flt3$^{+/ITD}$; Mx1-Cre; Dnmt3a$^{fl/fl}$	C57Bl/6	AML	TS	Yes	Meyer et al. (20)
EμSRα-tTA; Teto-Cre; Teto-MYC; Dnmt3b$^{fl/fl}$; Rosa26 LOXP$^{EGFP/EGFP}$	FVB/N	T-cell lymphoma	TS	Yes	Hlady et al. (21)
EμSRα-tTA; Teto-Cre; Dnmt3b$^{fl/fl}$; Rosa26 LOXP$^{EGFP/EGFP}$	FVB/N	No phenotype	N/A	Unknown	Hlady et al. (21)
Eμ-Myc;Dnmt3b$^{+/-}$	Unknown	B-cell lymphoma	TS	Unknown	Vasanthakumar et al. (22)
Mx1-Cre;Dnmt3b$^{fl/fl}$	C57Bl/6	No phenotype	N/A	Unknown	Challen et al. (15)

Unknown indicates that data were not currently available. ONC and TS stand for oncogene and tumor suppressor, respectively.
CLL, chronic lymphocytic leukemia; PTCL, peripheral T-cell lymphoma; HSC, hematopoietic stem cell; MDS, myelodysplastic syndrome; AML, acute myeloid leukemia; T-ALL, T-cell acute lymphoblastic leukemia; B-ALL, B-cell acute lymphoblastic leukemia; CMML, chronic myelomonocytic leukemia; PMF, primary myelofibrosis; ETP, early thymic progenitor acute lymphoblastic leukemia.

Questions Addressed In Dissertation

DNMT3A is a master epigenetic regulator of benign and malignant hematopoiesis. In human hematologic malignancy, mutations in *DNMT3A* occur in neoplasms of the myeloid and T-lymphoid lineages. In AML, ~60% of *DNMT3A* mutations are heterozygous missense mutations that result in the production of a dominant negative DNMT3A^{R882} protein [46]. The dominant negative DNMT3A^{R882} protein possesses only 20% methyltransferase activity and inhibits the tetramerization of wild type DNMT3A that is necessary for fully functional DNA methyltransferase activity [100]. Unlike myeloid malignancy, the dominant negative R882 mutation is rare in neoplasms of the T-lymphoid lineage (less than 20% of mutations), whereas nonsense, frameshift, and gene deletions account for approximately 50% of the *DNMT3A* lesions found in T-cell leukemia and lymphoma [46, 101]. Interestingly, biallelic (homozygous) mutations occur in T-cell acute lymphoblastic leukemia (T-ALL), whereas monoallelic (heterozygous) mutations occur in mature T-cell lymphomas. Notably, in each of these neoplasms there also exist mutations of other epigenetic regulators (e.g. *TET2, IDH1/2,* etc…), implying that epigenetic dysregulation is fundamental to the pathogenesis of these malignancies. These observations in humans, led us to ask several fundamental question.

Firstly, does inactivation of *DNMT3A* provoke hematologic malignancy? If so, how does this come to be? Secondly, as DNMT3A normally functions to maintain and establish DNA methylation, how does the DNA methylome change after loss of *DNMT3A*? Thirdly, as *DNMT3A* mutations are most frequently heterozygous, is *DNMT3A* a haploinsufficient tumor suppressor in hematologic malignancy? Fourthly, considering the difference in the spectrum of *DNMT3A* mutations between myeloid and lymphoid malignancies, does the number of mutated alleles affect the lineage of the resulting neoplasm?

To address these questions, we chose to use an *in vivo* approach in mice that examined the effects of *Dnmt3a* gene deficiency in HSPCs. To dissect the biological consequences of

homozygous and heterozygous *Dnmt3a* inactivation in malignant hematopoiesis, we generated *Dnmt3a* homozygous null (*Dnmt3a*$^{\Delta/\Delta}$) and *Dnmt3a* heterozygous deficient (*Dnmt3a*$^{+/-}$) mice and compared the presentations of hematologic malignancies between cohorts. Here we present that bi-allelic inactivation of *Dnmt3a* results in the presentation of mature lymphoid neoplasms resembling CLL (88% penetrance) and CD8+ PTCL (41% penetrance). In contrast, mono-allelic inactivation of *Dnmt3a* results in the presentation of CLL and PTCL at reduced penetrance (47% & 10%, respectively) and, rarely, a MPN (10% penetrance).

Consequently, inactivation of *Dnmt3a* in murine HSPCs does indeed provoke hematologic malignancy of B-lymphoid, T-lymphoid, and myeloid lineage. These neoplasms result from inactivation of one *Dnmt3a* allele without loss of heterozygosity, indicating that *Dnmt3a* is a haploinsufficient tumor suppressor. In contrast to monoallelic inactivation, we did not observe myeloid neoplasms following biallelic inactivation of *Dnmt3a*. This evidence agrees with that found in humans and suggests that either biallelic inactivation of Dnmt3a biases the lineage of the resulting neoplasm or, alternatively, myeloid neoplasms cannot exist without at least one wild type *Dnmt3a* allele.

To better understand the how loss of *Dnmt3a* results in the development of lymphoid neoplasms, particularly PTCL, we present here data examining the DNA methylome and transcriptome of *Dnmt3a*-deficient PTCL. As we are particularly interested in deregulation of gene expression, we investigated changes in promoter DNA methylation that may account for transcriptional upregulation of putative oncogenes as a possible mechanism explaining cellular transformation of T-lymphocytes. We present this data along with novel findings implicating deregulation of p53 in the pathogenesis of PTCL.

CHAPTER 2:

DNMT3A$^{\Delta/\Delta}$ AND DNMT3A$^{+/-}$ MICE DEVELOP CLL

Dnmt3a$^{\Delta/\Delta}$ and Dnmt3a$^{+/-}$ mice develop CLL

DNMT3A is a DNA methyltransferase capable of epigenetically repressing gene expression via methylation of genomic regulatory elements such as promoters and enhancers. Mutation or inactivation of epigenetic regulators, such as *DNMT3A*, is now known to be initiate the development of hematologic malignancies [6]. *DNMT3A* is frequently mutated in human myeloid and T-cell malignancies [22, 46, 102] and is functionally inactivated in mature B-cell neoplasms due to under-expression or biochemical inhibition [67, 79, 80]. In mice genetically deficient for *Dnmt3a*, we observe the development of leukemia resembling Chronic Lymphocytic Leukemia (CLL) and characterized by expansion of B220+CD19+CD5+ cells (B1) in bone marrow, blood, and spleen, demonstrating that *Dnmt3a*-deficiency plays a role in the transformation of hematopoietic cells of the B-cell lineage.

To test the tumor suppressor function of DNMT3A in the development of CLL, we compared cohorts of mice lacking one or two alleles of *Dnmt3a* in hematopoietic stem cells (HSCs). *Dnmt3a*–/– mice developed terminal CLL by 12 months of age whereas *Dnmt3a*+/– mice developed terminal CLL with reduced penetrance by 18 months of age without loss or mutation of the second allele, indicating that DNMT3A is a haplo-insufficient tumor suppressor [95]. Interestingly, *Dnmt3a*+/– mice present with pre-leukemic monoclonal B-cell lymphomatosis (MBL; 2-20% B1 cells per peripheral blood leukocytes) prior to transformation to CLL. This mirrors the development of CLL in humans, where MBL precedes malignant transformation. MBL is detected in ~5% of people over 40 years of age and 1-2% of these individuals will progress to CLL requiring treatment per year [63].

To understand the effect of *Dnmt3a*-deficiency on the CLL epigenome, we employed whole-genome bisulfite sequencing to compare *Dnmt3a*–/– and *Dnmt3a*+/– CLL with wild-type B1 cells. We observed a *Dnmt3a* gene dose-dependent decrease in genome-wide DNA methylation with hypomethylation occurring particularly in promoters and gene bodies – an epigenetic pattern

also observed in human CLL [67, 95]. As bi-allelic inactivation of *Dnmt3a* correlated with a greater magnitude of hypomethylation and an earlier onset of CLL compared to mono-allelic inactivation of *Dnmt3a*, we reason that the tumor suppressor function of DNMT3A is directly proportional to its catalytic activity and the ability to methylate DNA. As such, we reason that de-repression of gene transcription due to hypomethylation of functional genomic elements may contribute to the pathogenesis of this disease. As promoter hypermethylation is transcriptionally repressive and promoter hypomethylation is associated with a transcriptionally permissive state, we chose to profile genes hypomethylated and over-expressed relative to wild-type B1 cells. Using whole-genome bisulfite sequencing and RNA-seq, we identified 26 genes differentially hypo-methylated and over-expressed in both Dnmt3a–/– and Dnmt3a+/– CLL (HOC genes) relative to wild-type control B1 cells. Ingenuity pathway analysis of these HOC genes identified enrichment for hematological development, immunological disease, and cancer.

Among these 26 HOC genes several promising candidate genes such as *Pvt1* and *Pdcd1lg2* (PD-L2) were identified that may directly contribute the pathogenesis of CLL in mice. *Pvt1* is a long non-coding RNA containing several oncogenic microRNAs (miR-1206, miR-1204) whose overexpression correlates with that *c-Myc* and is universally found in Myc-associated translocations [103]. PD-L2 is a member of the B7 family of proteins that, much like its more well-known relative PD-L1, interacts with the T-cell co-inhibitory receptor PD-1 (CD279) to induce T-cell anergy or apoptosis, consequently permitting tumors to evade immune detection. PD-L1/2 are co-amplified in B-lymphoid malignancy and are over-expressed in human CLL [95, 104, 105]. Novel therapeutics targeting this pathway by using PD-L1 and PD-1 checkpoint inhibitors are currently underway for the treatment of CLL [105, 106]. Expanded analysis of hypomethylated and over-expressed genes in Dnmt3a–/– CLL but not Dnmt3a+/– CLL revealed two additional promising candidate genes: *Tlr2* and *Tnfrsf13b* (also known as TACI). *TLR2* is a particularly promising target as it harbors recurrent activating mutations in human CLL [107]. Moreover, additional mutations

were found in *TLR5* and *IRAK1* in addition to the previously known GOF mutations that most frequently occur in *MYD88* resulting in NFκB activation. This implicates the TLR / MYD88 / NFκB axis in murine CLL pathogenesis. Coinciding with stimulation of the NFκB pathway, *Tnfrsfr13b* encodes for the receptor TACI – one of two known receptors for the ligands APRIL (A Proliferation-Inducing Ligand) and BAFF (B-Cell Activating Factor) – whose activation results in intracellular activation of NFAT and NFκB [108]. Importantly, in murine transgenic over-expression models of both ligands APRIL (Tg-APRIL) and BAFF (Tg-BAFF) mice developed CLL-like diseases and signaling was dependent upon the receptor TACI [109-111].

CHAPTER 3:

DNMT3A IS A HAPLOINSUFFICIENT TUMOR SUPPRESSOR IN MURINE CD8+ PERIPHERAL T-CELL LYMPHOMA

The material presented in this chapter was published previously [112]:

Haney SL*, Upchurch GM*, Opavska J, Klinkebiel D, Hlady RA, Roy S, Dutta S, Datta K, Opavsky R. Dnmt3a Is a Haploinsufficient Tumor Suppressor in CD8+ Peripheral T Cell Lymphoma. PLoS Genetics. 2016 Sep 30;12(9):e1006334.

(* denotes equal contribution)

† Note: In the subsequent chapter, "data not shown" refers to data that is not presented here but can be located in online supplementary data files of the above publication.

Introduction

DNA methylation is an epigenetic modification involved in transcriptional regulation of gene expression. Three catalytically active DNA methyltransferases - Dnmt1, Dnmt3a, and Dnmt3b - are involved in the generation and maintenance of DNA methylation in mammalian cells. Dnmt3a and Dnmt3b are classified as *de novo* enzymes due to their methylation activity during early embryogenesis [85], whereas Dnmt1 binds hemi-methylated DNA and functions in the maintenance of methylation marks during cellular division [86, 113].

Recent studies suggest that all Dnmts may play roles in generating and maintaining DNA methylation. For instance, in mouse hematopoietic stem cells, Dnmt3a is responsible for maintaining DNA methylation in lowly methylated regions known as canyons [51]. In addition, Dnmt1 was shown to have cancer-specific *de novo* activity in a mouse model of MYC-induced T cell lymphomas [96], whereas Dnmt3a and Dnmt3b were primarily involved in maintenance methylation in tumors [92, 93]. However, a deeper understanding of individual Dnmt's activities in normal development and in cancer is still missing.

DNA methyltransferase 3a has emerged as a central regulator of hematopoiesis over the last several years. The interest in Dnmt3a was in particular fueled by recent findings of somatic mutations in human hematologic malignancies of myeloid and T cell origin [22, 46]. Given the importance of DNA methylation for differentiation of hematopoietic lineages [98] along with critical roles of Dnmt3a in differentiation and self-renewal of hematopoietic stem cells [25, 26], it is not unexpected that a disruption of Dnmt3a activity affects a variety of cell types and has the potential to transform hematopoietic lineages. For example, recent studies using the *Mx1-Cre* transgene to conditionally delete Dnmt3a in hematopoietic stem and progenitor cells (HSPCs) followed by transplantation into lethally irradiated recipients showed that a vast majority of mice develop myeloid disorders such as myeloid dysplastic syndrome and acute myeloid leukemia (69%)

with rare occurrences of CD4+CD8+ double positive T-ALL (<12%) or B-ALL (<4%) [87]. In addition, both myeloid deficiencies and neoplasms were observed in mice transplanted with Dnmt3a-null bone marrow obtained from *Mx1-Cre;Dnmt3a*$^{fl/fl}$ mice, altogether highlighting the importance of Dnmt3a in prevention of myeloid transformation [88, 89]. However, the role of Dnmt3a in differentiation into hematopoietic lineages and molecular functions in normal and malignant hematopoiesis in particular remain poorly understood.

To elucidate the role of Dnmt3a in normal and malignant hematopoiesis we used the *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ (*Dnmt3a*$^{\Delta/\Delta}$) mouse model to conditionally delete Dnmt3a in all cells of the hematopoietic compartment. Using this model, we previously showed that long-term Dnmt3a-defficiency resulted in the development of a chronic lymphocytic leukemia (CLL) around 1 year of age [79, 95]. In addition, we previously reported that combined inactivation of Dnmt3a and Dnmt3b results in the development of CLL and peripheral T cell lymphoma (PTCL) [79], however the molecular basis of PTCL is poorly understood. Here we expanded on our previous studies by observation of a larger cohort of *Dnmt3a*$^{\Delta/\Delta}$ mice. These studies revealed that while ~60% of mice succumb to CLL, ~40% of mice develop CD8+ mature peripheral T cell lymphoma either in combination with CLL or as a singular disease. Furthermore, we found that loss of one allele of Dnmt3a is sufficient to induce CD8+ PTCL in 10% of Dnmt3a heterozygous mice with tumors retaining expression of the *wild-type* allele. Molecular profiling of methylation and gene expression identified promoter hypomethylation as a major event in tumorigenesis of PTCL, which was frequently accompanied by upregulation of gene expression. Furthermore, we identified downregulation of tumor suppressor p53 not only in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ lymphomas but also in pre-tumor thymocytes, suggesting that p53 downregulation is likely relevant in the initiation/progression of lymphomagenesis.

Altogether, our data demonstrate that Dnmt3a is a haploinsufficient tumor suppressor in the prevention of CD8+ T cell transformation and highlight the importance of understanding of the roles of Dnmt3a target genes in disease pathogenesis.

Material and Methods

Mouse Studies. *EμSRα-tTA* and *Dnmt3a*$^{2loxP/2loxP}$ (*Dnmt3a*$^{fl/fl}$) mice were acquired from D.W. Felsher (Stanford University) and R. Jaenisch (Whitehead University), respectively. *ROSA26*^{EGFP} and *Teto-Cre* mice came from the Jackson Laboratories. All transgenic mice were generated using standard crosses. All mice used in these studies were of the FVB/N background and were generated using standard genetic crosses. To obtain mice with a germline transmission of the *Dnmt3a*$^{-}$ allele, we crossed *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$ mice with FVB mice, taking advantage of our observation that the *EμSRα-tTA* transgene is expressed in germ cells. To generate *Dnmt3a*$^{+/-}$ we subsequently bred out transgenes by crossing obtained mice with FVB mice. PCR-based genotyping of genomic DNA isolated from the tails was used to confirm genotypes. *Dnmt3a*$^{+/-}$ mice were harvested at the experimental end point of 16 months.

Human Lymphoma Samples. Human peripheral T cell lymphoma tissue samples were acquired from Cooperative Human Tissue Network, a National Cancer Institute supported resource (www.chtn.org).

FACS analysis. All analysis was performed at the Flow Cytometry Facility at the University of Nebraska Medical Center. Single cell suspensions were generated from mouse organs and labeled with fluorescently conjugated antibodies (eBioscience). Data was collected using the LSR II (BD Biosciences) and analyzed using BD FACSDiva software (BD Biosciences). Immunopheotypic criteria for normal and malignant cellular populations analyzed by flow cytometry are as follows: Cytotoxic T cells (EGFP-negative CD3+CD8+ without population expansion), CD8+ peripheral T cell lymphomas (EGFP+CD3+CD8+ with population expansion), B-1a cells (EGFP-negative CD19+B220+CD5+ without population expansion), chronic lymphocytic leukemia (EGFP+CD19+B220+CD5+ with population expansion). Clonality was assessed using the mouse Vβ TCR Screening Panel (BD bioscience) which uses FITC-conjugated

monoclonal antibodies to recognize mouse Vβ 2, 3, 4, 5.1 and 5.2, 6, 7, 8.1 and 8.2, 8.3, 9, 10b, 11, 12, 13, 14, and 17a T cell receptors.

Western Blot. Western blots were performed as previously described [6] with use of the following antibodies: Dnmt3a (H-295, Santa Cruz), γ-Tubulin (H-183, Santa Cruz), p53 (SC-6243, Santa Cruz), HDAC1 (ab7028, Abcam), and HSC-70 (SC-7298, Santa Cruz).

Combined Bisulfite Restriction Analysis (COBRA) and Bisulfite sequencing. COBRA analysis was carried out as described previously [93, 96]. Mouse and human bisulfite specific primers are shown in Table S11.

Whole Genome Bisulfite Sequencing (WGBS). DNA was isolated from FACS sorted CD8+ T cells obtained from the spleens of control FVB/N (CD8+CD3+) and *Dnmt3a*$^{\Delta/\Delta}$ terminally ill PTCL mice (EGFP+CD8+CD3+). The WGBS libraries were prepared and sequenced in the DNA Services facility at the University of Illinois at Urbana-Champaign, Roy J. Carver Biotechnology Center / W.M. Keck Center using two lanes for each sample on the Illumina HiSeq2500 sequencer with paired-end 160bp reads. Each lane produced over 310 million reads. Sequence tags were aligned with the mouse genome (Dec. 2011 mus musculus assembly mm10, Build 38) using the methylated sequence aligner Bismark [114] by the University of Nebraska Epigenomics Core facility. The resulting data file contains the percent methylation at each CpG measured. Each individual CpG was retained and percent methylation determined only if it was represented by ≥5 individual sequences. Correlation based, average linkage hierarchical clustering of genome location matching CpG methylation percentages per sample was performed using the R software package RnBeads [115]. Genome location matching differentially methylated cytosines (DMCs) and differentially methylated regions (DMRs) were determined using the R software package DSS [116]. DMCs were determined by first smoothing the raw percent methylation values based on a moving average algorithm and smoothing span of 500 bases. DMRs were then

determined based on average DMC methylation change of 30 percent or greater, at least 50 percent or greater individual DMC, p-values less than 0.05, minimum base pair length of 100, minimum of three DMLs represented, and the resulting DMRs were averaged if they were closer than 50 bases. Circos plots [117] were generated to visualize DMRs that had at least a 100 base overlap with genomic promoters defined as 1500 bases upstream of the transcription start site (TSS) to 500 bases downstream of the TSS. DMRs were aligned with the mouse genomic repeats. Genomic repeats were acquired from the UCSC Genome table browser based on the RepeatMasker program (http://www.repeatmasker.org). The repeat was retained if the overlap between the DMR and repeat was more than 25 percent of the length of the repeat. For analysis of enhancer methylation, DMRs were aligned using the bedtools intersect routine with the 48,415 2000bp enhancer annotated regions that were determined based on histone marks (high H3K4me1/2 and low H3K4me3) [118, 119]. WGBS data is available for download through the NCBI Gene Expression Omnibus (GSE78146).

Transcription factor enrichment. The mouse mm9 MotifMap database containing 2,237,515 transcription factor motifs [120, 121] (http://motifmap.ics.uci.edu/) was used to align transcription factors motifs present within promoters (-1500 to +500 TSS) that contained a significant hypo- or hypermethylated DMR using the Bedtools intersect routine [119]. For a control comparison 12 random sets of promoters (500 promoters each were used for the hypometylated controls and 50 for the hypermethylated) were selected from the UCSC known mm9 genes database using the Excel RANDBETWEEN function and sorting from highest to lowest number. The abundance of each transcription factor within the DMR promoters and random promoters were counted using the Excel COUNTIF function. P-values were calculated used a Wilcoxon sign rank test. Only $P<0.05$ were considered significant.

Reduced Representation Bisulfite Sequencing. Splenic CD3+CD8+ T cells were isolated by FACS sorting from two *Dnmt3a*$^{\Delta/\Delta}$ mice with PTCL. Age-matched control T cells were FACS-sorted from spleens of FVB/N mice (n=2). Genomic DNA was isolated using standard protocols. The RRBS libraries were prepared and sequenced at the Medical Genome Facility at the Mayo Clinic and ran on an Illumina HiSeq2500 sequencer. The Streamlined Analysis and Annotation Pipeline for RRBS data (SAAP-RRBS) was specifically designed to analyze RRBS data [122]. This software was used to align and determine the methylation status of CpGs associated with this type of restriction digest high throughput method. Sequences were initially aligned with genome mm9 then converted to mm10 using the UCSC Genome Browser Batch Coordinate Conversion (liftOver) utility. The methylation heat map was generated by taking the averages for all differentially methylated CpGs for a promoter (-1500 to +500 base pairs relative to the transcription start site). Promoters were only considered to be differentially methylated if one or more CpG sites showed a 30% change in methylation. RRBS data is available for download through the NCBI Gene Expression Omnibus (GSE78146).

RNA-seq. RNA was isolated as previously described [6] from FACS sorted CD8+ T cells obtained from spleens of control FVB/N (CD8+CD3+) and *Dnmt3a*$^{\Delta/\Delta}$ terminally ill PTCL mice (EGFP+CD8+CD3+). Library generation was performed using the TruSeq mRNA kit. The resulting libraries were sequenced on the Illumina HiSeq 2000 platform using paired-end 100bp runs (SeqMatic, Fremont, CA). The resulting sequencing data was first aligned using TopHat and mapped to the Mus musculus UCSC mm10 reference genome using the Bowtie2 aligner [123]. Cufflinks 2 was used to estimate FPKM of known transcripts, perform de novo assembly of novel transcripts, and calculate differential expression [124]. For differentially expressed genes, we considered those genes with a fold change ≥ 2 and a p-value < 0.05 to be significant. RNA-seq data is available for download through the NCBI Gene Expression Omnibus (GSE78146).

Quantitative Real-Time qRT-PCR. qRT-PCR was performed as previously described [93]. Mouse real time primer sequences used in experiments presented are shown in Table S11.

Statistical Methods. Data was compared in Microsoft Excel using Student's t-test ($p<0.05$ considered significant) or other appropriate statistical comparison listed elsewhere in materials and methods.

Histology. H&E staining was performed using standard protocols by the University of Nebraska Medical Center Tissue Science Facility.

Microarray. Microarray data was downloaded from the NCBI Gene Expression Omnibus. We compared gene expression of 5 normal Tonsil T cells samples (GSE65135) to 3 PTCL samples in which DNMT3A was reported to be mutated (GSE58445) [42]. Datasets were generated with Affymetrix U133 plus 2 arrays and analyzed using Affymetrix Expression Console and Transcriptome Analysis Console (v3.0). Data was analyzed using a one-way between-subject ANOVA to generate p-values and identify differentially expressed genes (p-value < 0.05 and fold change >1.5). Genes differentially expressed in human PTCL were compared to those genes identified as being over- or underexpressed in mouse *Dnmt3a*$^{+/-}$ or *Dnmt3a*$^{\Delta/\Delta}$ relative to CD8+ T cell controls (RNAseq, Fold change >2, $p<0.05$).

Ingenuity Pathway Analysis. All differentially expressed genes ($p<0.05$, fold change >2, analyzed by Cufflinks V2.0) for *Dnmt3a*$^{+/-}$ PTCL relative to *wild-type* CD8+ T cells and *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to *wild-type* CD8+ T cells were imported into IPA software. Core analysis were performed to identify top ranking pathways and categories for differentially expressed genes. Activated and inhibited pathways (Z-score>1.5, $p<0.05$) common to both *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL are shown in Fig S4. In Fig 3E, IPA core analysis was performed on highly expressed genes

(FPKM ≥ 10) in *wild-type* CD8+ T cells and the top subcategories obtained in Physiological System, Development and Functions were displayed (P<0.05, for all subcategories).

Gene Set Enrichment Analysis (GSEA). TopHat/Cufflinks/Cuffdiff RNA-seq gene-level read_group_tracking file was converted to GCT expression dataset and matching phenotype model using the Read_group_trackingToGct module (http://www.broadinstitute.org/cancer/software/genepattern/modules/docs/Read_group_trackingToGct/1). Gene Set Enrichment Analysis (GSEA, http://www.broadinstitute.org/gsea/index.jsp) was used to test the relationship between RNA-Seq mRNA expression and the Hallmark Signature gene sets (http://software.broadinstitute.org/gsea/msigdb/genesets.jsp?collection=H). From this we concentrated our effort on the Hallmark p53 pathway gene set (http://software.broadinstitute.org/gsea/msigdb/cards/HALLMARK_P53_PATHWAY.html) that consisting of 180+ genes involved in p53 pathway and network.

Retroviral shRNA Knockdown Experiments. HuSH 29-mer shRNA scrambled and shRNA Jdp2 in the retroviral vector pRFP-C-RS were purchased from Origene. Infections were performed as previously described [93], using a Dnmt3a-deficient MYC-induced CD4+ T cell lymphoma line [92]. Doubling time was calculated from each measured time point relative to the starting concentration of cells at Day 0. Each time point calculation of doubling time was considered a replicate measure and was averaged other measurements per experimental condition.

P53 Overexpression in Murine T Cell Lymphoma Cell Lines. The pMSCV-IRES-EGFP ("Vector") and pMSCV-bla-p53(WT)-IRES-EGFP ("p53"), a kind gift from Dr. Ute Moll [125], were transfected into the Phoenix$_{Eco}$ packaging cells and retrovirus was produced. Transductions were performed as previously described [93], using *wild-type* Dnmt3a and Dnmt3a-deficient MYC-induced CD4+ T cell lymphoma lines [92]. The maximum percent of EGFP expressing cells per cell population was observed at 48 hours post-transduction and all subsequent

EGFP data points were normalized to this time point. EGFP was measured periodically by flow cytometry on the LSRII available at the UNMC flow cytometry core facility. Cells were cultured in RPMI-1640 media supplemented with 10% FBS, 1% pen-strep-amphotericin B, 0.5% β-mercaptoethanol and split (1:3) to (1:5) every 3 days.

Study approval. This study was performed in accordance with the guidelines established by the Guide for the Care and Use of Laboratory Animals at the National Institutes of Health. All experiments involving mice were approved by the IACUC (Protocol number: 08-083-10-FC) at the University of Nebraska Medical Center.

Results

Dnmt3a$^{\Delta/\Delta}$ mice develop B220+CD19+CD5+ CLL and CD8+ PTCL.

We previously utilized quadruple transgenic mice, *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*; Rosa26LOXP*$^{EGFP/EGFP}$ (designated hereafter as *Dnmt3a*$^{\Delta/\Delta}$), to conditionally inactivate Dnmt3a in HSPCs as well as mature cells of all hematopoietic lineages (**Fig 1A**). In such genetic setting, Dnmt3a is deleted in only ~40% of hematopoietic cells due to restricted patterns of *EμSRα-tTA* expression and cells are marked by EGFP expression driven from a reporter gene [79, 95]. In the remaining 60% of hematopoietic cells the *EμSRα-tTA* transgene is not expressed and, as a result, the Dnmt3a conditional allele is not deleted and the EGFP reporter is not expressed. PCR based genotyping confirmed Dnmt3a deletion in EGFP–positive but not EGFP-negative stem cells as well as T-, B-, and myeloid cells (**Fig 1B**). Furthermore, Dnmt3a deletion was specific to cells of the hematopoietic lineages and was not observed in solid tissues (**Fig 2A**). Previously, we observed a small cohort of *Dnmt3a*$^{\Delta/\Delta}$ mice and reported the development of a CLL-like disease with a median survival of 371 days [79]. Here we utilized a larger cohort of 42 *Dnmt3a*$^{\Delta/\Delta}$ mice to observe the phenotypic consequences of Dnmt3a inactivation. Consistent with our previously reported data, we observed the development of a CLL-like disease in 61% of mice with a median survival of 306 days (**Fig 1C and 2B**). This disease was characterized by expansion of B220+CD19+CD5+IgM+ EGFP+ cells in the spleen, blood, bone marrow, peritoneal cavity (IP) and occasionally in the lymph nodes (**Fig 2C and 2D**). Interestingly, 12% of *Dnmt3a*$^{\Delta/\Delta}$ mice developed a different disease (MS= 295 days) characterized not only by splenomegaly but also by a significant enlargement of lymph nodes that was not observed in the CLL cases (**Fig 1C, 1D and 2B**). Histological analysis of spleens showed near complete effacement of the red pulp by massively expanded white pulp (**Fig 1E**). Small- to medium-sized cells were EGFP+, expressed markers of mature T cells – CD3, CD5, TCRβ and CD8 – and were negative for the expression of CD4, TCRγδ, NK-1.1 and CD16

(**Fig 1F, 1G and 2E**). To determine if *Dnmt3a*$^{\Delta/\Delta}$ PCTLs are clonal, we analyzed TCR-Vβ rearrangement by FACS analysis. All three *Dnmt3a*$^{\Delta/\Delta}$ lymphomas analyzed showed only one or two TCR-Vβ rearrangements, suggesting that they were clonally derived from a single cell following somatic recombination of the TCR-β locus (**Fig 2F and Table 2**). These phenotypes are most consistent with those observed in human cytotoxic peripheral T cell lymphomas not otherwise specified (PTCL-NOS).

Dnmt3a was efficiently deleted in lymph node cells of terminally ill mice as determined by immunoblot analysis using anti-Dnmt3a antibody (**Fig 1H**). When lymph node cells from terminally sick *Dnmt3a*$^{\Delta/\Delta}$ mice were injected into the IP cavity of sublethally irradiated FVB mice, recipients developed a CD3+CD8+ PTCL similar to that observed in the donor mice (**Fig 1I**), suggesting that *Dnmt3a*$^{\Delta/\Delta}$ cells have tumorigenic potential. Tumor burden (scored as average weights of spleens in terminally ill mice) was on average higher in PTCL than in CLL (**Fig 2G**).

In addition to the development of distinct disease types in individual mice, in 29% of mice we observed the simultaneous development of CLL and PTCL with a median survival of 293 days (**Fig 1C, 2B and 2H**). Interestingly, all CLL and PTCL cases presented with B220+CD19+CD5+ and CD3+CD8+ immunophenotypes, respectively, suggesting that normal B-1a B cells and cytotoxic CD8+ T cells are in particularly sensitive to cellular transformation in the absence of Dnmt3a. Altogether, these data suggest that Dnmt3a is a tumor suppressor gene in prevention of CLL and PTCL in mice.

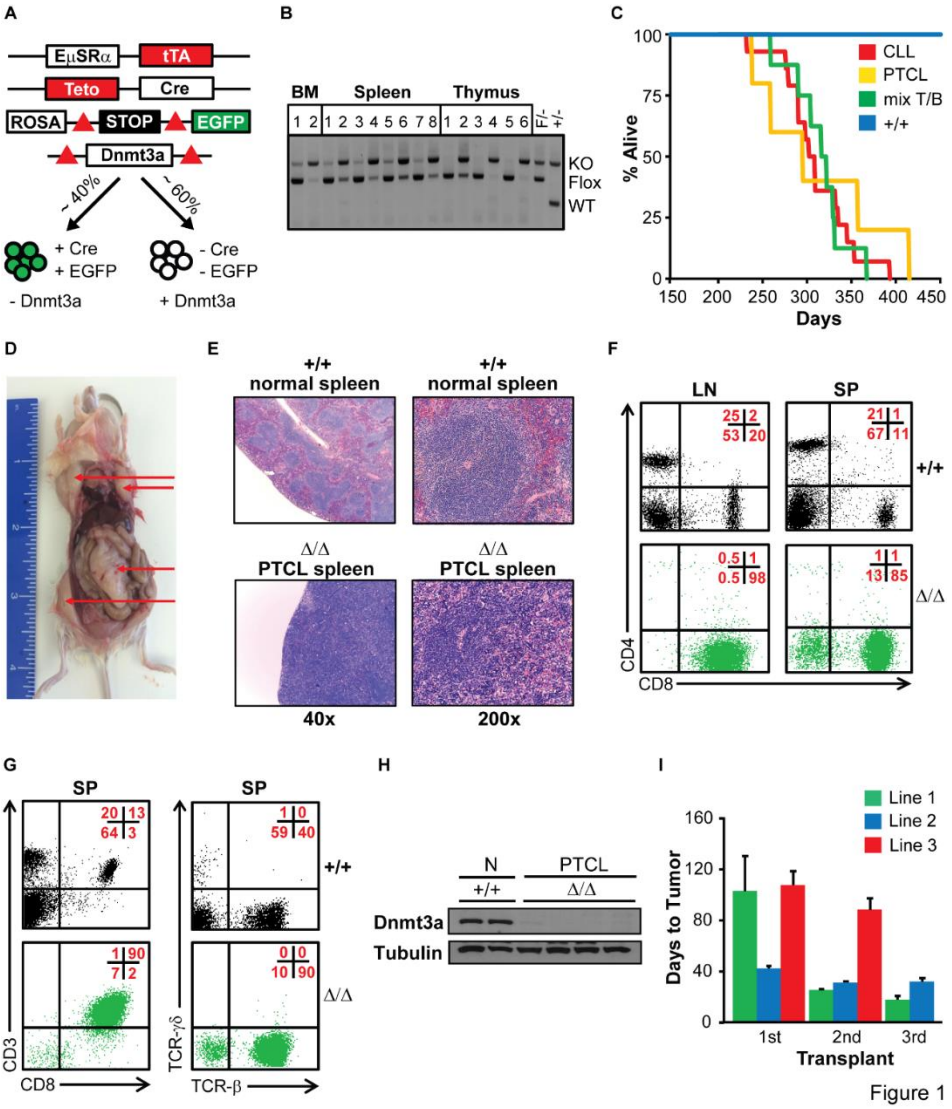


Figure 1

Figure 1. *Dnmt3a*$^{\Delta/\Delta}$ mice develop B220+CD19+CD5+ CLL and CD8+ PTCL. (A) Genetic setting used to conditionally delete Dnmt3a. The tetracycline activator protein (tTA) is expressed in ~40% of all hematopoietic cells (including stem cells). tTA promotes expression of the *Teto-Cre* transgene. Expression of Cre results in the excision of the stop cassette located upstream of the *Rosa26LOXP*^{EGFP} reporter locus and deletion of Dnmt3a within the same subpopulation of cells. Thus, inclusion of the EGFP transgene allows for monitoring cells expressing tTA, and Cre, as well as to identify cells deleted for Dnmt3a. (B) PCR-based genotyping of the Dnmt3a locus using DNA isolated from FACS-sorted EGFP-positive (+) and EGFP-negative (−) populations on cells obtained from the bone marrow (BM), spleen (SP) and thymus (TH) of 6-week-old *Dnmt3a*$^{-/-}$ mice. PCR reactions were performed on samples in the following order, BM: (1) -LSK, (2) +LSK, SP: (1) -B2, (2) +B2, (3) -B1, (4) +B1, (5) -myeloid, (6) +myeloid, (7) -erythroid, (8) +erythroid, TH: (1) -DP, (2) +DP, (3) -CD4, (4) +CD4, (5) -CD8, (6) +CD8. Immunophenotypes of sorted populations are as follows: LSK (Lineage-, Sca-1+, ckit+), B2 (B220+, CD5-), B1 (B220+, CD5+), Myeloid (CD11b+), erythroid (Ter119+), DP (CD4+, CD8+), CD4 (CD4+, CD8-), CD8 (CD4-, CD8+). Fragments from floxed (F), *wild-type* (WT) and knockout (KO) alleles are shown. DNA from conventional Dnmt3a heterozygous (F/- and +/-) mice were used as controls. (C) Kaplan Meyer survival curve for *Dnmt3a*$^{\Delta/\Delta}$ mice. Moribund mice were classified at the time of terminal harvest with CLL (red), PTCL (yellow), or mixed CLL/PTCL (green) based on the presence of B-1a or CD8+ T cells tumors, as determined by FACS. *Dnmt3a*$^{+/+}$ *wild-type* (+/+) controls are shown in blue. (D) A terminally ill *Dnmt3a*$^{\Delta/\Delta}$ mouse with PTCL. Generalized lymphadenopathy is denoted by arrows. (E) H&E stained sections of *Dnmt3a*$^{+/+}$ (normal) and *Dnmt3a*$^{\Delta/\Delta}$ (PTCL) spleens (40X and 200X). (F) CD8 and CD4 expression in cells isolated from the lymph nodes (LN) and spleens (SP) of *Dnmt3a*$^{+/+}$ control (+/+) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (Δ/Δ) mice as determined by FACS. Percentages of cells in each quadrant are shown in the top right in red. (G) CD3, CD8, TCR-β and TCR-γδ expression in cell isolated from the spleens (SP) of *Dnmt3a*$^{+/+}$ control (+/+) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (Δ/Δ) mice, as determined by FACS. (H) Immunoblot showing Dnmt3a protein levels in *Dnmt3a*$^{+/+}$ normal (N) controls (+/+) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (Δ/Δ) lymph node samples. γ-tubulin is shown as a loading control. (I) Time to tumor for primary (1), secondary (2) and tertiary (3) sub-lethally irradiated FVB-recipient mice serially transplanted with CD8+ PTCL tumors isolated from the lymph nodes of *Dnmt3a* $^{\Delta/\Delta}$ terminally sick mice. Data are presented as average time to tumor development. Three PTCL lines are shown.

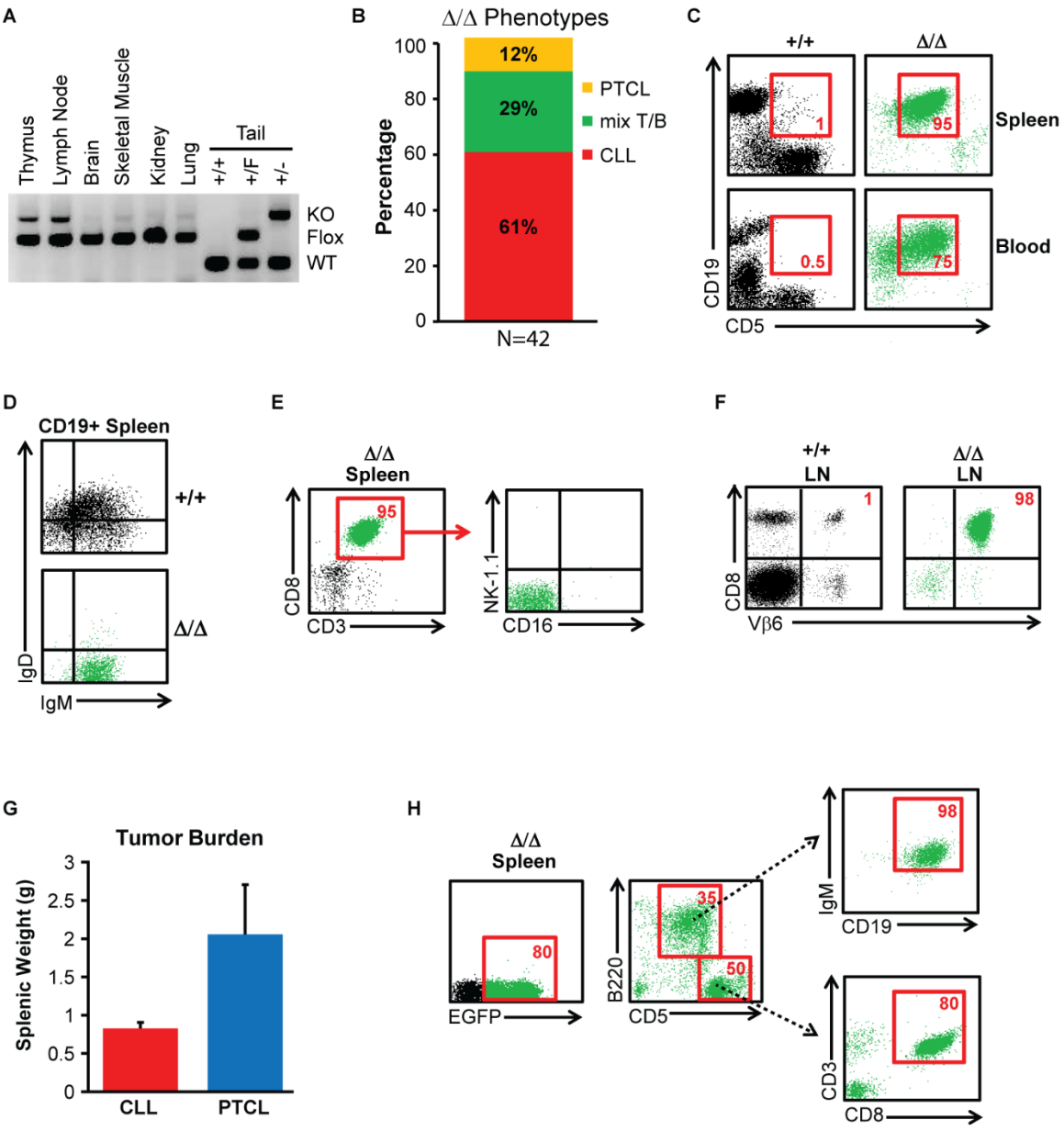


Figure 2. Loss of Dnmt3a induces CLL and PTCL in mice. (A) PCR based genotyping of the Dnmt3a locus using gDNA isolated from the thymus, lymph node, brain, skeletal muscle, kidney, and lung of *Dnmt3a*$^{\Delta/\Delta}$ mice. gDNA isolated from tails of *Dnmt3a*$^{+/+}$, *Dnmt3a*$^{+/F}$, and *Dnmt3a*$^{+/-}$ mice were used as controls. Knockout (KO), floxed, and wild-type (WT) bands are labeled. (B) Breakdown of phenotype spectrum in 42 *Dnmt3a*$^{\Delta/\Delta}$ mice. (C) CD19 and CD5 expression in cells isolated from the spleen and blood of *Dnmt3a*$^{+/+}$ control (+/+) and *Dnmt3a*$^{\Delta/\Delta}$ CLL (Δ/Δ) mice, as determined by FACS. Percentage B-1a cells are shown in the red box. (D) IgD and IgM expression in cells isolated from the spleen of *Dnmt3a*$^{+/+}$ control (+/+) and *Dnmt3a*$^{\Delta/\Delta}$ CLL (Δ/Δ) mice, as determined by FACS. (E) FACS analysis showing CD8+CD3+ *Dnmt3a*$^{\Delta/\Delta}$ PTCL cells do not express Nk-1 and CD16 markers. (F) Representative FACS diagram showing clonal TCR-Vβ expression in a *Dnmt3a*$^{\Delta/\Delta}$ tumor. Control lymph node (+/+) is shown for reference. (G) Tumor burden as determined by average weight of spleens in *Dnmt3a*$^{\Delta/\Delta}$ CLL (red) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (blue). (H) FACS diagram showing the simultaneous expansion of B220+CD5+CD19+IgM+ B-1a cells and CD8+CD3+ T cells in the spleen of a *Dnmt3a*$^{\Delta/\Delta}$ mouse.

Table 2. Summary of TCR Vβ expression in PTCL samples.

		TCRβ Variable Chain in CD8+ Tumor Cells															
Genotype	Tissue	Vβ 8.1, 8.2	Vβ 5.1, 5.2	Vβ 6	Vβ 11	Vβ 9	Vβ 9	Vβ 13	Vβ 3	Vβ 17a	Vβ 7	Vβ 14	Vβ 2	Vβ 10b	Vβ 12	Vβ 4	Vβ 8.3
Dnmt3a(-/-)	LN	(–)	(–)	(+)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)
Dnmt3a(-/-)	LN	(–)	(–)	(+)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)
Dnmt3a(-/-)	LN	(–)	(–)	(+)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)
Dnmt3a(+/-)	LN	(–)	(–)	(–)	(+)	(–)	(–)	(–)	(–)	(–)	(+)	(–)	(–)	(–)	(–)	(–)	(–)
Dnmt3a(+/-)	LN	(+)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)
Dnmt3a(+/-)	LN	(+)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)	(–)

Table 2. Summary of TCR-Vβ expression in PTCL samples. *Dnmt3a*$^{+/-}$ (n = 3) and *Dnmt3a*$^{\Delta/\Delta}$ (n = 3) PTCL lymph node (LN) samples were analyzed by flow cytometry for the expression of 15 different TCR-Vβ surface markers. (-) indicates negative expression, whereas (+) denotes positive expression of TCR-Vβ markers.

Dnmt3a is a haploinsufficient tumor suppressor in the prevention of CD8+ PTCL in mice.

We have recently reported that mice harboring a conventional knockout allele of Dnmt3a (*Dnmt3a*$^{+/-}$ mice) develop either CLL, myeloproliferative disorder or remain healthy by 16 months of age [95]. Here we expanded these studies by observing a larger cohort of 30 *Dnmt3a*$^{+/-}$ and 20 control *Dnmt3a*$^{+/+}$ mice. Interestingly, we found that 3 out of 30 analyzed mice developed CD8+ T cell lymphomas, which were indistinguishable from those observed in *Dnmt3a*$^{\Delta/\Delta}$ mice (**Fig 3A-C**). None of the control mice were affected by lymphoma and remained healthy during the observational period. Serial transplantation of *Dnmt3a*$^{+/-}$ lymphoma cells induced PTCL within 2 months in secondary and tertiary transplanted mice, illustrating their selective advantage to grow and induce disease (**Fig 3D**). *Dnmt3a*$^{+/-}$ lymphomas retained approximately 50% expression of Dnmt3a, suggesting that the remaining allele is expressed in fully transformed cells (**Fig 3E).** Like *Dnmt3a*$^{\Delta/\Delta}$ lymphomas, *Dnmt3a*$^{+/-}$ lymphomas were also clonal (**Fig 3F and Table 2**). Importantly, sequencing analysis of cDNA generated from two independent *Dnmt3a*$^{+/-}$ PTCL samples revealed no mutations in the coding sequence of Dnmt3a (data not shown), demonstrating that the expressed Dnmt3a allele is in the *wild-type* configuration. Similarly, we did not find any mutations in the coding sequences of two genes that are commonly mutated in human T cell malignancies, Tet2 and RhoA, and their expression was not changed in Dnmt3a-deffcient lymphomas, suggesting that changes in the activity of these genes may not be involved in the transformation of T cells in this model (**Fig 4** and data not shown). Altogether, these data suggest that Dnmt3a is a haploinsufficient tumor suppressor gene in the prevention of CD8+ T cell lymphomas and CLL in mice.

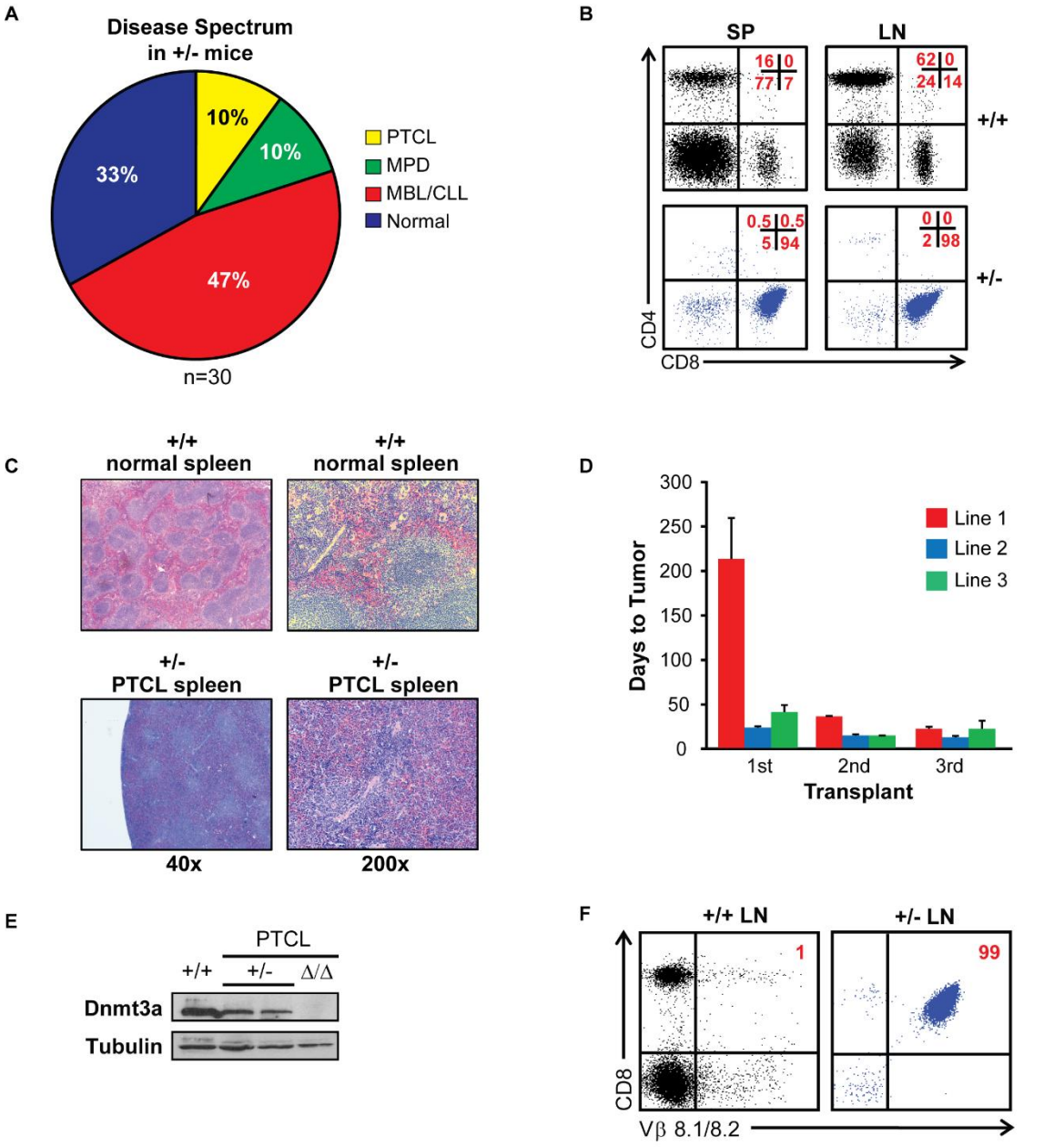


Figure 3. Dnmt3a is a haploinsufficient tumor suppressor in the prevention of CD8+ PTCL in mice. (A) Percentage of conventional heterozygous *Dnmt3a*$^{+/-}$ mice developing PTCL (yellow), MPD (green; MPN), CLL (red), or no disease (blue) at time of harvest, as determined FACS. (N=30). (B) FACS analysis of CD8 and CD4 expression in cells isolated from the spleens (SP) and lymph nodes (LN) of *Dnmt3a*$^{+/+}$ control (+/+) and *Dnmt3a*$^{+/-}$ PTCL (+/-) mice. Percentages of cells in each quadrant are shown in the top right in red. (C) H&E stained sections of *Dnmt3a*$^{+/+}$ (normal) and *Dnmt3a*$^{+/-}$ (PTCL) spleens (40X and 200X). (D) The time to tumor development for primary (1), secondary (2) and tertiary (3) sub-lethally irradiated FVB-recipient mice serially transplanted with CD8+ PTCL tumors isolated from the lymph nodes of a *Dnmt3a*$^{+/-}$ terminally sick mice. Data are presented as average time to tumor development. Three PTCL lines are shown. (E) Immunoblot showing Dnmt3a protein levels in *Dnmt3a*$^{+/+}$ controls (+/+) and *Dnmt3a*$^{+/-}$ PTCL (+/-) samples. *Dnmt3a*$^{\Delta/\Delta}$ PTCL (Δ/Δ) is shown as a negative control. γ-tubulin is shown as a loading control. (F) Representative flow diagram showing CD8 and TCR-Vβ 8.1/8.2 expression in *Dnmt3a*$^{+/+}$ controls (+/+) and *Dnmt3a*$^{+/-}$ PTCL LN samples.

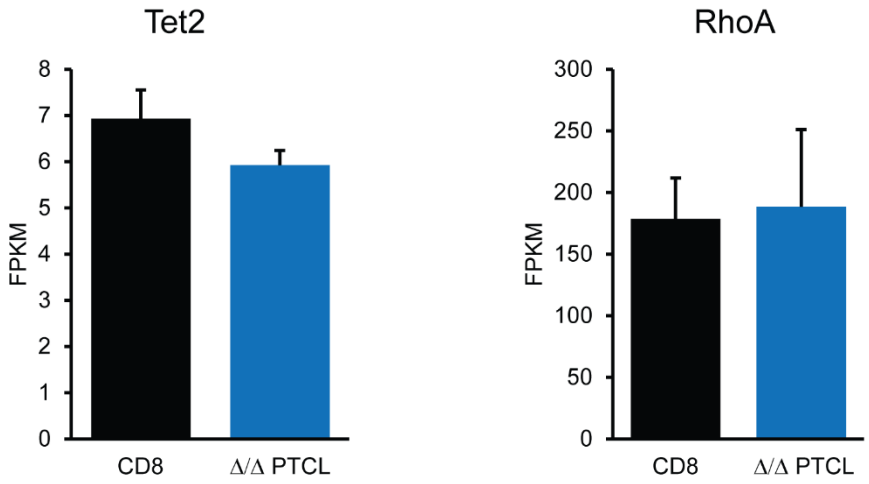


Figure 4. Levels of Tet2 and RhoA are unchanged in *Dnmt3a*$^{\Delta/\Delta}$ PTCL. Expression data from RNAseq (FPKM) for Tet2 and RhoA transcript levels in *Dnmt3a*$^{+/+}$ CD8+ T cells and *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples.

A majority of promoters are methylated and inactive in normal mouse CD8+ T cells.

To determine the nature of deregulated molecular events during PTCL development in *Dnmt3a*$^{\Delta/\Delta}$ mice, we performed global methylation analysis using whole genome bisulfite sequencing (WGBS) and gene expression profiling by RNA-seq on CD8+ T cells isolated from *Dnmt3a*$^{+/+}$ spleens, as this cellular population is immunophenotypically the closest normal counterpart of CD8+CD4- PTCLs. Methylation analysis revealed that 75% of 13,859,068 CpG dinucleotides were heavily methylated (≥76%), while only 6% were methylated at low levels (≤25%) (**Fig 5A**). The remaining 19% of CpG were methylated at intermediate levels (25% to 75%). Likewise, 44% of core promoters (-300 to +150 bp relative to transcription start site; TSS) were heavily methylated (≥76%), while 29% were lowly methylated (≤25%) (**Fig 5B**). Analysis of methylation across core promoter regions revealed that over 13,000 genes had a mean methylation value greater than 50%, suggesting that the majority of promoters in CD8+ T cells are heavily methylated (**Fig 5C** and data not shown). A combined gene expression and methylation analysis revealed that the majority of genes with low levels of promoter methylation were expressed, whereas genes with high levels of promoter methylation were largely repressed, suggesting that promoter methylation correlates with gene expression (**Fig 5D** and data not shown).

Ingenuity pathway analysis (IPA) of highly expressed genes in CD8+ T cells revealed the top subcategories of genes significantly associated with *organismal survival*, *hematological system*, *tissue morphology*, *hematopoiesis*, *lymphoid tissue structure* (**Fig 5E**), underlining their link to hematopoietic system. Altogether, these data reveal that a significant number of promoters are hypermethylated and inactive in normal CD8+ T cells, highlighting both the importance of DNA methylation in differentiation and potential for deregulation of these genes upon inactivation of DNA methyltransferases.

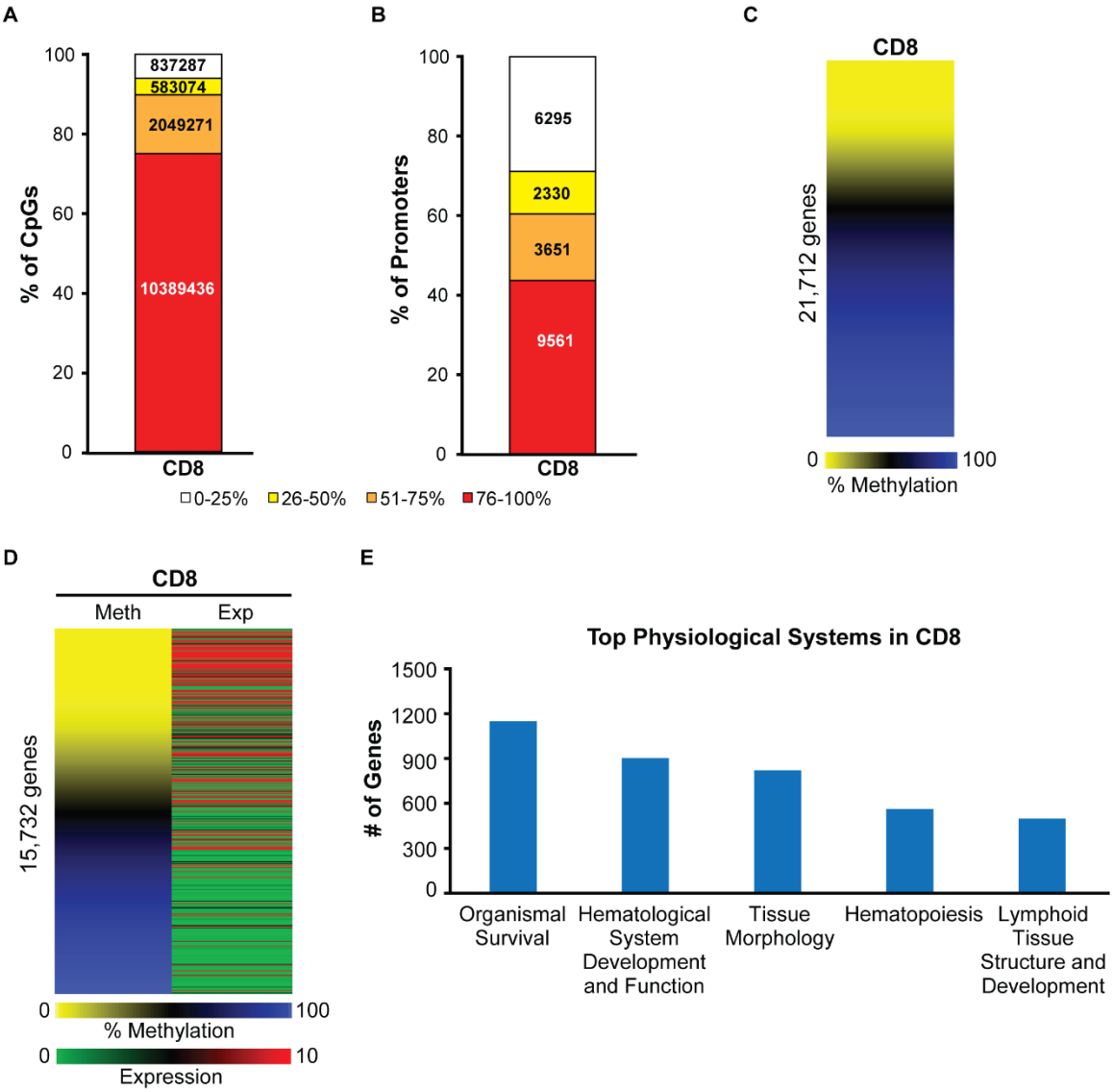


Figure 5. A majority of promoters are methylated and inactive in normal mouse CD8+ T cells. (A) CpG methylation in *wild-type* CD8+ cells, as determined by WGBS. Individual CpGs were placed into quartiles based on percent methylation (0-25%, 26-50%, 51-75%, and 76-100%). (B) Percent methylation shown by quartiles for core promoter regions (-300bp to +150bp relative to the TSS) in *wild-type* CD8+ cells. Methylation percentages for all CpGs across the 450bp region were averaged to give a mean methylation value for each gene's core promoter. (C) A heat map displaying methylation status of 21,712 promoters in *wild-type* CD8+ as determined by WGBS. Methylation percentage for individual CpGs were annotated to the promoter regions −300bp to +150bp relative to the transcription start site (TSS). Methylation percentages for all CpGs across the 450bp region were averaged to give a mean methylation value for each gene promoter. Lowly methylated promoters are shown in yellow and highly methylated promoters in blue (D) Heat map presentation of gene-matched promoter methylation (as analyzed in panel C) and corresponding transcriptional expression (averaged FPKM values) in *wild-type* CD8+ cells, as determined by WGBS and RNA-seq for 15,732 genes. Highly expressed genes are denoted in red and lowly expressed genes are denoted in green. (E) Ingenuity Pathway analysis (IPA) of highly expressed genes (FPKM ≥ 10) in *wild-type* CD8+ cells. The top subcategories obtained in Physiological System, Development and Functions are displayed (P<0.05, for all subcategories).

***Dnmt3a*$^{\Delta/\Delta}$ PTCL is characterized by genome wide hypomethylation.**

To determine the effects of Dnmt3a loss on the cancer methylome, we next performed WGBS on DNA isolated from *Dnmt3a*$^{\Delta/\Delta}$ PTCL cells. Out of ~14 million CpG dinucleotides analyzed, we observed decreased methylation in 1,263,413 (9%) CpGs and increased methylation in 155,977 (1%) CpGs (**Fig 6A** and data not shown**)**. By analysis of differentially methylated cytosines (DMCs) we found that the majority of DMCs were present in gene bodies and intergenic regions, where hypomethylation was 8 fold higher than hypermethylation (**Fig 6B**). Although the vast majority of changes in methylcytosine levels occurred in gene bodies and intergenic regions, we detected an overall decrease in methylation in both long and short promoter regions (-1500 to +500 bp and -300 to +150 bp relative to TSS, respectively) in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8+ T cell controls (**Fig 6C** and data not shown). Likewise, analysis of differentially methylated regions (DMRS) found significant changes in the methylation of long promoters, with 500 hypomethylated DMRS and 50 hypermethylated DMRS identified in PTCL relative to CD8+ T cell controls (**Fig 6D** and data not shown). Similarly, short promoters were hypomethylated (132) more so than hypermethylated (19) in PTCL (**Fig 6D** and data not shown). Like with promoters, hypomethylated DMRS were 10-and 18-fold higher than hypermethylated DMRS in gene bodies and enhancers, respectively (**Fig 6D** and data not shown). Extensive hypomethylation was also observed in repeat elements, with LINE elements showing the largest degree of hypomethylation (~4 fold) compared to hypermethyation (**Fig 6E**).

Next, we were interested in determining if differentially methylated promoter regions shared particular transcription factor binding motifs. This analysis revealed a significant enrichment for AML1, NF-κB, and OCT1 binding motifs in hypomethylated promoters (**Fig 6F**). This suggests a possible involvement of these factors in maintenance methylation performed by Dnmt3a. Similarly, AP-2rep, SOX5, and myogenin binding motifs were enriched in gene promoters

hypermethylated in tumors, possibly implying role of these factors in cancer-specific aberrant methylation (**Fig 6G**). Further functional analysis will be required to test an involvement of these proteins in deregulated methylation in mouse lymphomas.

Promoter and gene-body hypomethylation is present throughout the genome of *Dnmt3a*$^{\Delta/\Delta}$ PTCLs.

Locus-specific analysis revealed that hypo- and hypermethylated DMRS associated with promoters and gene bodies were relatively equally distributed across the genome, with the highest number of hypomethylated promoters present on chromosomes 11 and 5, lowest numbers on chromosomes X and 12 (**Fig 7** and data not shown). Interestingly, very few differentially methylated promoters were detected on the X chromosome, suggesting that Dnmt3a is dispensable for maintenance methylation in these areas of the genome (**Fig 7** and data not shown). Altogether, our data suggest that disease development in the absence of Dnmt3a results in decreased methylation across the genome with a significant number of gene promoters affected whose untimely activation may contribute to the malignant transformation of CD8+ T cells.

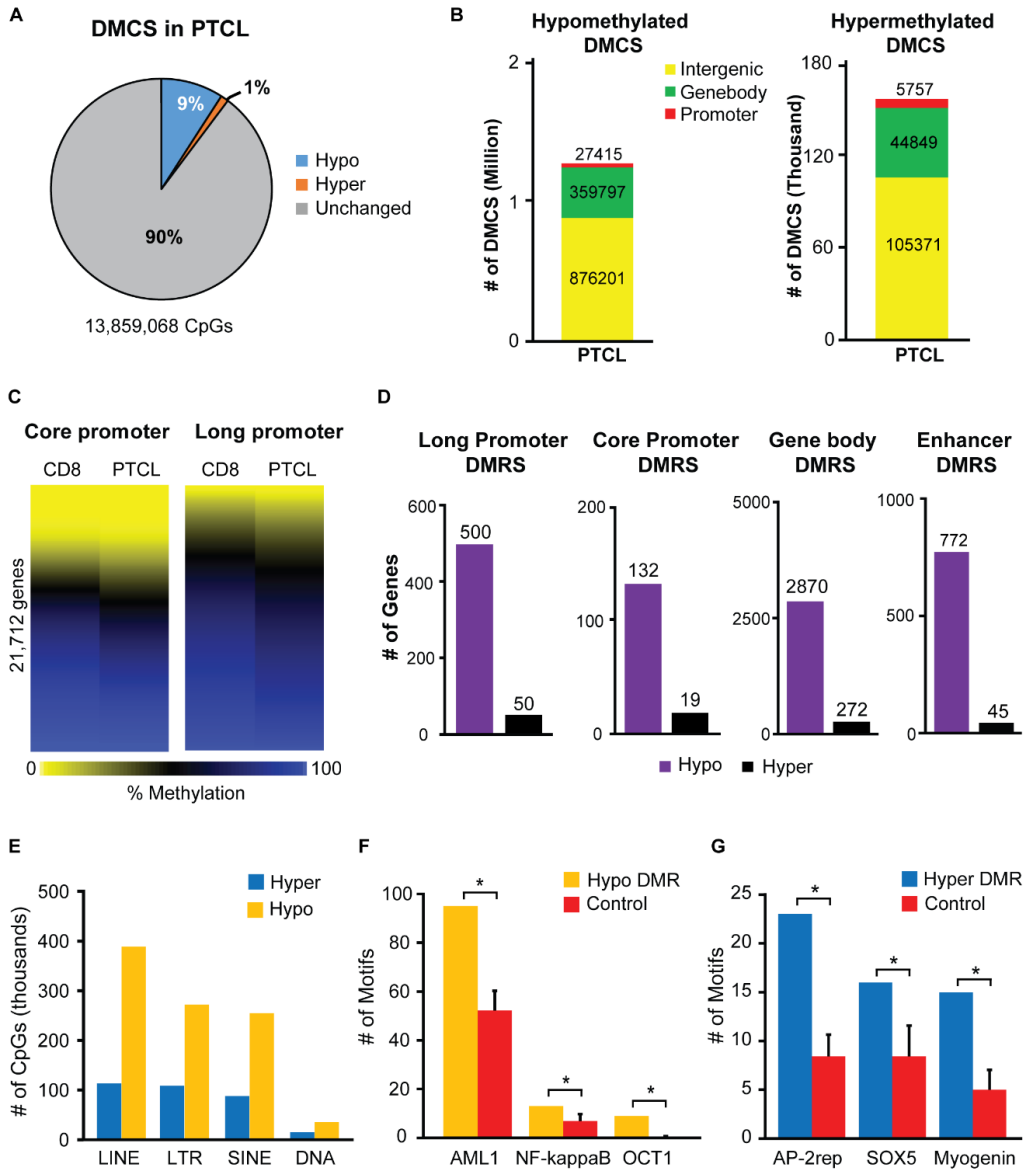


Figure 6. ***Dnmt3a*$^{\Delta/\Delta}$ PTCL is characterized by genome wide hypomethylation.** (A) Percentage of differentially methylated CpGs (DMCS) in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to *wild-type* CD8+ cells. Differentially methylated CpGs (DMCS) are defined as either hypo- (blue) or hypermethylated (orange) by a ≥30% change in percent CpG methylation in tumor samples compared to *wild-type* control samples. CpGs not meeting these criteria are shown in gray (unchanged). (B) Genomic location of hypomethylated (left) and hypermethylated (right) DMCS in *Dnmt3a*$^{\Delta/\Delta}$ PTCL, as compared to *wild-type* CD8+ cells. DMCS were annotated to long gene promoters (-1500 to +500bp relative to TSS), gene bodies, or intergenic regions. (C) Heat map comparing the methylation status of 21,712 promoters in *wild-type* CD8+ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples, as determined by WGBS. The averaged percent CpG methylation at core promoter regions (-300bp to +150bp relative to the TSS, *left*) and at long promoter regions (-1500bp to +500bp relative to the TSS, *right*) are displayed. (D) The number of genes with differentially methylated regions (DMRS) in their long promoters, core promoters, gene bodies and predicted enhancers in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8+ control. (E) The number of hypo- and hypermethylated CpGs present in LINE, LTR, SINE, and DNA repeat elements in *Dnmt3a*$^{\Delta/\Delta}$ PTCL, as compared to *wild-type* CD8+ cells. (F) The number of AML1, NF-κB, and OCT1 transcription factor motifs present in hypomethylated promoters in *Dnmt3a*$^{\Delta/\Delta}$ PTCL, as compared to *wild-type* CD8+ cells (red). The average motif count of 12 randomly generated control promoter sets is shown in red, with error bars denoting standard deviation. (*) denotes $p<0.05$ by a Wilcoxon rank test. (G) The number of AP-2rep, SOX5, and myogenin transcription factor motifs present in hypermethylated promoters in *Dnmt3a*$^{\Delta/\Delta}$ PTCL, as compared to *wild-type* CD8+ cells (blue). The average motif count of 12 randomly generated control promoter sets is shown in red, with error bars denoting standard deviation. (*) denotes $p<0.05$ by a Wilcoxon rank test.

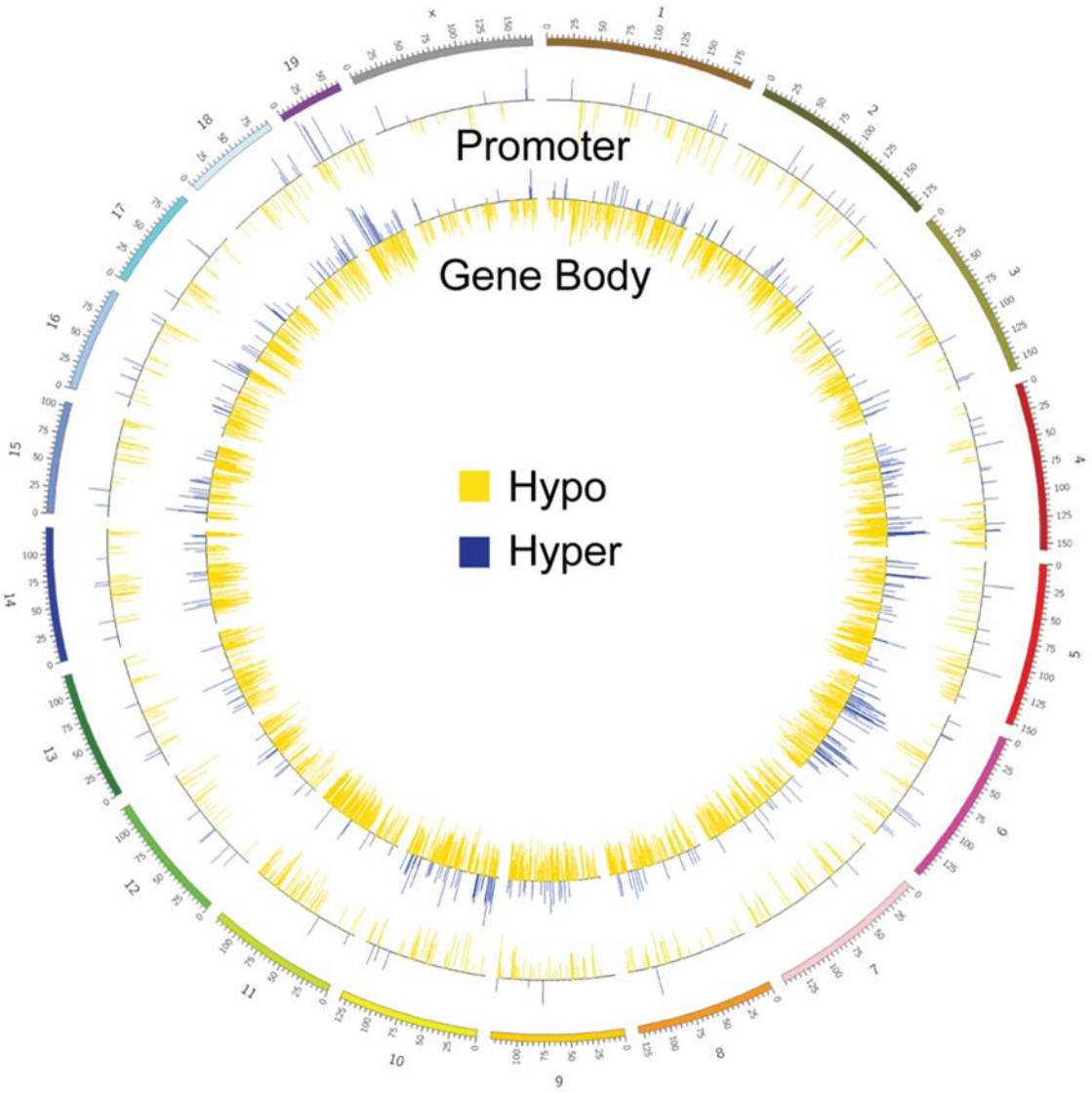


Figure 7. Promoter and gene-body hypomethylation is present throughout the genome of *Dnmt3a*$^{\Delta/\Delta}$ PTCLs. Circos plot of DMRS annotated to promoters and gene bodies in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to *wild-type* CD8+ cells. DMRs aligning to promoter (outer circle) and gene body (inner circles) are displayed in relation to its chromosomal position in the mouse genome. Hypomethylated DMRS are indicated by yellow lines and hypermethylated DMRS are indicated by blue lines.

Promoter hypomethylation is conserved across multiple *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ mouse lymphomas.

To determine whether the methylation landscape generated by WGBS is specific to the PTCL sample profiled or rather represents common changes that occur in Dnmt3a-deficient lymphoma, we validated hypo- and hypermethylated promoters using reduced representation bisulfite sequencing (RRBS) on additional normal CD8+ T cells and *Dnmt3a*$^{\Delta/\Delta}$ T cell lymphomas. This analysis confirmed hypomethylation of 90 gene promoters identified by WGBS in *Dnmt3a*$^{\Delta/\Delta}$ PTCL sample (**Fig 8A, 8B** and data not shown). In addition, 31 out of 38 gene promoters were confirmed to be hypermethylated by RRBS (**Fig 8A, 8B** and data not shown). The lesser extent to which hypomethylated promoters were confirmed by RRBS is not surprising in view of the inherent bias of the RRBS method which tends to underestimate the number of hypomethylated events in promoters with low CG content [126]. In fact, analysis of CpG content across DMRS revealed that hypomethylated promoters represent regions of lower CpG content when compared to hypermethylated promoters (**Fig 8C**). Altogether, these data are in good agreement with results obtained from WGBS, which show large scale promoter hypomethylation in Dnmt3a-defficient PTCLs.

To further validate data obtained by global methods on *Dnmt3a*$^{\Delta/\Delta}$ PTCL and to assess if methylation patterns are conserved in *Dnmt3a*$^{+/-}$ PTCL, we performed locus-specific methylation analysis using Combined Bisulfite Restriction Analysis (COBRA) for 11 selected genes in multiple independent tumor samples from *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ mice. Consistently with results obtained by WGBS, promoters of *Coro2a, Cxcr5, Ikzf3, Il2Rβ, Jdp2, Lpar5, Oas3, Ppil1, Pvt1, RacGAP1, and Wnt8a* were found to be hypermethylated in normal CD8+ T cells but hypomethylated in three independent *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples (**Fig 9**). Furthermore, all 11 promoters were hypomethylated in *Dnmt3a*$^{+/-}$ PTCL samples, suggesting that loss of a single

Dnmt3a allele is sufficient to induce patterns of promoter hypomethylation similar to those observed in *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples (**Fig 9**). To determine if promoter hypomethylation observed in *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ PTCL occurs as a result of Dnmt3a inactivation in normal CD8+ T cells due to the lack of Dnmt3a's *de novo* or maintenance activity, we analyzed promoter methylation in CD8+ T cells isolated from 8-week old *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ mice. For all 11 genes tested, promoters were hypermethylated in *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ CD8+ T cells to a similar degree as in *Dnmt3a*$^{+/+}$ CD8+ T cells, suggesting that partial or complete inactivation of Dnmt3a does not affect the methylation status of these genes in normal CD8+ T cells during development (**Fig 9**). Altogether, our data demonstrate that changes in promoter methylation identified using WGBS likely represent tumor-specific events occurring in mouse PTCL driven by mono or bi-allelic loss of Dnmt3a.

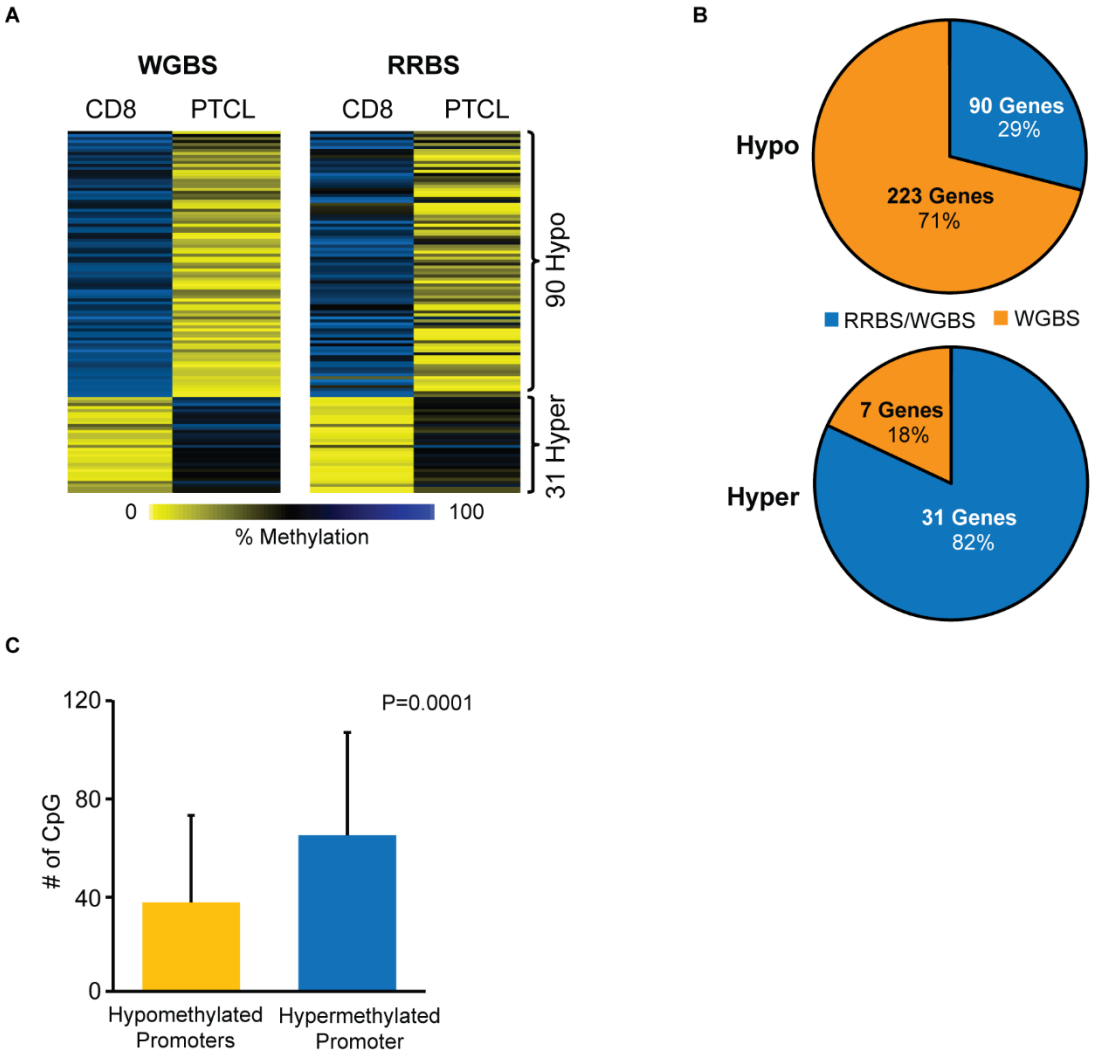


Figure 8. RRBS analysis. (A) Heat map displaying 90 hypomethylated and 31 hypermethylated promoters identified by WGBS and confirmed by RRBS. RRBS data is shown as the average percent methylation of DMCS annotated to long promoters (-1500 to +500 relative to TSS) for *Dnmt3a*$^{+/+}$ CD8+ T cells (n=2) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (n=2). DMCS are defined by a ≥30% change in percent methylation in tumor samples compared to wild-type control samples. (B) RRBS confirmation of differentially methylated promoters identified by WGBS. Hypomethylated (top) and hypermethylated (bottom) genes confirmed by RRBS are shown in blue. Differentially methylated gene promoters identified by WGBS, but not confirmed RRBS are shown in orange. (C) The average number of CpG dinucleotides present in hypo- and hypermethylated promoter regions (-500 to +1500 bp relative to TSS) in *Dnmt3a*$^{\Delta/\Delta}$ PTCL, as compared to *Dnmt3a*$^{+/+}$ CD8+ T cells. Error bars show standard deviation.

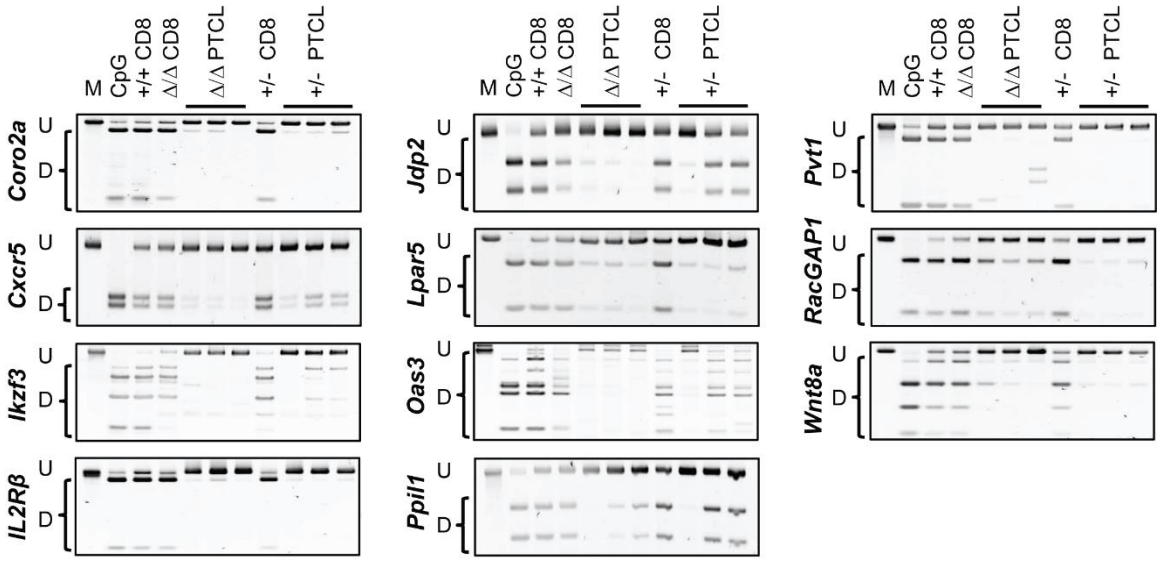


Figure 9. Promoter hypomethylation is conserved across multiple *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ mouse lymphomas. COBRA analysis of promoter methylation for *Coro2a, Cxcr5, Ikzf3, IL2Rβ, Jdp2, Lpar5, Oas3, Ppil1, Pvt1, Racgap1* and *Wnt8a* in *wild-type* CD8+, *Dnmt3a*$^{\Delta/\Delta}$ pre-tumor CD8+, *Dnmt3a*$^{\Delta/\Delta}$ PTCL, *Dnmt3a*$^{+/-}$ pre-tumor CD8+, and *Dnmt3a*$^{+/-}$ PTCL samples. PCR fragments amplified from bisulfite-treated genomic DNA were digested with the restriction enzymes *BstU*I, *Taq*1 or *Tai*I. Undigested (U) and digested (D) fragments correspond to unmethylated and methylated DNA, respectively. Control PCR fragments generated from fully methylated mouse genomic DNA that is undigested (M) or digested (CpG) are shown.

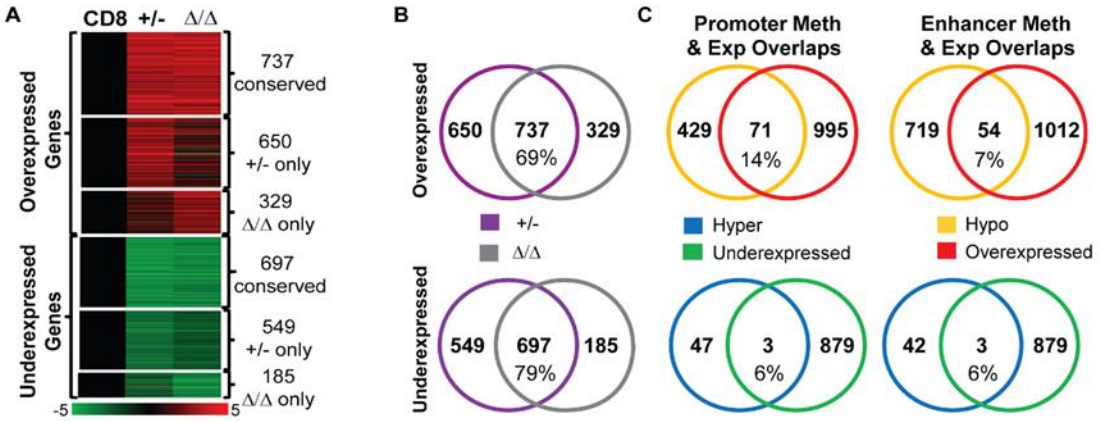
Gene expression is deregulated in *Dnmt3a*$^{\Delta/\Delta}$ PTCL.

To better understand deregulated molecular events in PTCL induced by mono or bi-allelic loss of Dnmt3a we performed global gene expression profiling of *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ PTCLs using RNA-seq. Comparison of gene expression patterns obtained from lymphomas to patterns obtained from normal CD8+ T cells revealed that *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ PTCLs shared strikingly similar expression profiles. In total, 737 (69%) overexpressed and 697 (79%) underexpressed genes were conserved between *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ PTCLs relative to CD8+ T cell controls (**Fig 10A, 10B** and data not shown). We also identified 329 upregulated and 185 downregulated genes specific to *Dnmt3a*$^{\Delta/\Delta}$ PTCL, as well as 650 upregulated and 549 downregulated genes specific to *Dnmt3a*$^{+/-}$ PTCL (**Fig 10A, 10B** and data not shown). Altogether, these data suggest that molecular events driving T cell transformation in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ mice are likely conserved.

IPA of differentially expressed genes in PTCL identified 3 Inhibited Pathways common to both *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (*Tec Kinase signaling, Type I Diabetes Mellitus Signaling, 4-1BB Signaling in T Lymphocytes*) and 1 commonly Activated Pathways (*Cyclins and Cell Cycle regulation*) in *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ lymphomas (**Fig 11**). The top 5 categories for "diseases and disorders" were identical for both *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ tumors (Inflammatory response, Immunological disease, Connective tissue disorder, Inflammatory disease and Skeletal and muscular disorders), further illustrating the similarities between their molecular landscapes.

Comparison of methylation and gene expression revealed that 71 genes (14%) whose promoters were hypomethylated in PTCL were associated with overexpression (**Fig 10C and 10D,** referred to herein as HOT genes – Hypomethylated and overexpressed in TCL). In contrast, we detected only three genes - *CD226*, *Fhit*, and *Emp1* - whose hypermethylation correlated with underexpression, suggesting that most of the cancer-specific hypermethylation has little effect on

gene expression and tumor progression (**Fig 10C**). Further analysis revealed 54 genes (7%) whose predicted enhancer regions were hypomethylated in PTCL and were also overexpressed, whereas only 3 genes with hypermethylated enhancers were downregulated (**Fig 10C and 12**). Altogether, these data demonstrate that hypomethylation affects gene expression on a broader scale than hypermethylation in mouse *Dnmt3a*$^{\Delta/\Delta}$ PTCL and thus may functionally contribute to a disease development.



D

Gene	CD8 % Meth	PTCL % Meth	Exp	Gene	CD8 % Meth	PTCL % Meth	Exp
1700025G04Rik	0.79	0.2	6.7	Klrd1	0.8	0.1	2.4
1700048O20Rik	0.51	0.06	12.5	Klre1	0.85	0.075	24.2
2310001H17Rik	0.76	0.15	2.0	Lpar5	0.85	0.18	9.9
2810006K23Rik	0.88	0.12	3.9	Lyrm9	0.53	0.05	4.0
4921525O09Rik	0.77	0.2	9.0	Mmp14	0.86	0.22	20.6
Abi3	0.76	0.12	3.3	Myo1c	0.91	0.23	2.2
Acot7	0.69	0.09	3.0	Myo6	0.61	0.07	7.0
Alpk3	0.71	0.11	4.7	N4bp1	0.91	0.34	4.1
Amica1	0.87	0.12	6.0	Nod1	0.76	0.13	2.3
Apobr	0.7	0.07	10.3	Oas2	0.91	0.42	21.3
Arap3	0.67	0.09	3.8	Oas3	0.82	0.14	10.8
Arl4d	0.815	0.085	9.1	Osbpl3	0.78	0.17	15.9
Atp8b4	0.81	0.01	4.0	P2rx7	0.86	0.19	4.1
AW112010	0.78	0.11	4.2	Per3	0.87	0.12	4.7
B4galt5	0.75	0.06	2.4	Pif1	0.74	0.27	26.7
Celsr1	0.83	0.28	2.1	Plac8	0.73	0.06	19.6
Cipc	0.91	0.21	2.5	Plekha8	0.71	0.1	2.5
Cln3	0.7	0.07	3.3	Ppil1	0.55	0.04	5.6
Coro2a	0.82	0.125	4.0	Pvt1	0.95	0.23	2.7
Fah	0.62	0.06	4.0	Reep5	0.63	0.14	3.0
Fcgrt	0.94	0.5	3.2	Rnasel	0.92	0.08	3.4
Gdf11	0.67	0.12	4.8	Rnf43	0.82	0.08	4.8
Gimap4	0.64	0.02	2.9	Samd3	0.82	0.1	8.3
Gimap7	0.7	0.03	2.7	Sh2d1a	0.85	0.37	2.9
Gm19705	0.66	0.19	7.4	Sla	0.85	0.19	2.9
Gpnmb	0.71	0.04	198.9	Sntb2	0.9	0.37	2.0
Ifitm10	0.83	0.08	3.3	Stat1	0.85	0.14	7.5
Ikzf3	0.94	0.1	2.0	Tm6sf1	0.58	0.03	2.4
Il18rap	0.85	0.26	2.5	Tmem37	0.68	0.02	6.9
Il2rb	0.88	0.1	2.5	Tmevpg1	0.88	0.18	12.0
Impa1	0.74	0.1	2.9	Trim14	0.88	0.34	2.6
Islr	0.84	0.21	1251.4	Ttll11	0.74	0.07	14.5
Jdp2	0.85	0.21	4.4	Wdr86	0.92	0.42	12.2
Keap1	0.86	0.31	2.9	Wnt8a	0.755	0.17	307.5
Klhl30	0.74	0.11	83.8	Zfp808	0.7	0.11	2.0
Klrc1	0.77	0.07	4.0				

Figure 10. Gene expression is deregulated in *Dnmt3a*$^{\Delta/\Delta}$ PTCL. (A) Heat map of RNA-seq global expression data displaying differentially expressed genes in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (≥2 fold change and a p-value < 0.05) relative to *wild-type* CD8+ cells, 737 genes were overexpressed and 697 genes were underexpressed in both *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL. 650 overexpressed and 549 underexpressed genes were specific to *Dnmt3a*$^{+/-}$ PTCL, whereas 329 overexpressed and 185 underexpressed genes were only observed in *Dnmt3a*$^{\Delta/\Delta}$ PTCL. A color bar displays fold change in gene expression with overexpressed shown in red and underexpressed in green. (B) Venn diagram showing overlap in over- and underexpressed genes between *Dnmt3a*$^{+/-}$ PTCL (purple) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (gray). (C) (left) Venn diagram showing overlap between the number of differentially expressed genes (red = overexpression; green = underexpression) and the number of differentially methylated gene promoters (yellow = hypomethylation; blue = hypermethylation) in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to *wild-type* CD8+ controls. (right) Overlap between differentially expressed genes and differentially methylated enhancers regions. (D) List of genes hypomethylated and overexpressed in *Dnmt3a*$^{\Delta/\Delta}$ PTCL (HOT genes). Promoter methylation for *wild-type* CD8+ cells and *Dnmt3a*$^{\Delta/\Delta}$ PTCL is shown in blue and yellow, respectively. Corresponding changes in gene expression for *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to control CD8+ samples are shown in red.

Pathway	Het # mol	Het z-score	Mut # mol	Mut z-score	Activity
Cyclins and Cell Cycle Regulation	19	2.828	17	2.5	Activated
Tec Kinase signaling	37	-1.567	27	-1.528	Inhibited
Type I Diabetes Mellitus Signaling	24	-1.789	18	-1.94	Inhibited
4-1BB Signaling in T Lymphocytes	9	-2.828	7	-2.45	Inhibited

Top Diseases and Disorders	Het # mol	Mut # mol
Inflammatory response	357	273
Imunological disease	230	166
Connective tissue disorder	81	86
Inflammatory disease	123	108
Skeletal and muscular disorders	81	68

Figure 11. IPA analysis. Summary of top categories, including "pathways" and "diseases and disorders", derived from Ingenuity pathway analysis (IPA) of genes differentially expressed in both *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to *Dnmt3a*$^{+/+}$ CD8+ T cell controls. P<0.05 for all categories.

Gene	CD8 % Meth	PTCL %Meth	Exp		Gene	CD8 % Meth	PTCL %Meth	Exp
4933432G23Rik	0.73	0.19	4.1		Lgals3bp	0.89	0.20	5.2
A930006K02Rik	0.61	0.07	6.0		Lhfp	0.88	0.39	3.6
Bsn	0.69	0.13	13.8		Lmnb1	0.82	0.37	3.1
Cdc6	0.58	0.06	14.1		Mgll	0.79	0.05	4.2
Cdca7l	0.78	0.22	2.6		Mina	0.84	0.09	3.2
Cdkn2b	0.80	0.31	15.7		Myo5a	0.85	0.13	2.5
Celsr1	0.78	0.27	2.1		Nabp1	0.73	0.26	2.5
Chrne	0.58	0.07	74.2		Ndrg1	0.86	0.16	5.2
Clic5	0.78	0.19	6.5		Nedd4	0.77	0.24	4.0
Crybg3	0.81	0.12	3.9		Osbpl3	0.82	0.18	15.9
Dtnbp1	0.66	0.06	2.3		Pfkp	0.92	0.31	2.2
Entpd1	0.87	0.20	7.5		Plcb3	0.59	0.02	2.0
Fads6	0.63	0.18	87.9		Reep5	0.84	0.20	3.0
Fam185a	0.74	0.11	3.9		Rhoq	0.57	0.03	3.7
Fignl1	0.52	0.05	11.5		Rmi2	0.85	0.13	11.1
Galm	0.71	0.18	2.4		Rnf43	0.63	0.10	4.8
Gimap4	0.78	0.17	2.9		Rtp4	0.93	0.42	19.9
Gimap7	0.51	0.05	2.7		Sdc3	0.70	0.13	3.7
Gm12504	0.81	0.08	3.2		Setbp1	0.80	0.20	7.0
Grhl3	0.82	0.06	5.7		Socs1	0.75	0.25	2.9
Grtp1	0.80	0.33	7.7		Socs2	0.78	0.27	6.9
Hdac11	0.59	0.02	4.2		St3gal6	0.88	0.18	4.0
Kbtbd11	0.81	0.21	5.6		Stat1	0.87	0.16	7.5
Kcnq5	0.75	0.11	2.5		Tmevpg1	0.82	0.17	12.0
Kif14	0.87	0.31	13.7		Tspan9	0.79	0.06	12.5
Klhl5	0.74	0.07	2.5		Uprt	0.76	0.16	3.7
Lactb2	0.73	0.12	2.9		Zbp1	0.90	0.35	7.3

Figure 12. Overexpressed genes with enhancer hypomethylation. List of genes that are overexpressed and whose predicted enhancer regions are hypomethylated in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to *Dnmt3a*$^{+/+}$ CD8+ T cell controls. Percent methylation derived from WGBS for enhancer regions for *Dnmt3a*$^{+/+}$ CD8+ T cells and *Dnmt3a*$^{\Delta/\Delta}$ PTCL is shown in blue (representing high levels of methylation) and yellow (representing low levels of methylation). Corresponding fold changes in gene expression (determined by RNA-seq) for *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to control *Dnmt3a*$^{+/+}$ CD8+ samples are shown in red.

Gene expression changes are partially conserved between mouse and human PTCL.

To determine the extent of similarity between mouse and human disease on the molecular level, we compared gene expression signatures obtained from mouse PTCL to those derived from human PTCL. We utilized microarray data obtained on a set of five normal tonsil T cells and three human PTCL samples with predicted inactivating Dnmt3a mutations [42]. When we compared expression of genes deregulated in human PTCL to those genes deregulated in either *Dnmt3a*$^{+/-}$ or *Dnmt3a*$^{\Delta/\Delta}$ PTCL, we identified 316 (28%) overexpressed and 415 (36%) underexpressed genes that were shared between human and *Dnmt3a*$^{+/-}$ PTCL (**Fig 13A, 13B** and data not shown). Fewer genes (252 overexpressed and 239 underexpressed) were shared between human PTCL and *Dnmt3a*$^{\Delta/\Delta}$ PTCL, suggesting that the transcriptome of lymphomas induced by loss of a single Dnmt3a allele more so resembles human disease than those that arise due to full inactivation of Dnmt3a (**Fig 13A, 13B** and data not shown). The extent of overlap in up-and downregulated genes was significant for all comparisons ($P<0.01$), suggesting that similar molecular events may drive PTCL in both species.

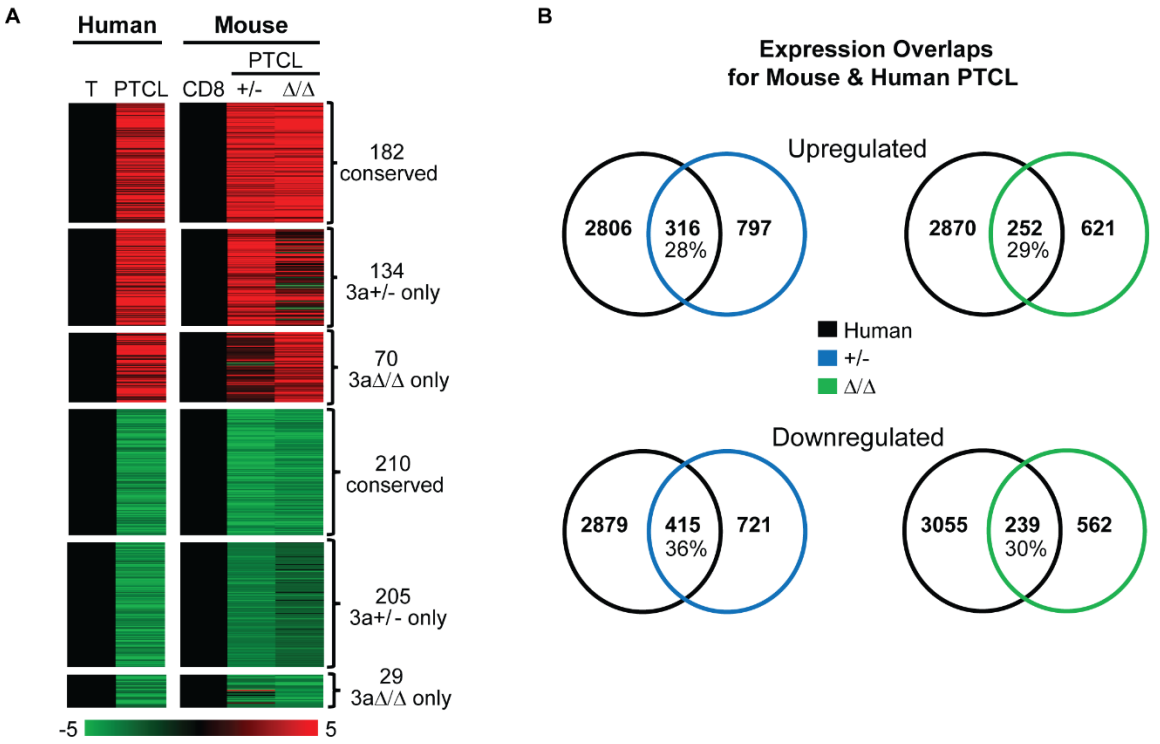


Figure 13. Gene expression changes are partially conserved between mouse and human PTCL. (A) Heat maps derived from global expression profiling for genes differentially expressed in both human PTCL (PTCL) relative to normal tonsil T cells (T) and in *Dnmt3a*$^{+/-}$ PTCL (+/-) and/or *Dnmt3a*$^{\Delta/\Delta}$ PTCL (Δ/Δ) relative to normal *Dnmt3a*$^{+/+}$ CD8+ cells. 182 genes were commonly overexpressed in human PTCL, *Dnmt3a*$^{+/-}$ PTCL, and *Dnmt3a*$^{\Delta/\Delta}$ PTCL, while 210 genes were commonly underexpressed in all three tumor types. 134 overexpressed and 205 underexpressed genes were specific to human PTCL and *Dnmt3a*$^{+/-}$ PTCL, whereas 70 overexpressed and 29 underexpressed genes were only observed in human PTCL and *Dnmt3a*$^{\Delta/\Delta}$ PTCL. For microarray data (human samples), genes with a fold change >1.5 and a P-value <0.05 were considered significant. For RNA-seq data (mouse samples), genes with a fold change >2 and a p-value <0.05 were considered significant. A color bar displays fold change in gene expression with overexpressed shown in red and underexpressed in green. (B) Venn diagrams showing overlaps in gene expression between human PTCL (black) and mouse *Dnmt3a*$^{+/-}$ PTCL (blue) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (green), as determined in panel A of the figure.

***Jdp2* is hypomethyalted and overexpressed in mouse and human PTCL**.

Because promoter hypomethylation resulted in upregulated gene expression in PTCL we next asked whether any of the HOT genes may have potential oncogenic functions in the development of T cell lymphomas and whether such genes are also hypomethylated and overexpressed in human PTCLs. One such candidate gene with oncogenic function in the T cell compartment - Jun Dimerization Protein 2 protein (Jdp2) - is a component of the AP-1 transcription factor that was reported to negatively regulate *Trp53* and promote the development of T cell leukemia in mice [127]. Consistently with global WGBS data, Jdp2 was hypomethylated in *Dnmt3a*$^{\Delta/\Delta}$ PTCLs as determined by COBRA (**Fig 14A**). Likewise, with RNA-seq data, analysis of *Jdp2* transcript levels by qRT-PCR confirmed overexpression of *Jdp2* in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples (**Fig 14B**). Next, we analyzed the methylation status of the JDP2 promoter in human CD3 T cells and PTCL samples and found that like in Dnmt3a-deficient mouse PTCL, the JDP2 promoter is hypomethylated in human PTCL relative to controls (**Fig 14C**). To determine whether overexpression of JDP2 occurs in human PTCL, we compared transcript levels in a set of 8 human PTCLs to normal CD3+ T cells. This analysis showed ~20-1,700-fold increase in JDP2 levels in human PTCL relative to normal T cells (**Fig 14D**). These data demonstrate that JDP2 promoter hypomethylation correlates with its overexpression in human PTCL. To evaluate the role of Jdp2 in tumor maintenance we used a shRNA construct to knockdown the levels of Jdp2 in a Dnmt3a-deficient MYC-induced PTCL cell line. However, decrease in the levels of Jdp2 in this cell line did not affect cellular growth, suggesting Jdp2 is not required for tumor maintenance in such setting (**Fig 15A and 15B**). Overall, our data indicate that methylation likely plays a role in the regulation of Jdp2 in mouse and human PTCL.

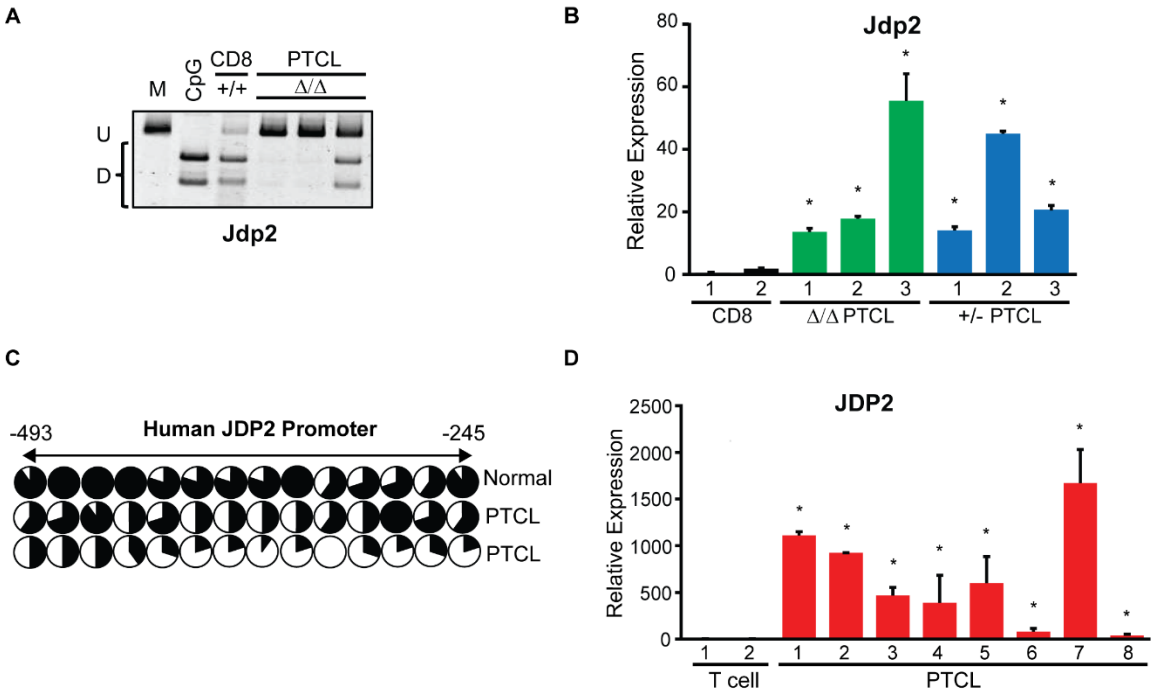


Figure 14. Jdp2 is hypomethylated and overexpressed in human and mouse PTCL. (A) COBRA analysis of mouse Jdp2 promoter methylation in three independent *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples. Undigested (U) and digested (D) fragments correspond to unmethylated and methylated DNA, respectively. Control PCR fragments generated from fully methylated mouse genomic DNA that is undigested (M) or digested (CpG) are shown. (B) Normalized gene expression of Jdp2 in mouse CD8+ control, *Dnmt3a*$^{+/-}$ PTCL, and *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples by qRT-PCR. Data presented are the average of two independent experiments. Error bars show standard deviation and an asterisk (*) denotes a $p<0.05$ (student t-test). (C) Bisulfite sequencing of the JDP2 promoter in normal human CD3+ T cells and in two independent human PTCL samples. Circles represent individual CpGs within the promoter. Black and white areas denote the relative portion of methylated and un-methylated sequence reads at a CpG, respectively. (D) Normalized gene expression of JDP2 in normal human CD3+ T cells and human PTCL samples, by qRT-PCR. Data presented are the average of two independent experiments. Error bars show standard deviation and an asterisk (*) denotes a $p<0.05$ (student t-test).

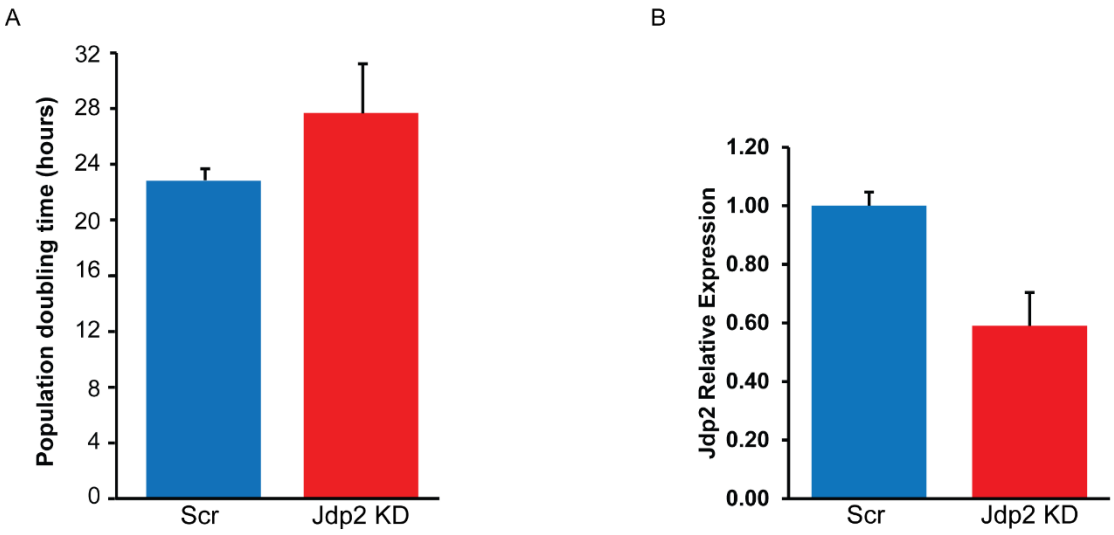


Figure 15. Knockdown of Jdp2 in MYC-induced *Dnmt3a*$^{-/-}$ T cell lymphoma does not affect cellular growth *in vitro*. (A) Average population doubling time for a *Dnmt3a*$^{-/-}$ MYC-induced T cell lymphoma line infected with either scrambled shRNA (blue) or shRNA against Jdp2 (red). Error bars show standard deviation. (B) Normalized gene expression of *Jdp2* transcript levels as determined by qRT-PCR for a *Dnmt3a*$^{-/-}$ MYC-induced T cell lymphoma line infected with either scrambled shRNA (blue) or shRNA against Jdp2 (red). Error bars show standard deviation.

P53 is downregulated in pretumor thymocytes and in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCLs.

Because Jdp2 was reported to negatively regulate p53 transcript levels we analyzed *Trp53* expression by qRT-PCR. Despite a 10-70-fold increase in *Jdp2* levels in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples, we did not observe any effects on *Trp53* transcript levels (**Fig 16**), suggesting that in this setting Jdp2 overexpression has no direct effect on p53 transcription. However, analysis of p53 protein levels showed downregulation of p53 in all *Dnmt3a*$^{+/-}$ tumors and 3/4 *Dnmt3a*$^{\Delta/\Delta}$ tumors, suggesting that Jdp2 may regulate p53 at the protein level or p53 downregulation in tumors occurs independently of Jdp2 overexpression (**Fig 17A**). Gene Set Enrichment Analysis (GSEA) using RNA-seq data from normal *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL revealed significant downregulation of the p53 pathway genes in both settings (**Fig 17B, 17C** and data not shown). To determine when p53 is downregulated during lymphomagenesis, we next measured protein levels in spleens, lymph nodes and thymi isolated from 9 months old *Dnmt3a*$^{+/+}$ and *Dnmt3a*$^{+/-}$ mice. At this age, *Dnmt3a*$^{+/-}$ mice did not show any sign of lymphomagenesis or cellular changes in the hematopoietic compartment (**Fig 18**). Interestingly, whereas p53 levels in spleens and lymph nodes were similar between *Dnmt3a*$^{+/+}$ and *Dnmt3a*$^{+/-}$ settings, p53 was downregulated in thymocytes of *Dnmt3a*$^{+/-}$ mice (**Fig 17D**). To determine whether downregulation of p53 occurs as a direct response to Dnmt3a monoallelic loss we analyzed p53 protein levels in splenocytes and thymocytes of 6 weeks old *Dnmt3a*$^{+/+}$ and *Dnmt3a*$^{+/-}$ mice. This analysis revealed no apparent differences in p53 levels, suggesting that loss of one allele of Dnmt3a is insufficient to downregulate p53 at this time point (**Fig 17E**). Because we did not succeed in establishing CD8+ lymphoma cell lines, we utilized previously generated MYC-induced *Dnmt3a*$^{+/+}$ and *Dnmt3a*$^{-/-}$ T cell lymphoma cell lines [92], along with MSCV-IRES-p53-GFP overexpressing both p53 and EGFP [21], to evaluate the role of p53 in lymphomagenesis. Overexpression of p53 induced selection against EGFP-positive cells in both *Dnmt3a*$^{+/+}$ and *Dnmt3a*$^{-/-}$ T cell lymphoma cell lines, suggesting that exogenous p53 inhibited cellular proliferation *in vitro* (**Fig 17F**). This result suggests that low p53 levels are

important for tumor maintenance *in vitro* and therefore downregulation of p53 *in vivo* is likely an important event in tumorigenesis. Altogether, these data suggest that downregulation of p53 is chronologically an intermediate event in lymphomagenesis and therefore likely a relevant in initiation/progression of lymphomagenesis and may be mediated by upregulation of Jdp2.

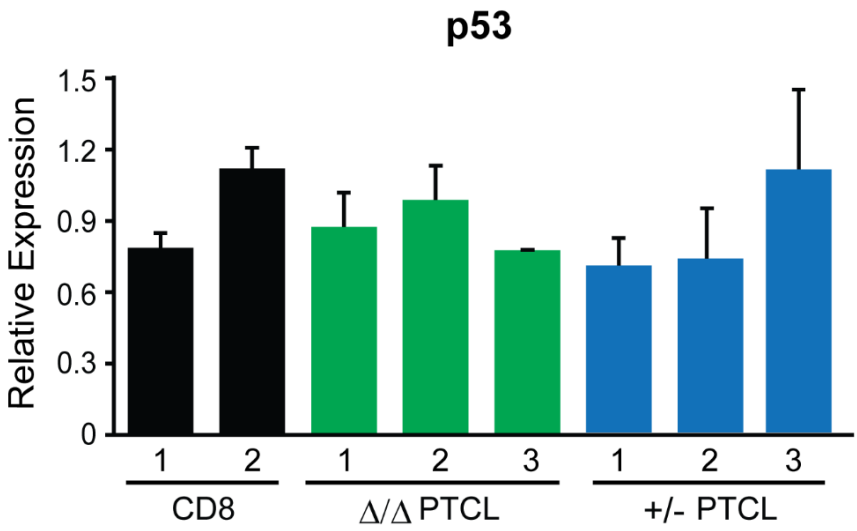


Figure 16. *Tp53* transcript levels are unchanged in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL. Normalized gene expression of *Tp53* transcript levels as determined by qRT-PCR in mouse *Dnmt3a*$^{+/+}$ CD8+ T cell control, *Dnmt3a*$^{+/-}$ PTCL, and *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples. Data presented are the average of two independent experiments. Error bars show standard deviation.

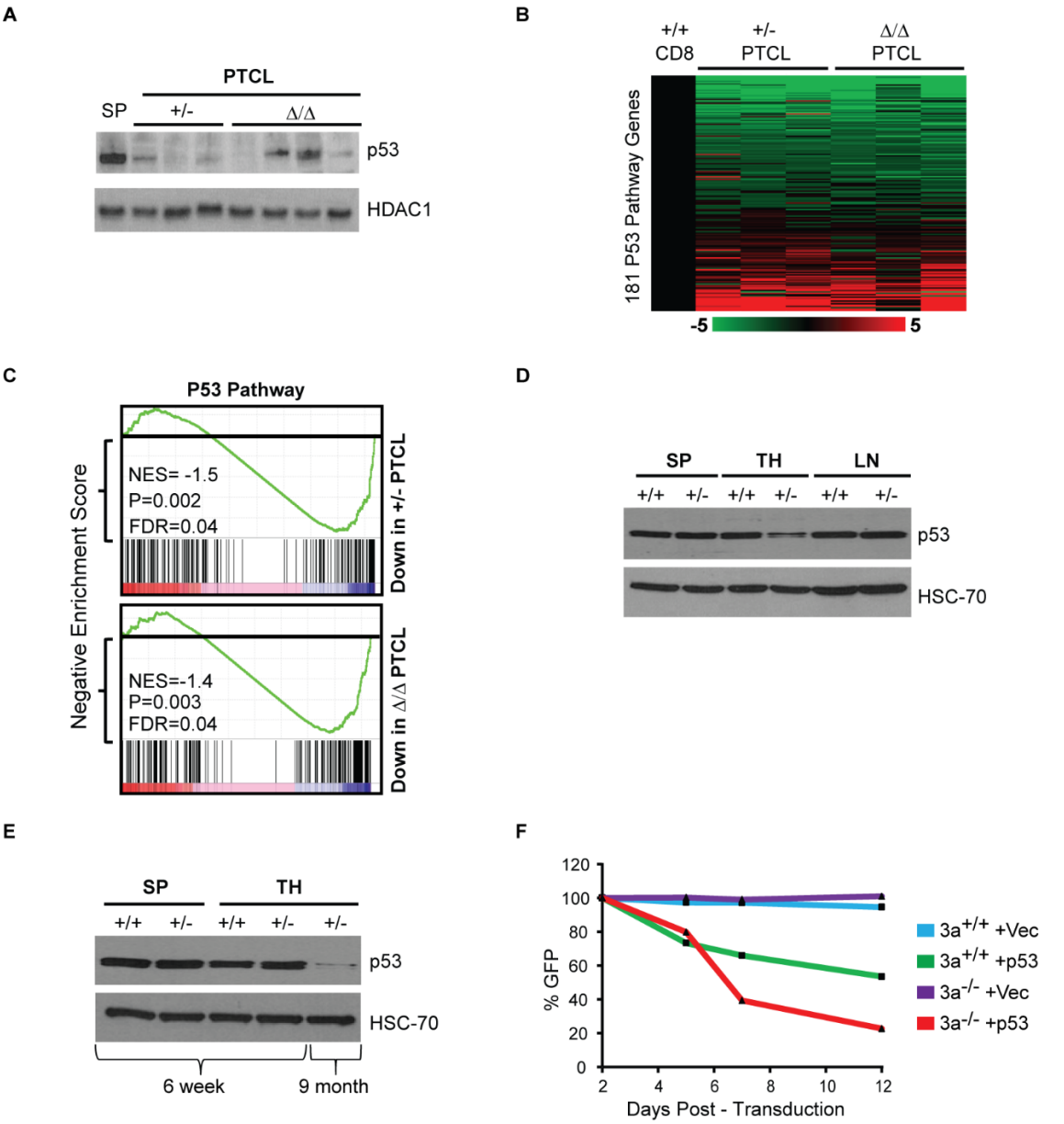


Figure 17. P53 is downregulated in pretumor thymocytes and in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCLs. (A) Immunoblot showing p53 protein levels in *Dnmt3a*$^{+/+}$ control spleen (SP), *Dnmt3a*$^{+/-}$ PTCL (+/-) *Dnmt3a*$^{\Delta/\Delta}$ PTCL (Δ/Δ) samples. HDAC1 is shown as a loading control. (B) Heatmap showing fold change in gene expression data derived from RNA-seq of 181 p53 pathway genes identified through GSEA for *Dnmt3a*$^{+/-}$ PTCL and *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8+ controls. (C) RNA-seq expression profiles on *Dnmt3a*$^{+/-}$ PTCL, and *Dnmt3a*$^{\Delta/\Delta}$ PTCL were subjected to GSEA to identify enriched signatures. Each group was run as a pairwise comparison to normal CD8+ cells. In both tumor groups, the p53 pathway was identified as being significantly downregulated relative to controls. Normalized enrichment scores (NES), false discovery rate (FDR) and P-values are shown for each analysis. Black bars indicate individual genes within the pathway. Red indicates genes with high expression and blue indicates low expression in tumors relative to controls. (D) Immunoblot showing p53 protein levels in *Dnmt3a*$^{+/+}$ (+/+) and *Dnmt3a*$^{+/-}$ (+/-) cells isolated from the spleen (SP), thymus (TH), and lymph node (LN) of 9 month old mice with no signs of lymphoma development. HSC-70 is shown as a loading control. (E) Immunoblot showing p53 protein levels in *Dnmt3a*$^{+/+}$ (+/+) and *Dnmt3a*$^{+/-}$ (+/-) cells isolated from the spleen (SP) and thymus (TH) of 6 week and 9 month old mice with no signs of lymphoma development. HSC-70 is shown as a loading control. (F) Relative percentage of EGFP positive cells as determined by FACS for *Dnmt3a*$^{+/+}$ (+/+) and *Dnmt3a*$^{-/-}$ (-/-) lymphoma cell lines infected with either MSCV-IRES-EGFP (Vec) or MSCV-IRES-p53-EGFP (p53). EGFP measurements were taken at multiples time points and normalized by the percentage of EGFP positive cells observed 2 days post-transduction.

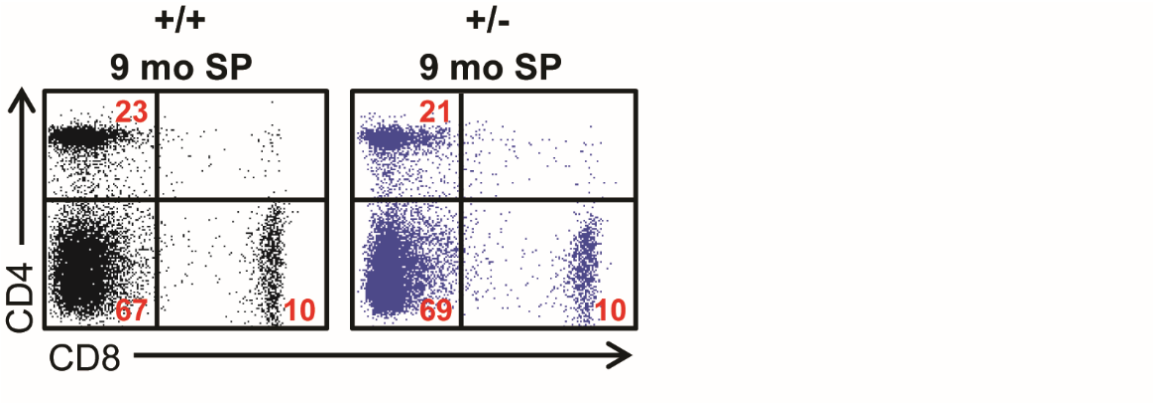


Figure 18. CD8+ T cells are not expanded in the spleen of a 9 months old *Dnmt3a*$^{+/-}$ mouse. CD4 and CD8 expression in cells isolated from the spleen of 9 months old *Dnmt3a*$^{+/+}$ (+/+) and *Dnmt3a*$^{+/-}$ (+/-) mice, as determined by FACS. Percentage of cells in each quadrant are shown in red.

Discussion

In this study we show that loss of Dnmt3a in HSPCs in *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*; Rosa26LOXP*$^{EGFP/EGFP}$ mice not only results in the development of CLL as we reported previously [79, 95] but also in the development of peripheral T cell lymphomas in ~40% of *Dnmt3a*$^{\Delta/\Delta}$ mice either alone or in combination with CLL. We further show that not only complete inactivation but also a reduction in Dnmt3a levels results in the development of PTCL in 10% of *Dnmt3a*$^{+/-}$ mice. Lymphomas that develop in both *Dnmt3a*$^{\Delta/\Delta}$ and *Dnmt3a*$^{+/-}$ mice are exclusively CD8+CD4- mature T cell lymphomas. Importantly, *Dnmt3a*$^{+/-}$ PTCLs retain expression of the Dnmt3a *wild-type* allele. Thus, consistent with heterozygous mutations of Dnmt3a found in human T cell malignancies, Dnmt3a is a haploinsufficient tumor suppressor gene in the prevention of mouse mature CD8+CD4- T cell lymphomas.

The growing number of various malignant phenotypes observed in the hematopoietic system with Dnmt3a-deficiency in mice raises questions about the nature of deregulated events induced by Dnmt3a inactivation. Because Dnmt3a is a methyltransferase, we were interested in finding whether genes deregulated in a methylation dependent manner could provide clues towards understanding the pathobiology of *Dnmt3a*$^{\Delta/\Delta}$ PTCLs. A combined analysis of global methylation and gene expression identified promoter hypomethylation as a major deregulated event in PTCL development in the absence of Dnmt3a with as many as 500 genes hypomethylated in lymphomas. Of these genes, expression of as many as 71 genes (14%) was upregulated in tumors. Since Dnmt3a has now been shown to be a tumor suppressor in the prevention of a number of hematologic malignancies in a variety of biological settings [79, 87-89, 95], it is therefore possible that promoter hypomethylation along with gene upregulation may be either a contributing factor or even the primary event driving the initiation/progression of tumor development. In such a scenario, proto-

oncogenes are silenced in normal cells but are progressively hypomethylated and overexpressed resulting in cellular transformation. Analysis of data derived from *Dnmt3a*$^{\Delta/\Delta}$ lymphomas identified several putative drivers of T cell transformation whose promoters were hypomethylated and overexpressed in tumors (HOT genes). One such HOT gene is the Interleukin-2 receptor *Il2rb*, a component of the IL-2 signaling pathway that is important for the growth of T lymphocytes. Inappropriate activation of this pathway may promote unchecked proliferation of T cells, thus contributing to tumorigenesis [128]. The HOT gene, *Stat1*, participates in cytokine signaling in T cells and has been reported to be significantly overexpressed in human PTCL-NOS [129]. In a mouse model of *v-abl*-induced leukemia, *Stat1*$^{-/-}$ mice were partially protected from the development of leukemia, demonstrating that *Stat1* possesses tumor-promoting activity [130]. Another HOT gene, *Trim14* was demonstrated to have oncogenic function in tongue squamous cell carcinoma cell lines by activating the NF-κB pathway [131]. Whereas HOT genes represent good candidates to explain the tumor suppressor function of Dnmt3a, demonstration of a causative oncogenic role in initiation/progression of lymphomagenesis for any of these genes is challenging as it requires long-term *in vivo* experiments in mice. Thus, only future functional studies can address the importance of these genes in the pathogenesis of *Dnmt3a*$^{\Delta/\Delta}$ PTCLs.

An additional HOT gene with predicted oncogenic activity is Jun Dimerization Protein 2 (JDP2), which we found to be hypomethylated and overexpressed not only in mouse PTCL but also in human PTCLs. Jdp2 protein is a component of the AP-1 transcription factor complex that represses transactivation mediated by the Jun family of proteins and it plays a role in AP-1-mediated responses in UV-induced apoptosis and cell differentiation [132]. Jdp2 was reported to promote liver transformation as JDP transgenic mice displayed potentiation of liver cancer, higher mortality and increased number and size of tumors [133]. Importantly, Jdp2 was identified in a screen for oncogenes able to collaborate with the loss of p27^{kip1} cyclin-dependent inhibitor to induce lymphomas [134]. Altogether these data along with our findings suggest that upregulation of Jdp2

induced by loss of Dnmt3a might be a contributing factor to the development of PTCL. Because Jdp2 was reported to negatively regulate *Trp53* on a transcriptional level and promote the development of T cell leukemia in mice [127] we tested whether Trp53 levels are affected in *Dnmt3a*$^{\Delta/\Delta}$ PTCLs. Despite a 15-70 fold increase in *Jdp2* levels we did not observe any changes in *Trp53* transcript levels, suggesting that in this setting Jdp2 overexpression has little effect on p53 transcription. However, western blot revealed decrease in p53 protein in the majority of tumor samples, suggesting that Jdp2 may regulate p53 by other mechanisms or p53 downregulation occurs through an independent pathway not involving Jdp2. Regardless of the mechanism by which p53 is downregulated in tumors, decreased p53 protein is likely contributing to CD8+ T cell transformation due to its strong tumor suppressor function in T cell compartment. For example, it was previously reported that *Trp53*$^{-/-}$ mice are highly susceptible to spontaneous tumor development, with the majority of mice developing immature CD4+CD8+ thymic lymphomas [135]. To the best of our knowledge, there are no studies in mice demonstrating the tumor suppressor function of p53 specifically in CD8+ T cell lymphomas. However, a loss of the region containing the p53 gene on chromosome 17 was observed in human primary cutaneous CD8+ cytotoxic T cell lymphoma, suggesting that low p53 levels could be involved in the pathogenesis of human CD8+ PTCL [136]. The fact that p53 was downregulated in thymocytes isolated from 9 month old, but not 6 week old, *Dnmt3a*$^{+/-}$ tumor-free mice indicates that p53 downregulation is chronologically an intermediate event in lymphomagenesis and this strongly suggest that this event is relevant in the initiation/progression of CD8+ PTCL. Consistent with downregulation of p53 protein levels, GSEA revealed suppression of p53 pathway genes in both *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL tumors, such as GADD45a, ZFP36L1, and KLF4. Studies using *Gadd45a*$^{-/-}$ mice found that ablation of Gadd45a in lymphoma-prone AKR mice decreased the latency and increased the incidence of T cell lymphomas, while deletion of Gadd45a on a p53 deficient background altered the tumor spectrum to heavily favor the development of T cell lymphomas [137]. Similarly, mice deficient for ZFP36L1 and ZFP36L2 displayed altered T cell development and readily succumbed

to CD8+ T cell acute lymphoblastic leukemia [138]. KLF4 was identified to be mutated in pediatric T-ALL patients [139] and was shown to induce apoptosis in primary T-ALL cells [140].

These results suggest the downregulation of p53 target genes may contribute to T cell transformation in Dnmt3a-deficient mice. Altogether, these data indicate that downregulation of p53 is an important event during lymphomagenesis in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ mice.

Promoter hypomethylation and p53 downregulation may not be the only relevant events involved in the development of PTCL in Dnmt3a-defficient mice. An additional DNA methylation change that could contribute to the development of PTCL in *Dnmt3a*$^{\Delta/\Delta}$ mice is promoter hypermethylation, as it has been linked to the inactivation of tumor suppressor genes [141, 142]. Although such changes would not be linked to Dnmt3a directly as inactivation of this enzyme is an initiating event of tumorigenesis, promoter hypermethylation mediated by other DNA methyltransferase and subsequent gene silencing could still drive tumorigenesis. In particular, in our previous studies we observed upregulation of Dnmt3b in Dnmt3a-defficient MYC-induced T cell lymphomas, suggesting that such an event may result in aberrant *de novo* methylation [92]. Surprisingly, despite identification of 50 genes whose promoters are hypermethylated in PTCL relative to CD8+ T cell controls, only *Fhit, CD226,* and *Emp1* were underexpressed. This raises a possibility that silencing of these genes contributes to PTCL development. *Fhit* is a predicted tumor suppressor gene that is frequently deleted in B cell malignancies, including Burkitt's lymphoma and primary effusion lymphoma [143, 144]. Furthermore, *in vivo* studied using *Fhit*$^{+/-}$ mice found that loss of a single allele of *Fhit* increased susceptible to carcinogen-induced tumor development in the esophagus and forestomach, further demonstrating the role of Fhit as a tumor suppressor [144]. CD226 is expressed on different hematopoietic cells including CD8+ T cells and contributes to their activation, expansion and differentiation but its deficiency in mice did not induce lymphomas, suggesting that this gene may not be a tumor suppressor gene [145]. Similarly, *Emp1*

overexpression correlated with enhanced cell proliferation and poor prognosis in B cell precursor ALL leukemia, suggesting an oncogenic function of this gene at least in some hematologic malignancies [146]. However, the possible role of *Fhit, CD226,* and *Emp1* as tumor suppressors in CD8+ *Dnmt3a*$^{\Delta/\Delta}$ PTCL is unclear. Thus, the role of hypermethylation and silencing in disease development and progression in mouse *Dnmt3a*$^{\Delta/\Delta}$ PTCL will require further investigation.

One of the interesting findings presented here is the exclusive sensitivity of CD8+ T cell to transformation in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ mice. This is not a consequence of impaired T cell development as we previously reported that loss of Dnmt3a does not affect the development of hematopoietic lineages [79]. Therefore, the reason as to why CD8+ but never CD4+ or CD4+CD8+ T cells become transformed in the absence of Dnmt3a is unclear at present. We speculate that the epigenome of CD8+ T cells is more dependent on Dnmt3a than other T cell types or CD8+ T cells may acquire genetic alterations that collaborate with epimutations more readily than other T cells. Of note, a differential sensitivity of T cell subtypes to transformation has been observed in response to infection by HTLV-1, which predominantly transforms CD4+ T cells, while HTLV-2 mainly transforms CD8+ T cells [147, 148]. Further studies will have to clarify whether the methylome of CD4 cells is more resistant to the lack of Dnmt3a as well as the nature of events responsible for CD8+ T cell transformation.

Another interesting finding from our study is association of transcription factor (TF) binding motifs with regions hypomethylated and hypermethylated in Dnmt3a-defficient PTCL. Analysis of TF binding sites found motifs for three TFs - AML1, NF-κB, and OCT1 – that were enriched in hypomethylated DMRS, suggesting their potential role in maintenance methylation mediated by Dnmt3a. In such a scenario, interaction of these factors with Dnmt3a may determine which specific loci Dnmt3a is targeted to. Interestingly, the p50 subunit of the NF-κB transcription factor was reported to interact with Dnmt3a in a glioblastoma cell line [149]. Similarly, we also

observed association of binding sites for Ap-2rep, SOX5, and myogenin with hypermethylated DMRS. Whether any of these transcription factors play role in aberrant promoter hypo- or hypermethylation remains to be determines.

Altogether, our data identify Dnmt3a as a critical tumor suppressor gene in the prevention of B- and T cell malignancies and link decreased Dnmt3a levels to decrease in p53, which may functionally contribute to the development of CD8+ PTCL. These data along with its documented role in prevention of myeloid malignancies defines Dnmt3a as a protector of the methylome critical for safeguarding normal hematopoiesis.

CHAPTER 4:

LOSS OF DNMT3A INDUCES CLL AND PTCL WITH DISTINCT METHYLOMES AND TRANSCRIPTOMES IN MICE.

The material presented in this chapter was published previously [150]:

Haney SL*, Upchurch GM*, Opavska J, Klinkebiel D, Appiah AK, Smith LM, Heavican TB, Iqbal J, Joshi S, Opavsky R. Loss of Dnmt3a induces CLL and PTCL with distinct methylomes and transcriptomes in mice. Scientific Reports. 2016 Sep 28;6:34222.

(* denotes equal contribution)

† Note: In the subsequent chapter, "data not shown" refers to data that is not presented here but can be located in online supplementary data files of the above publication.

Introduction

Cytosine methylation of DNA is an epigenetic modification affecting gene transcription and the integrity of the mammalian genome. Basic methylation patterns are established and maintained by catalytic activity of three DNA methyltransferases: DNMT1, DNMT3A, and DNMT3B. Promoter methylation is associated with transcriptional repression and plays a role in a variety of normal physiologic processes, including X-chromosome inactivation, genomic imprinting, differentiation and hematopoiesis [82, 83].

Non-Hodgkin's lymphoma (NHL) is a heterogeneous group of lymphoid malignancies that arise from transformation of B, T, and NK cells. The majority of NHLs are B-cell lymphomas, the most common of which is chronic lymphocytic leukemia/small lymphocytic lymphoma (CLL/SLL): an indolent low-grade lymphoproliferation of mature B-cells [58]. However, T cell lymphomas develop in ~10% of NHL patients [151]. Approximately 30% of these T cell malignancies will be diagnosed as peripheral T cell lymphoma-not otherwise specified (PTCL-NOS): a group of high-grade mature T cell neoplasms not classified by other WHO criteria [152]. Both CLL and PTCL are life-threatening conditions that present in late adulthood and despite recent advances in chemotherapy, these diseases remain refractory to cure. A better understanding of deregulated molecular landscapes and a contribution of individual changes to the development of these two NHLs is needed to generate new therapeutic approaches.

Two types of molecular changes likely involved in the pathogenesis of CLL and PTCL are genetic alterations and epimutations such as de-regulated cytosine methylation. It was reported that an average of 45 somatic mutations are present in human CLL samples, with most genes mutated in less than 5% of cases [153]. About one third of cases did not have recurrent mutations, suggesting a high degree of heterogeneity and no clear genetic drivers of CLL. Interestingly, global gene expression profiling identified two DNA methyltransferases, Dnmt3b and Dnmt3a, as the top 1% of underexpressed genes in human CLL [154], suggesting that the DNA methylation landscape

may be deregulated. Indeed, DNA methylation profiling revealed a substantial genome wide promoter and gene-body hypomethylation in tumors relative to normal B cells [67]. Consistently with possible roles of Dnmt3a and Dnmt3b in CLL development we have previously reported that conditional inactivation of Dnmt3a in hematopoietic stem cells and progenitors in *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26 LOXP* $^{EGFP/EGFP}$ (*Dnmt3a*$^{\Delta/\Delta}$) mice resulted in the development of chronic lymphocytic leukemia (CLL) around 1 year of age and is accelerated when Dnmt3b is deleted as well [79]. Such data strongly suggest that Dnmt3a represses genes in normal B cells likely through promoter methylation whose up-regulation upon hypomethylation may contribute to the development of CLL.

The mutational landscape of TCL appears to be less diverse than in CLL. Interestingly, one of the most frequently mutated gene is DNMT3A, suggesting a possible involvement in disease development [42]. Although no comparable large scale profiling of the methylation landscape has been performed on TCL to date, our functional studies utilizing *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26 LOXP* $^{EGFP/EGFP}$ mice demonstrated that Dnmt3a likely play a role in pathogenesis of PTCL. Although these mice primarily develop CLL, a 30% of mice develop CD8-positive PTCL either in combination with CLL or by itself [79, 95]. However, the nature of deregulated events in both mouse diseases and how they relate to human diseases remains poorly understood.

To better understand the molecular changes occurring in Dnmt3a-deficient mice we performed global methylation profiling using whole genome bisulfite sequencing (WGBS) and gene expression profiling using RNA-seq on CLL and PTCL tumors isolated from *Dnmt3a*$^{\Delta/\Delta}$ mice, as well as control B-1a and CD8+ T cells. This analysis revealed that while normal B-1a and CD8+ T cells had remarkably similar methylomes and transcriptomes, Dnmt3a loss induced unique changes in DNA methylation and gene transcription in CLL and PTCL. Importantly, analysis of available expression data from human CLL and PTCL samples, revealed a significant overlap

between human and mouse transcriptomes and methylomes, suggesting *Dnmt3a*$^{\Delta/\Delta}$ mice may serve as a useful tool to identify oncogenic drivers contributing to CLL and PTCL pathogenesis in humans.

Material and Methods

Mouse Studies. *Dnmt3a* $^{\Delta/\Delta}$ (*EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$) and *Dnmt3a* $^{+/+}$ (*EμSRα-tTA;Teto-Cre;Dnmt3a*$^{+/+}$*;Rosa26LOXP*$^{EGFP/EGFP}$) mice were generated as previously described [79, 92].

FACS. FACS analysis was performed as described previously [95]. Briefly, B-1a cells were identified as being positive for B220, CD19, and CD5 cell surface markers, while CD8+ T cell were positive for CD8 and CD3 cell surface markers. Mice diagnosed with MBL had 2-20% B-1a cells in blood, mice diagnosed with CLL had >20% B-1a cells in blood, mice diagnosed with PTCL had >25% CD3+CD8+ cells in the spleen. B cell subsets shown in Figure 5c were FACS sorted from the spleen and bone marrow of 8-week old FVB/N mice using the following markers: splenic B-1a (B220+, CD19+, CD5+), splenic marginal zone B cells (CD23+, CD21-, IgM-low), splenic follicular B cells (CD23-, CD21-, IgM-high), bone marrow derived immature B cells (B220+, IgM+, IgD-) and bone barrow derived mature B cells (B220+, IgM+, IgD+). Analysis was performed at the UNMC Flow Cytometry Facility.

Western Blot. Western blots were performed as previously described [93] with use of the following antibodies: Dnmt3b (52A1018, Imgenex), Dnmt1 (H-300, Santa Cruz), and HDAC1 (ab7028, Abcam).

Combined Bisulfite Restriction Analysis (COBRA). COBRA analysis was carried out as described previously [93, 96]. Primers used in this study are as follows:

Sgk3:	TATTTTTTGTTTATAAATGTTGGTGG	(F)
	AAACACTAAAAAAATAACATCATTCTC	(R)
Nfam1:	GGTTGAGTAAGAGTAAAGAAGATAGGTTAT	(F)
	AACAAAAAAACAAACCCAAAAATAT	(R)

Pstpip2:	TGTATGTTTTTGGAGAGTAATAGGGT	(F)
	TCCTTCCTAACAAAACCAATATCTT	(R)
Plscr1:	AGGAAATATAATAGAGGTGAAGATAGTAGG	(F)
	CCATAACACTCCAATACTAAAAAAAA	(R)
Crtam:	AAGAGTTTTTATTGGGTTTTTATTTTT	(F)
	ACTCACAAAAATCCACTTCCTAAATAC	(R)
Sgk1:	TGTTTTTATGGGTATTAAATGGTATA	(F)
	TACAACCTAAATTTACTATCAAAATTACTT	(R)
Samd3:	TTGGAAGTTATTTTTGATATAGATTTTGT	(F)
	TCACTTTCCAAAAAAACATAAACAC	(R)
Oas3:	TGGGGTTGTATAAAGGATTAAGGTA	(F)
	ACACACACATACACAAAAATACAAACA	(R)
Fam169b:	TTGGAAATTGAATTTTAATTTATTAGGAT	(F)
	AAAAAAACCACTCTTCAACCTATCA	(R)
Pvt1:	GGAGTTTTAAGTGGGATTTTTTAAA	(F)
	CATACAATCACTCCCTAACAAAATAAA	(R)
Dbi:	GTAAAGGTAGGGTTAGGGTTGTTGT	(F)
	TTCCTCTTTCAAAAATAACAAAAAAAA	(R)
Il5ra:	GGAGAATTTATGTTTTTTTAGAGTGTT	(F)
	ATCATCCCATTAATCATTTATATTTTTATA	(R)
Zbtb32:	AAAGATTAGGTGTGTGTGATTTAGATTT	(F)
	AAATTTCAAAAACTTTTTATCCCTTC	(R)

Whole genome Bisulfite sequencing (WGBS). Splenic B-1a cells (EGFP+CD5+CD19+B220+) were FACS sorted from Dnmt3a$^{+/+}$ control (N=1), and a *Dnmt3a*$^{\Delta/\Delta}$ mouse with CLL (N=1). Splenic CD8+ T cells (EGFP+CD8+CD3+) were FACS sorted from *Dnmt3a*$^{+/+}$ control (N=1), and a *Dnmt3a*$^{\Delta/\Delta}$ mouse with PTCL (N=1). Genomic DNA was isolated using standard protocols. The WGBS was performed in DNA Services facility at the University of Illinois at Urbana-Champaign, Roy J. Carver Biotechnology Center using two lanes for each sample on the Illumina HiSeq2500 sequencer with paired-end 160bp reads. Each lane produced over 310 million reads. Sequence tags were aligned with the mouse genome (Dec. 2011 mus musculus assembly mm10, Build 38) using the methylated sequence aligner Bismark [114] by the University of Nebraska Epigenomics Core facility. The resulting data file contains the percent methylation at each CpG measured. Each individual CpG was retained and percent methylation determined only if it was represented by ≥ 5 individual sequences. Correlation based, average linkage hierarchical clustering of genome location matching CpG methylation percentages per sample was performed using the R software package RnBeads [115]. Genome location matching differentially methylated cytosine (DMCs) and differentially methylated regions (DMRs) were determined using the R software package DSS [116]. DMCs were determined by first smoothing the raw percent methylation values based on a moving average algorithm and smoothing span of 500 bases. DMRs were then determined based on average DMC methylation change of 30% or greater, at least 50% or greater individual DMC P-values less than 0.05, minimum base pair length of 100, minimum of three DMCs represented, and the resulting DMRs were averaged if they were closer than 50 bases. DMRs were aligned with the mouse genomic repeats. Genomic repeats were acquired from the UCSC Genome table browser based on the RepeatMasker program (http://www.repeatmasker.org). The repeat was retained if the overlap between the DMR and repeat was more than 25 percent of the length of the repeat. WGBS data is available for download through the NCBI Gene Expression Omnibus (GSE78146).

RNA-seq. RNA was isolated from FACS sorted splenic B-1a cells (EGFP+CD5+CD19+B220+) from Dnmt3a$^{+/+}$ control (N=2), and *Dnmt3a*$^{\Delta/\Delta}$ CLL (N=8). Splenic CD8+ T cells (EGFP+CD8+CD3+) were FACS sorted from *Dnmt3a*$^{+/+}$ control (N=2) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (N=3). RNA was isolated as previously described [93]. Library generation was performed using the TruSeq mRNA kit. The resulting libraries were sequenced on the Illumina HiSeq 2000 platform using paired-end 100bp runs (SeqMatic, Fremont, CA). The resulting sequencing data was first aligned using TopHat (version 1.0.0) and mapped to the Mus musculus UCSC mm10 reference genome using the TopHat 2 aligner. Cufflinks 2 was used to estimate FPKM of known transcripts, perform de novo assembly of novel transcripts, and calculate differential expression. For differentially expressed genes, we considered those genes with a fold change ≥ 2 and a q-value < 0.05 to be significant. RNA-seq data is available for download through the NCBI Gene Expression Omnibus (GSE78146). RNA-seq data obtained from 5 normal human B cell samples and 15 human CLL tumors was downloaded from the Gene Expression Omnibus (GSE70830) and used to identify differentially expressed genes in human CLL ($q<0.05$ was considered significant, as determined by Cufflinks).

Microarray. Microarray data from 5 normal Tonsil T cells (GSE65135) were downloaded from NCBI Gene Expression Omnibus and compared to 15 cytotoxic PTCL samples [42]. Datasets were generated with Affymetrix U133 plus 2 arrays and analyzed using Affymetrix Expression Console and Transcriptome Analysis Console (v3.0). Data was analyzed using a one-way between-subject ANOVA to generate P-values and identify differentially expressed genes (P-value < 0.05 and fold change >1.5).

Statistical Analysis. Comparison of overlap between mouse and human datasets was performed using a statistical simulation in which we randomly selected samples from the human genes list and mouse gene list without replacement. The number of times the sample human and mice genes overlapped was computed. This process ran for 100,000 times and for runs where the

number of times the genes overlapped was greater or equal to the number observed it was counted. The P-value was computed by dividing the number of overlaps greater than that observed by 100,000. Continuous variables were compared using 2-sample Student's *t* tests; results are presented as mean with error bars representing standard deviation. The Kaplan-Meier method was used to estimate disease-free survival distribution. All other statistical methods are described in the relevant section of the material and methods.

Quantitative Real-Time qRT-PCR. qRT-PCR was performed as previously described [93]. Primer sequences used in experiments presented here are as follows:

Human Primers:

STAT1: ACAACCTGCTCCCCATGTCT (F), GCAGGAGGGAATCACAGATG (R)

JDP2: CCGGGAGAAGAACAAAGTCG (F), CGGTTCAGCATCAGGATGAG (R)

ZBTB32: GGAGACGCACTACCGAGTCC (F), GAAGGTGGAGCGGATGGTC (R)

ZBTB38: TGGAGGACTCAGAACCAAGG (F), CCGTGTCACTGTGAAAGTCG (R)

Mouse Primers:

Stat1: GGCCCCGAATTTGACAGTAT (F), ACCAGCAGTGCTCAGCAAAT (R)

Jdp2: GCCGGGAAAAGAACAAAGTC (F), GCGGTTGAGCATCAGGATAA (R)

Zbtb32: TCAGCCCTTGGCAGATAGAA (F), CACAGAGGGCATCGATAGGG (R)

Zbtb38: CGAGTGATTTCTGCCAAAGC (F), CACATAAGATGCCCCGAATG (R)

Study approval. This study was performed in accordance with the guidelines established by the Guide for the Care and Use of Laboratory Animals at the National Institutes of Health. All experiments involving mice were approved by the IACUC (Protocol number: 08-083-10-FC) at the University of Nebraska Medical Center. The Institutional Review Board of UNMC approved the

human studies, and patients provided informed consent prior to the study in accordance with the Declaration of Helsinki.

Results

Cell-Autonomous Tumor Suppressor Function of Dnmt3a in the prevention of CLL.

To determine the role of Dnmt3a in hematopoiesis we used the *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ (*Dnmt3a*$^{\Delta/\Delta}$) mouse model to conditionally delete Dnmt3a in hematopoietic stem and progenitor cells (HSPCs) cells. In this system, deletion of Dnmt3a occurs in EGFP-positive cells (30-50% of all HSPCs) whereas EGFP-negative cells (50-70% of all HSPCs) retain both conditional Dnmt3a alleles (*Dnmt3a*$^{fl/fl}$) and therefore behave like *wild-type* cells [79]. Using this model, we previously showed that a long-term *Dnmt3a*-deficiency resulted in the development of a chronic lymphocytic leukemia (CLL) in 68% of mice, CD8-positive mature T cell lymphomas (PTCL) in 14% mice and mixed CLL/PTCL in 18% cases within one year of age [79, 95]. **(Fig. 19a;** and data not shown).

To examine if the levels of other DNA methyltransferases were changed upon deletion of Dnmt3a, we analyzed protein levels of Dnmt1 and Dnmt3b in Dnmt3a-deficient PTCL and CLL samples. While we did not observe any consistent changes in the levels of Dnmt1 in either tumor type, we did see down regulation of Dnmt3b in PTCL, but not CLL samples, relative to their controls (wild-type thymocytes and splenocytes, respectively) (**Fig. 19b, 20**). Such molecular change may contribute to the pathogenesis of PTCL in particular because Dnmt3b loss accelerates MYC-induced lymphomagenesis as well as promotes the development of T cell lymphomas in *Dnmt3a*$^{\Delta/\Delta}$ mice [79, 93].

To determine whether loss of Dnmt3a promotes tumorigenesis in a cell-autonomous way, we used adoptive transfer to introduce bone marrow cells isolated from 6 weeks old *Dnmt3a*$^{+/+}$ control and *Dnmt3a*$^{\Delta/\Delta}$ mice into lethally irradiated FVB recipient mice. Analysis of hematopoiesis in recipient mice revealed that both EGFP-negative and EGFP-positive cells contributed to reconstitution of hematopoietic system with no signs of disease 2 months after adoptive transfer

(**Fig. 19c, 21** and data not shown). Injected mice were observed for signs of tumorigenesis and harvested when terminally sick. Control mice remained disease free during the 500 days observational period (**Fig. 19d**). In contrast, 50% of *Dnmt3a*$^{\Delta/\Delta}$ mice (4 of 8) became terminally ill with CLL with a median survival of 311 days (**Fig. 19d-g**). The remaining 50% of *Dnmt3a*$^{\Delta/\Delta}$ mice developed monoclonal B cell lymphopoiesis (MBL), a less advanced form of CLL, in which the percentage of B-1a cells in the blood ranges from 2-20% (**Fig. 19d-g).** EGFP-positive CLL or MBL splenic cells induced disease in sub-lethally irradiated secondary FVB recipients, suggesting that tumors contain leukemia-initiating cells (data not shown). Altogether, these results show that the Dnmt3a's tumor suppressor function in prevention of CLL resides within the hematopoietic compartment and is likely cell-autonomous. Due to the lower frequency of PTCL development in Dnmt3a-deficient mice, larger cohorts of mice will need to be analyzed to determine whether Dnmt3a tumor suppressor function in prevention of PTCL is also autonomous to hematopoietic system.

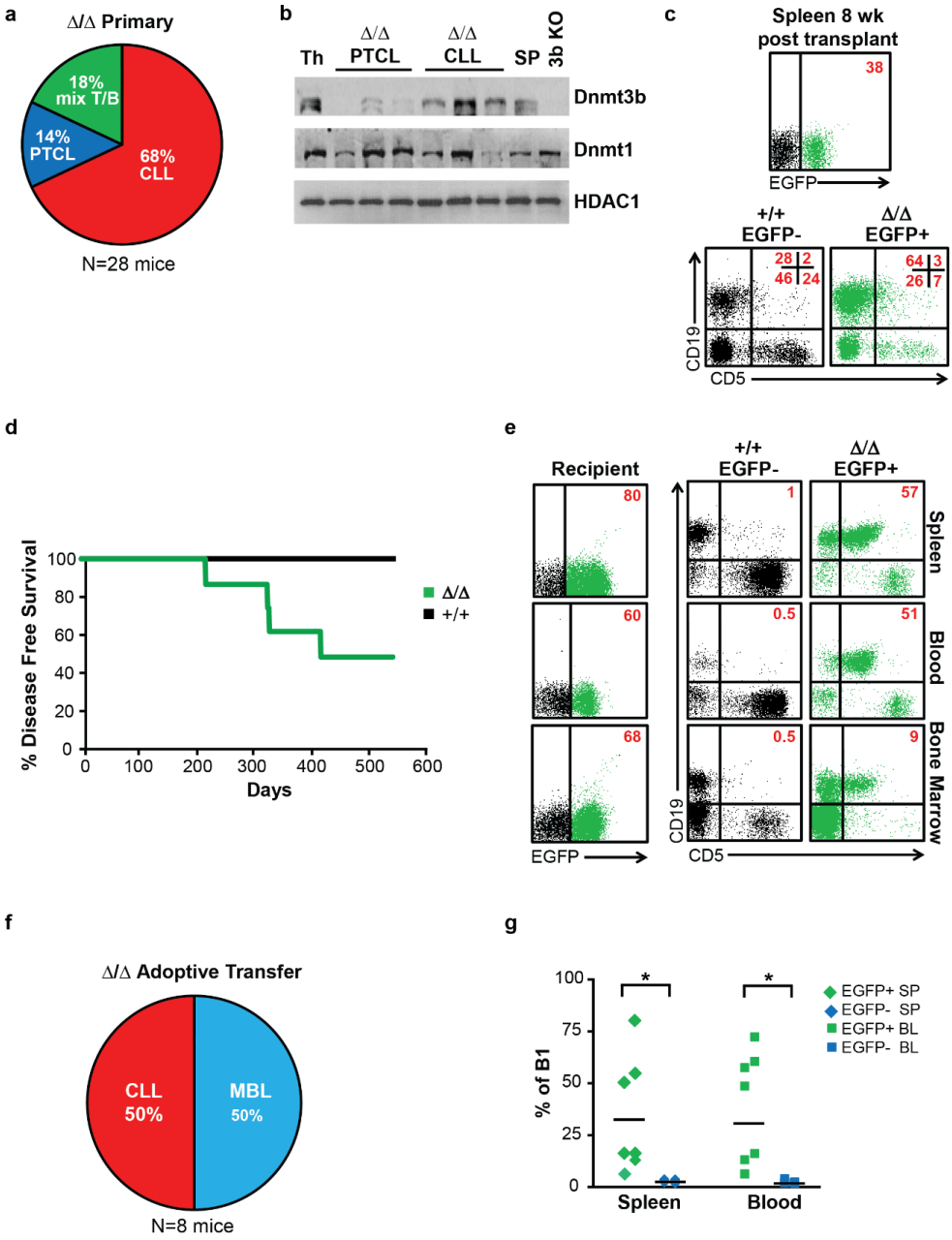


Figure 19. Dnmt3a's tumor suppressor function is cell autonomous. (a) Percentage of *Dnmt3a*$^{\Delta/\Delta}$ mice diagnosed with CLL (red), PTCL (blue), or mixed CLL/PTCL (green) at time of harvest, as determined FACS. N=28. **(b)** Immunoblot analysis of Dnmt3b and Dnmt1 proteins in *Dnmt3a*$^{+/+}$ normal thymus (Th) and spleen (SP), *Dnmt3a*$^{\Delta/\Delta}$ PTCL, and *Dnmt3a*$^{\Delta/\Delta}$ CLL samples. *Dnmt3b*$^{-/-}$ (3b KO) cells were used as a negative control. HDAC1 is shown as a loading control. **(c)** Representative FACS diagram showing CD19 and CD5 expression in EGFP- (black) and EGFP+ (green) cellular populations isolated from the spleen of a lethally irradiated FVB recipient mouse injected with bone marrow *Dnmt3a*$^{\Delta/\Delta}$ bone marrow. The mouse was harvested 8 weeks post injection. Percentage of cells in each quadrant are shown in the top right in red. **(d)** Kaplan-Meier survival curves for FVB mice lethally irradiated and injected with *Dnmt3a*$^{+/+}$ or *Dnmt3a*$^{\Delta/\Delta}$ (N=8) bone marrow cells. **(e)** Representative FACS diagram showing CD19 and CD5 expression in EGFP- (black) and EGFP+ (green) cell isolated from the spleen, blood and bone marrow of a lethally irradiated FVB recipient mice injected with *Dnmt3a*$^{\Delta/\Delta}$ bone marrow. The mouse was harvested 9 months post injection when terminally ill. Percentage of cells staining positive in each quadrant are shown in the top right in red. **(f)** Percentage of lethally irradiated wild-type mice injected with *Dnmt3a*$^{\Delta/\Delta}$ bone marrow cell that were diagnosed with CLL (red) or MBL (blue) at time of harvest, as determined FACS. N=8. **(g)** Percentage of EGFP+ (green) and EGFP- (blue) B-1a cells (B220+CD19+CD5+, determined by FACS) in the spleens and blood of lethally irradiated FVB recipient mice injected with *Dnmt3a*$^{\Delta/\Delta}$ bone marrow. P<0.05 is indicated by (*), two-tailed Student's *t*-test.

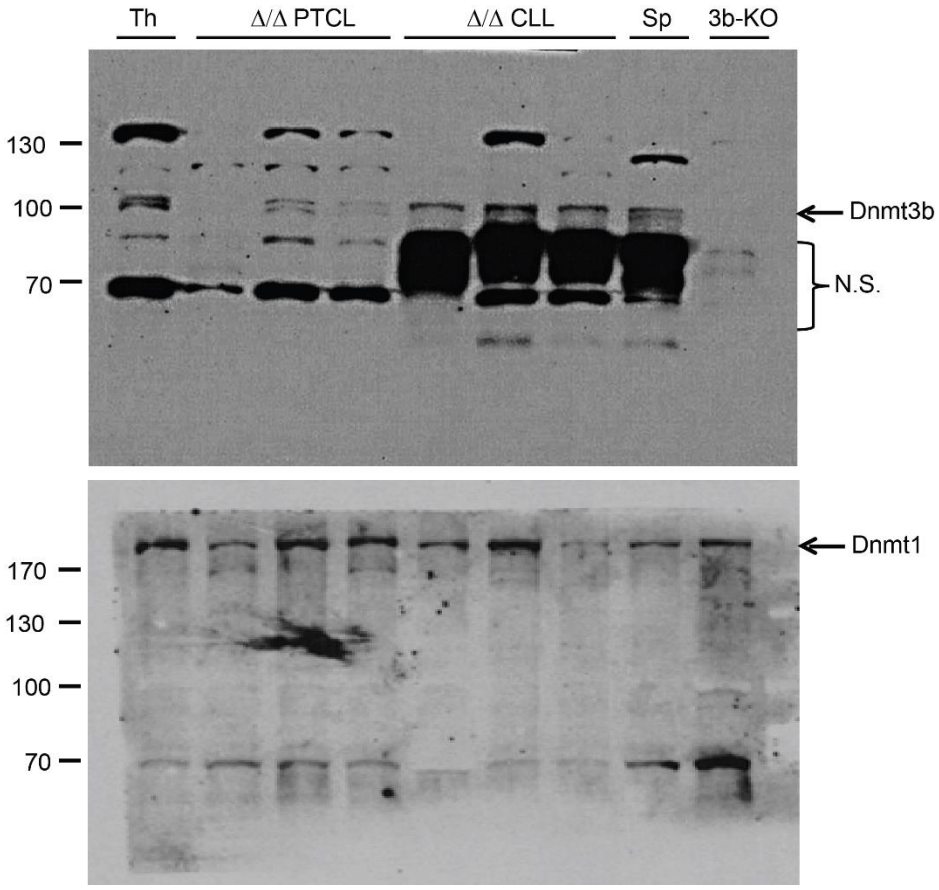


Figure 20. Full length immunoblots as presented in Figure 19b. Immunoblot analysis of Dnmt3b and Dnmt1 proteins in *Dnmt3a*$^{+/+}$ normal thymus (Th), *Dnmt3a*$^{+/+}$ normal spleen (Sp), *Dnmt3a*$^{\Delta/\Delta}$ PTCL, and *Dnmt3a*$^{\Delta/\Delta}$ CLL samples. *Dnmt3b*$^{-/-}$ (3b KO) T cell lymphoma line was used as a negative control for Dnmt3b immunoblot. The location of Dnmt3b and Dnmt1 proteins are marked by arrows. Non-specific bands (N.S.) present in the Dnmt3b KO control are labelled.

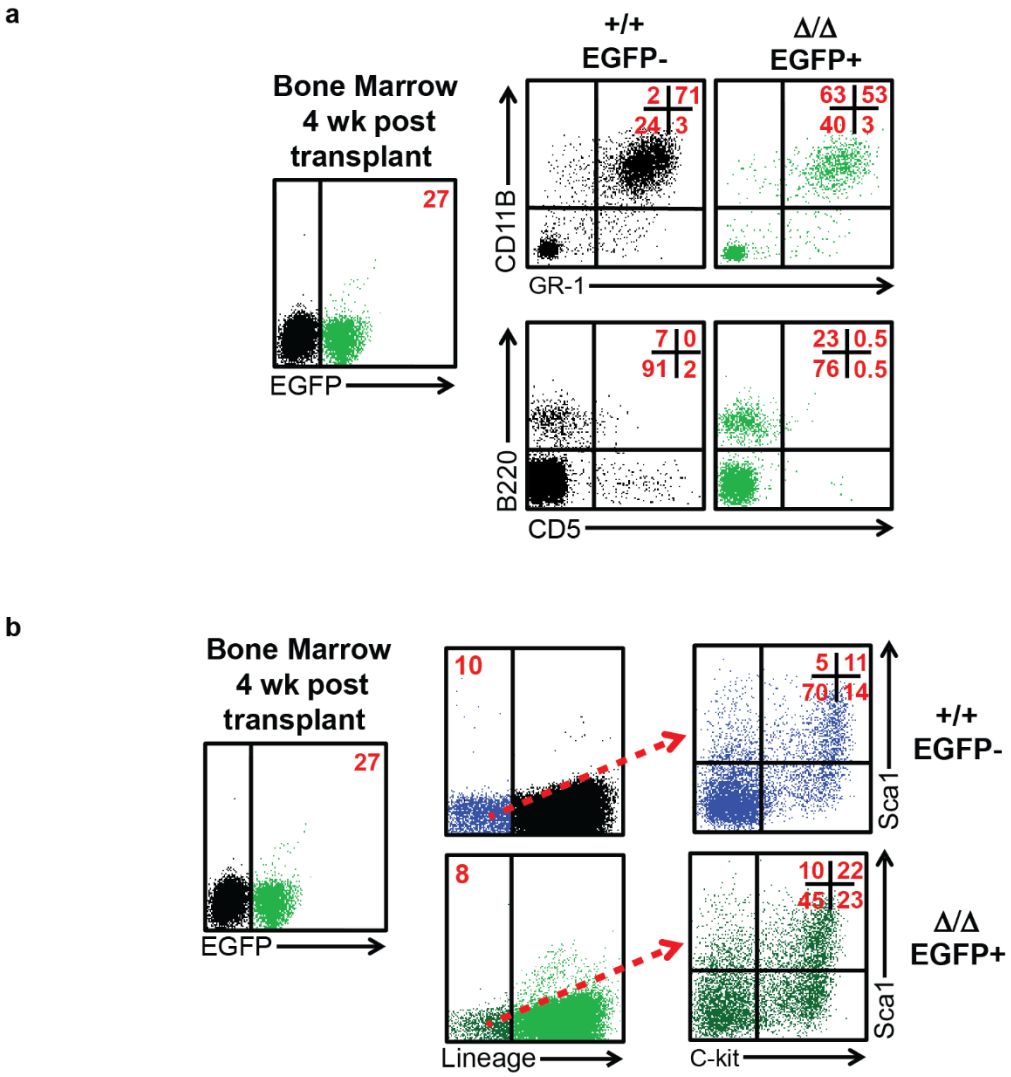


Figure 21. Dnmt3a's tumor suppressor function is cell autonomous to the hematopoietic system. (a) FACS diagram showing CD11b and GR-1 expression (top) and B220 and CD5 expression (bottom) in EGFP- (black) and EGFP+ (green) cells isolated from the bone marrow of a lethally irradiated FVB recipient mice injected with *Dnmt3a*$^{\Delta/\Delta}$ bone marrow. The mouse was harvested 4 weeks post injection. Percentage of cells staining positive in each quadrant are shown in red. **(b)** FACS diagram showing percentage of lineage-Sca1+ckit- cells in EGFP- (blue) and EGFP+ (green) cells isolated from the bone marrow of a lethally irradiated FVB recipient mice injected with *Dnmt3a*$^{\square/\square}$ bone marrow. The mouse was harvested 4 weeks post injection. Percentage of cells staining positive in each quadrant are shown in red.

DNA methylome and transcriptome is highly similar between normal mouse B-1a and CD8 cells.

To determine the nature of deregulated molecular events during CLL and PTCL development in *Dnmt3a*$^{\Delta/\Delta}$ mice we performed global methylation analysis by Whole Genome Bisulfite Sequencing (WGBS) and gene expression profiling by RNA-seq on B220+CD19+CD5+ (B-1a cells) and CD8-positive T cells isolated from *Dnmt3a*$^{+/+}$ spleens, as these cellular populations are immunophenotypically the closest normal counterparts of CLL and PTCL cells, respectively. Analysis of methylation in 15,533,510 matched CpG dinucleotides distributed across the genome revealed a high degree of methylation in both cell types, with higher levels of methylation in B-1a cells. For example, more CpG dinucleotides were methylated at levels ≥ 76% in B- than in T-cells (79% and 75% in B-1a and CD8+ T cells, respectively) **(Fig. 22a**). In contrast, fewer CpG dinucleotides were methylated at low levels (≤ 25%) in B- than T- cells (5.5% and 6% in B-1a and CD8+ T cells, respectively). Increased methylation in B-1a cells relative to CD8+ T cells was also detected in core promoters. More promoters were methylated at levels ≥ 76% in B- than T-cells (48% and 44% in B-1a and CD8+ T cells, respectively) (**Fig. 22b**). A smaller number of promoters was methylated at low levels (≤ 25%) in B- than in T-cells (28% and 29% in B-1a and CD8+ T cells, respectively) (**Fig. 22b**). Methylation of 21,712 core promoters (-300 +150bp relative to TSS) was remarkably similar on a locus specific level between these two normal cell types, although some promoters were differentially methylated in B-1a B cells and CD8+ T cells (**Fig. 22c, Fig. 23,** and data not shown). A combined gene expression and methylation analysis revealed that the majority of genes with low levels of promoter methylation were expressed, whereas genes with high levels of promoter methylation were largely repressed (**Fig. 22d, Fig. 24,** and data not shown). Promoter methylation was well conserved between B-1a and CD8+ cells, with 90% of hypomethylated and 83% of hypermethylated promoters common between B-1a and CLL (**Fig. 22e**). Similarly, more than 90% of highly and lowly expressed genes were shared between B-1a

and CD8+ cells (**Fig. 22e**). Ingenuity pathway analysis (IPA) of highly expressed genes in B-1a and CD8 cells revealed the same top subcategories of genes significantly associated with *organismal survival*, *hematologic system*, *tissue morphology*, *hematopoiesis*, *lymphoid tissue structure* (**Fig. 22f, Fig. 25**), highlighting a large overlap in genes commonly expressed in B1a and CD8 cells, as well as their link to hematopoietic system. However, IPA analysis of genes only overexpressed in individual cell types (503 and 289 genes in B-1a and CD8, respectively) revealed specific differences. For example, genes associated with categories such as *proliferation of B lymphocytes*, *quantity of B lymphocytes* or *quantity of IgG* were only detected using data obtained from B-1a cells (**Fig. 26**). Similarly, categories such as *quantity of CD8+ T lymphocyte*, *activation of T lymphocytes* or *morphology of thymus gland* were only detected using data obtained from CD8 cells (**Fig. 26**). Altogether, these data show large scale hypermethylation in both normal cell types, with B-1a cell showing higher levels of methylation. In addition, methylation and gene expression in normal splenic B-1a cells and CD8 cells is largely shared with the exception of a subset of genes specifically involved in processes associated with unique functions of these cell types.

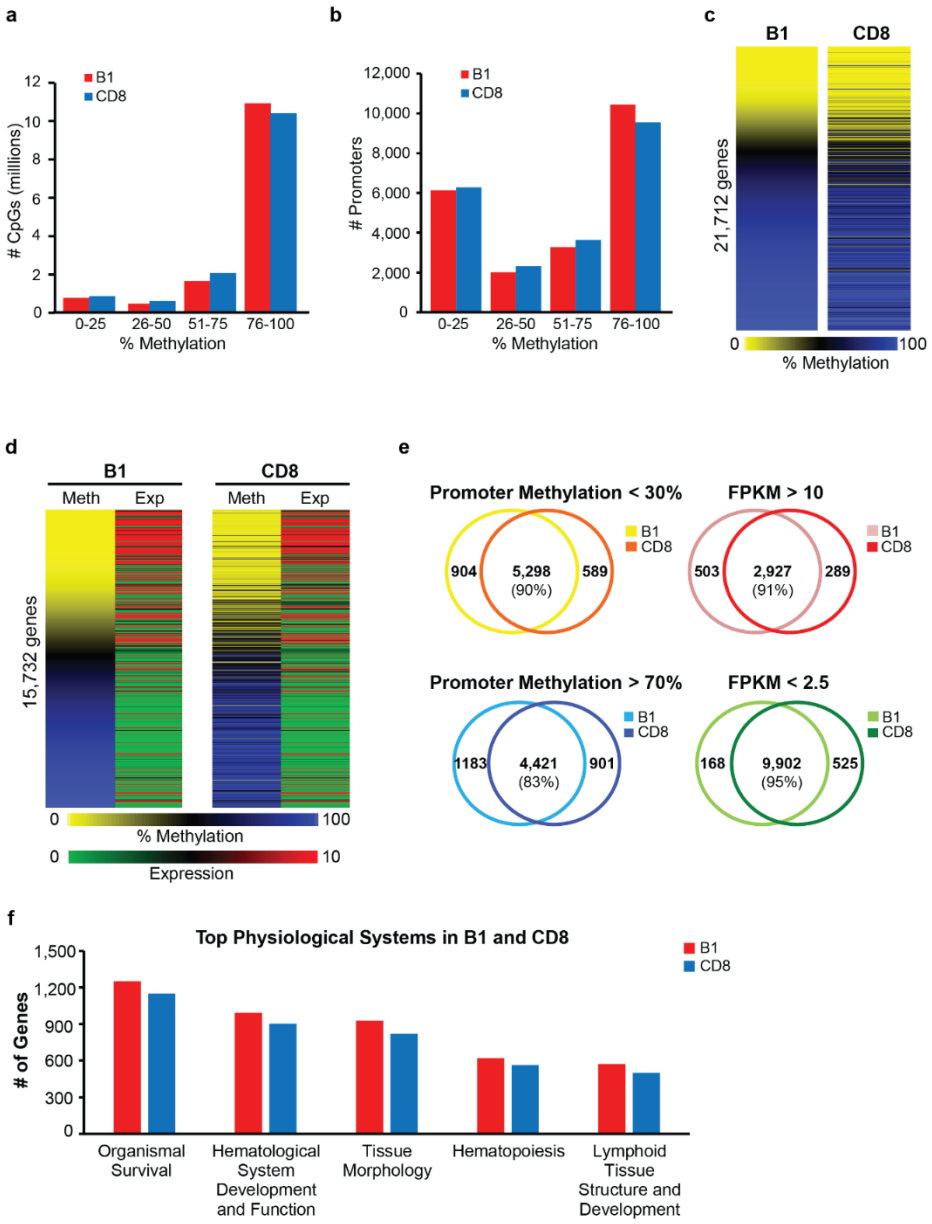


Figure 22. Transcriptome and methylome of normal B-1a and CD8 cells. (a) Breakdown of CpG methylation as determined by WGBS in B-1a (B1) and CD8 samples. Individual CpGs were placed into 4 categories based on percent methylation (0-25%, 26-50%, 51-75%, and 76-100%). **(b)** Breakdown of promoter methylation for 21,712 genes in B1 and CD8 samples. Methylation percentage for individual CpGs were annotated to the core promoter regions (-300bp to +150bp relative to the TSS). Methylation percentages for all CpGs across the 450bp region were averaged to give a mean methylation value for each gene promoter. Promoters were placed in 4 categories based on percent methylation (0-25%, 26-50%, 51-75%, and 76-100%). **(c)** Methylation status of 21,712 promoters in B1 and CD8 samples as determined by WGBS. Mean promoter methylation was determined as described in part b of the figure legend. Hypomethylation is shown in yellow and hypermethylation in blue. **(d)** Heat map presentation of promoter methylation (analyzed as in Fig. 22b) and corresponding gene expression (presented as average FPKM values as determined by RNA-seq) in mouse splenic B1 and CD8 cells for 15,732 genes. Genes with high FPKM values are shown in red and genes with low FPKM values are shown in green. Heat maps are organized in the same gene order to match data for methylation and gene expression. **(e)** (left) Venn diagram showing number of unique and overlapping hypomethylated (methylation <30%) and hypermethylated (methylation >70%) promoters in B1 and CD8 samples. (right) Venn diagram showing number of unique and overlapping highly expressed (FPKM >10) and lowly expressed (FPKM <2.5) genes in B1 and CD8 samples. **(f)** Ingenuity Pathway analysis (IPA) of highly expressed genes (FPKM ≥10) in B1 (red) and CD8 (blue) samples. The top subcategories obtained in Physiological System, Development and Functions are displayed ($P<0.05$, for all subcategories).

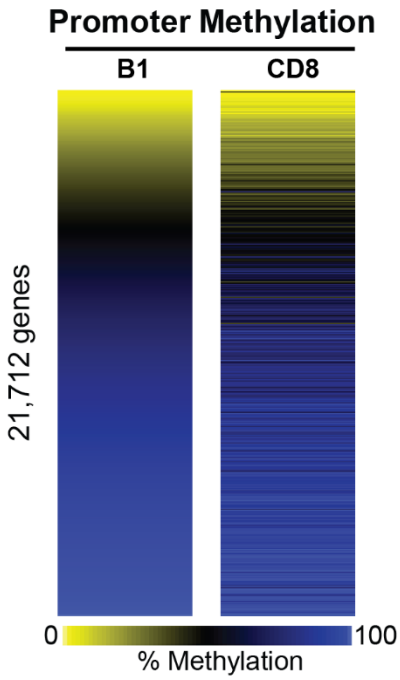


Figure 23. Methylome of long promoter regions in B-1a and CD8. Methylation status of 21,838 promoters in B-1a and CD8 samples as determined by WGBS. Methylation percentage for individual CpGs were annotated to the promoter regions -1,500bp to +500bp relative to the transcription start site. Methylation percentages for all CpGs across the 2000bp region were averaged to give a mean methylation value for each gene promoter. Hypomethylation is shown in yellow and hypermethylation in blue.

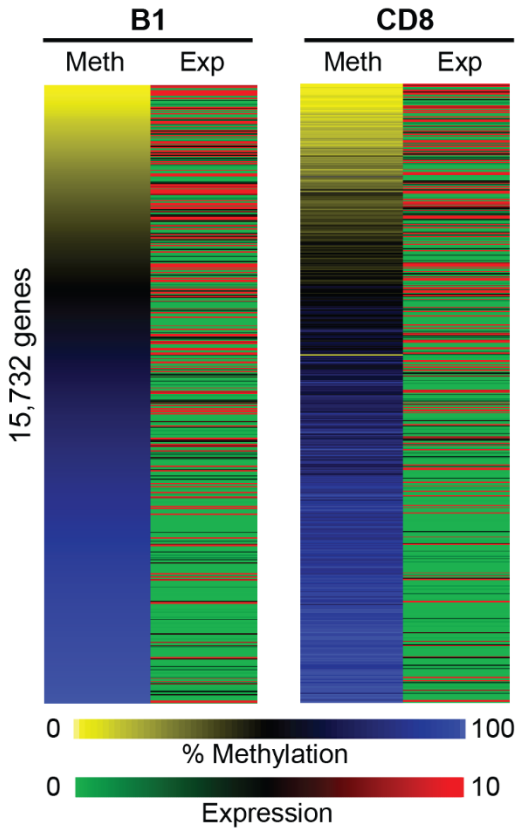


Figure 24. Methylome of long promoter regions with corresponding gene transcription in B-1a and CD8 cells. Heat map presentation of 2,000bp promoter methylation (analyzed as in Figure 23) and corresponding gene expression (presented as average FPKM values as determined by RNA-seq) in mouse splenic B1 and CD8 cells for 15,732 genes. Genes with high FPKM values are shown in red and genes with low FPKM values are shown in green. Heat maps are organized in the same gene order to match data for methylation and gene expression.

B-1 Highest Expressed Genes	
Physiological System Development and Function	**# Genes**
Hematological system development and function	994
Tissue Morphology	927
Organismal survival	1251
Hematopoiesis	619
Lymphoid tissue structure and development	571
Pathways	**# Genes**
EIF2 signaling	132
Protein Ubiquitination pathway	155
Glucocorticoid receptor signaling	155
Regulation of eIF4 anf p70S6K signaling	100
CD28 signaling in T helper cells	84
Diseases and Disorders	**# Genes**
Infectious disease	956
Organismal injury and abnormalities	1108
Inflammatory response	616
Immunological disease	1037
Cancer	1011

CD8 Highest Expressed Genes	
Physiological System Development and Function	**# Genes**
Organismal survival	1151
Hematological system development and function	903
Tissue Morphology	822
Hematopoiesis	563
Lymphoid tissue structure and development	498
Pathways	**# Genes**
EIF2 signaling	130
Protein Ubiquitination pathway	157
Regulation of eIF4 anf p70S6K signaling	96
Mitochondrial dysfunction	101
Glucocorticoid receptor signaling	141
Diseases and Disorders	**# Genes**
Infectious disease	842
Organismal injury and abnormalities	1525
Inflammatory response	526
Cancer	1427
Hematological disease	840

Figure 25. Summary of Ingenuity Pathway analysis (IPA) of all highly expressed genes (FPKM ≥ 10) in B-1a and CD8 control samples. P-values were less than 0.05 for all categories.

B-1 specific Genes	
Physiological System Development and Function	**# Genes**
Hematological system development and function	127
Tissue Morphology	107
Humoral immune response	75
Hematopoiesis	82
Lymphoid tissue structure and development	81
Pathways	**# Genes**
B Cell Receptor Signaling	22
PI3K Signaling in B Lymphocytes	17
Role of NFAT in Regulation of the Immune Response	17
Primary Immunodeficiency Signaling	9
FcRIIB Signaling in B Lymphocytes	8
Diseases or Functions Annotation	**# Genes**
Quantity of leukocytes	83
Proliferation of B lymphocytes	40
Quantity of mononuclear leukocytes	71
Quantity of lymphocytes	69
Quantity of immunoglobulin	42

CD8 Specific Genes	
Physiological System Development and Function	**# Genes**
Hematological system development and function	62
Immune cell trafficking	36
Digestive system development and function	8
Hepatic system development and function	3
Organ development	20
Pathways	**# Genes**
Role of IL-17A in Psoriasis	2
S-methyl-5'-thioadenosine Degradation II	1
Leukotriene Biosynthesis	2
Mismatch Repair in Eukaryotes	2
D-myo-inositol (1,4,5)-trisphosphate Degradation	2
Diseases or Functions Annotation	**# Genes**
Activation of leukocytes	24
Function of oval cells	3
Insulin-dependent diabetes mellitus	21
Systemic autoimmune syndrome	37
Cytolysis of lymphocytes	8

Figure 26. Summary of Ingenuity Pathway analysis (IPA) of those genes specifically expressed in either B-1a or CD8, but not the other. Genes with an FPKM ≥ 10 were used for analysis. P-values were less than 0.05 for all categories.

***Dnmt3a*$^{\Delta/\Delta}$ CLL and *Dnmt3a*$^{\Delta/\Delta}$ PTCL have distinct methylomes.**

To determine the effects of loss of Dnmt3a on cancer methylomes we next performed WGBS on DNA isolated from *Dnmt3a*$^{\Delta/\Delta}$ CLL and PTCL cells (n=1 biological sample per tumor type or normal control). Out of ~14 million CpG dinucleotides analyzed, we observed decreased methylation in 3,676,916 (26.5%) CpGs and increased methylation in 75,662 (0.5%) of CpGs in CLL relative to B-1a (**Fig. 27**). Although methylation changes in PTCL were less pronounced than in CLL, hypomethylation was still the major change, as decreased methylation was observed in 1,263,413 (9%) CpGs and increased methylation in 155,977 (1%) CpGs (**Fig. 27**). Unsupervised hierarchical cluster analysis using methylation readouts for all individual CpGs revealed tight clustering of normal cells, whereas tumors were both significantly distinct from both their normal counterparts and each other **(Fig. 28)**. Although the vast majority of changes in methylcytosine levels occurred in gene bodies and intergenic regions in both disease settings, we also observed significant changes in promoter methylation (**Fig. 29a**). Interestingly, the number of hypomethylated cytosines in promoters was almost 2.2 fold higher in CLL (61,356) than in PTCL (27,415) (**Fig. 29a**). In contrast, the number of hypermethylated cytosines in promoters were almost equal between CLL (5,831) and PTCL (5,757) (**Fig. 29b**). Notably, a relatively small number of cytosines associated with long promoters (-1,500 to +500bp relative to TSS) were commonly hypo- and hypermethylated in CLL and PTCL (15% and 22%, respectively; **Fig. 29c,** and data not shown). In contrast, differentially methylated cytosines in gene-bodies were more commonly shared between the two disease settings (48% hypomethylated and 33% hypermethylated CpGs **Fig. 29d**). Consistent with the greater number of hypomethylated CpGs observed in CLL, analysis of differentially methylated regions revealed that the number of hypomethylated promoters was ~ 3-fold higher in CLL relative to PTCL in both long and core (-300 to +150 relative to TSS) promoters (**Fig. 29e-g,** and data not shown). Interestingly, only 126 out of 500 long promoters hypomethylated in PTCL were also hypomethylated in CLL and this trend could also be seen in core promoters

(**Fig. 29g,** and data not shown). The difference in gene body hypomethylation was even more pronounced with more than 7-fold more genes affected in CLL than in PTCL (**Fig. 29h,** and data not shown). These data suggest that Dnmt3a may play broader functions in promoter and gene body methylation in B cells than in T cells. In addition to hypomethylation we also observed promoter and gene body hypermethylation in CLL and PTCL (**Fig. 29f-h**). Like with hypomethylation, most hypermethylation events were tumor type specific.

Analysis of the spatial distribution of DNA methylation changes revealed that areas around the TSS were –not surprisingly – the least methylated areas in both normal and tumor cells (**Fig. 30**). Consistently with analysis of individual elements, the degree of hypomethylation was higher in CLL than in PTCL, with no substantial differences observed in the spatial distributions of methyl cytosines across gene elements (**Fig. 30**).

Locus-specific analysis revealed that hypo- and hypermethylated promoters were relatively equally distributed across the genome, with exception of the X chromosome in which very few differentially methylated promoters were detected in CLL or PTCL (**Fig. 31**). The highest number of hypomethylated promoters in CLL were present on chromosomes 11 and 9, whereas in PTCL chromosomes 11 and 5 were most affected (**Fig. 31**). Like in human CLL [67], large-scale gene body hypomethylation (21,157 genes) was observed in the CLL genome, along with hypermethylation of 196 genes bodies. The trend was similar, but less pronounced in the PTCL methylome, with little overlap between tumor types (**Fig. 29h**).

To further validate data obtained by global methods we performed locus-specific methylation analysis using Combined Bisulfite Restriction Analysis (COBRA) on promoters specifically hypomethylated in either CLL or PTCL. Similar to WGBS, promoters of *Sgk3, Nfam1, Pstpip2, and Plscr1* were found to be hypermethylated in normal B-1a cells, normal CD8 T cells, and *Dnmt3a*$^{\Delta/\Delta}$ PTCL, but were hypomethylated in three independent *Dnmt3a*$^{\Delta/\Delta}$ CLL samples (**Fig. 32a**). We also confirmed that the promoters of *Crtam, Sgk1, Samd3, Oas3,* and *Fam169b* were

hypermethylated in normal B-1a, normal CD8 T cells, and *Dnmt3a*$^{\Delta/\Delta}$ CLL, but were hypomethylated in *Dnmt3a*$^{\Delta/\Delta}$ PTCL (**Fig. 32b**). Overall, our data demonstrate that changes in promoter methylation identified using WGBS likely represent common events occurring in mouse CLL and PTCL and demonstrate the cell-type specific patterns induced by loss of Dnmt3a.

To determine if promoter hypomethylation is cancer-specific or is present in other normal cell types, we evaluated the methylation status of promoters in a range of B cell subsets (immature B cells, mature B cells, B-1a, CD19+, follicular B cells, and marginal zone B cells), T cells (CD4 and CD8), and myeloid cells. Interestingly, for the majority of genes analyzed, the promoters were heavily methylated in all normal cell types (**Fig. 32c**) suggesting that promoter hypomethylation does not occur during differentiation but it rather is tumor-type specific event. Altogether, our data suggest that Dnmt3a may have cell-type specific functions in maintaining promoter and gene-body methylation that seems to be more pronounced in B-1a cells than in CD8+ cells, but which may nonetheless be critical to prevent transformation of both normal cell types.

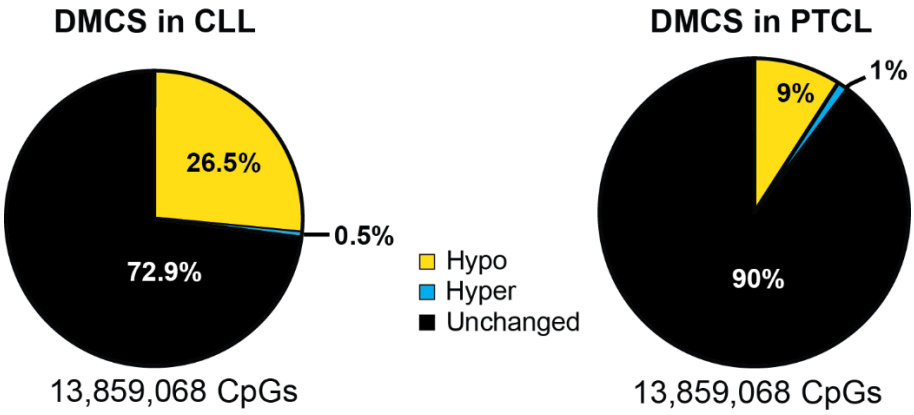


Figure 27. Differentially methylation cytosines (DMCS) in CLL and PTCL. Methylation status of 13,859,068 CpGs in *Dnmt3a*$^{\Delta/\Delta}$ CLL relative B-1a (left) and *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8+ T cells (right). Hypomethylated (yellow) and hypermethylated (blue) CpGs had a 30% or greater decrease or increase in methylation, respectively. CpGs whose methylation was unchanged are shown in black.

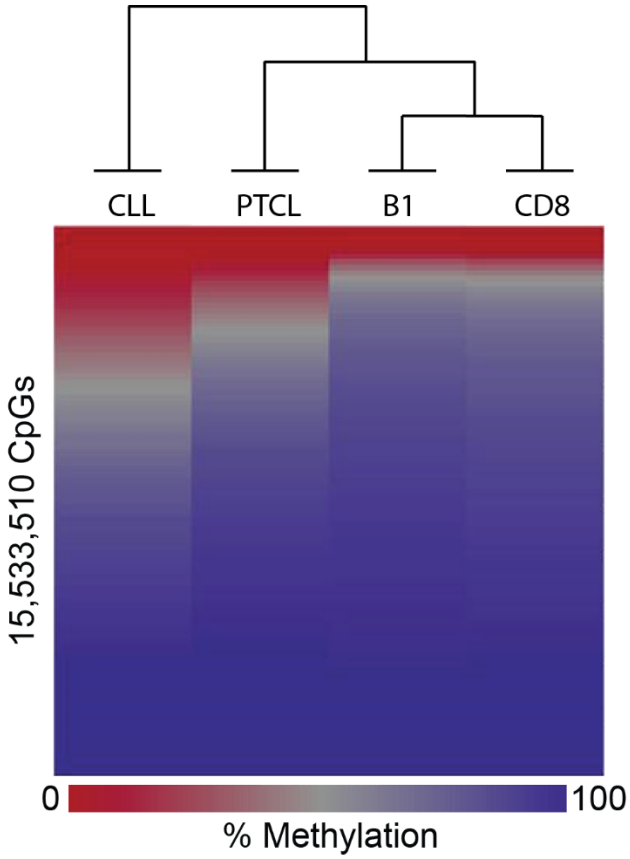


Figure 28. Methylation status of 15,533,510 CpG dinucleotides in control B-1a, control CD8, and Dnmt3a-deficient CLL and PTCL cells as determined by whole-genome bisulfite sequencing (WGBS). Hypomethylated CpGs are shown in red whereas hypermethylated CpGs are in blue. Hierarchical clustering of samples derived from WGBS datasets are shown at the top.

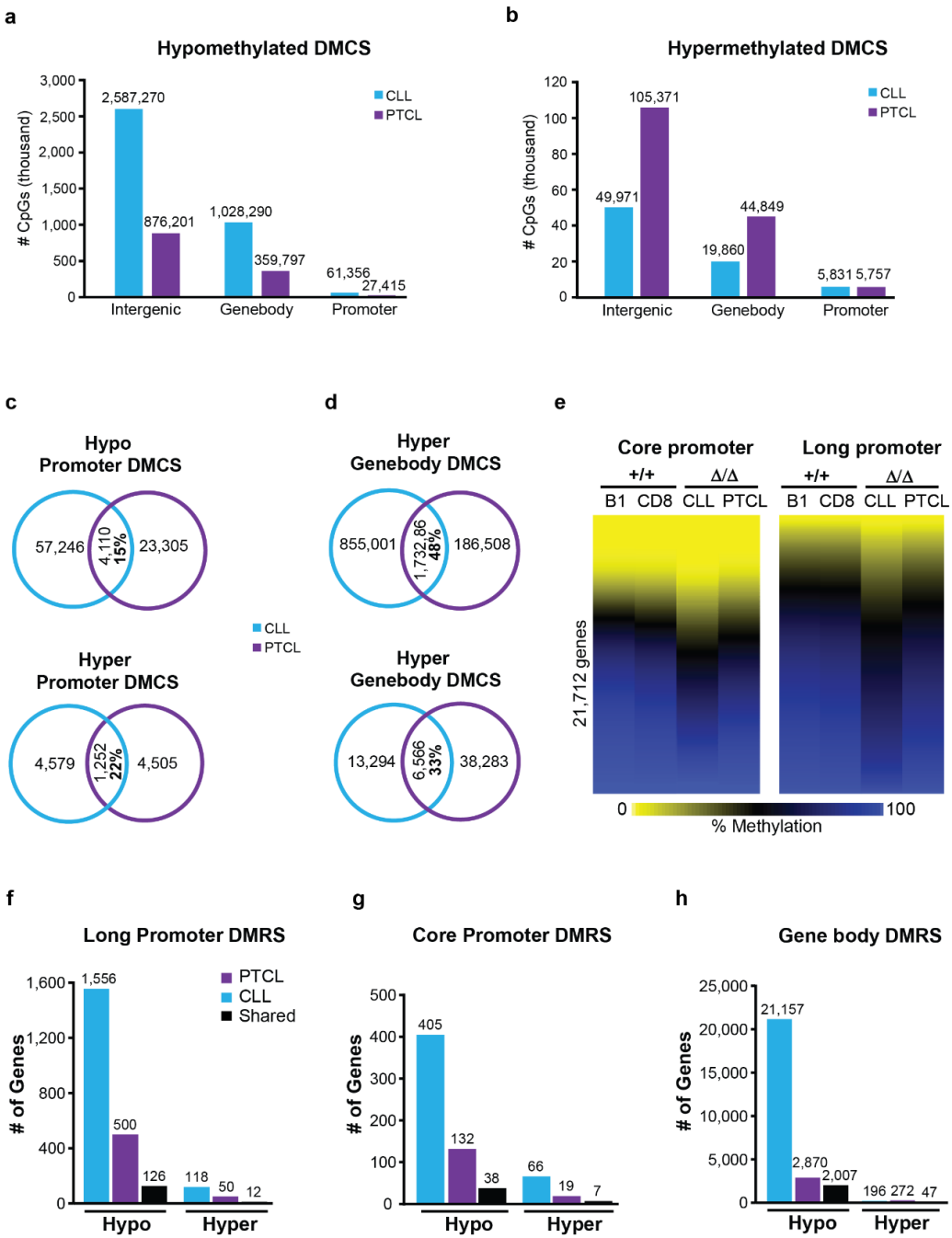


Figure 29. DNA methylome of *Dnmt3a*$^{\Delta/\Delta}$ CLL and PTCL. (a) Breakdown of hypomethylated and **(b)** hypermethylated CpGs (differentially methylated CpGs; DMCS) in *Dnmt3a*$^{\Delta/\Delta}$ CLL (relative to B-1a) and PTCL (relative to CD8) categorized by genomic location. Differentially methylated CpGs (defined as a methylation difference of ≥30%) were annotated to gene promoters (-1,500 to +500bp relative to TSS), gene-bodies, or intergenic regions. **(c)** Venn diagram showing unique and overlapping hypomethylated (methylation <30%) and hypermethylated (methylation >70%) CpGs annotated to long promoters (-1,500 to +500bp relative to TSS) and **(d)** gene-bodies in CLL (blue) and PTCL (purple). **(e)** Heat map displaying the methylation status of 21,712 promoters as determined by WGBS in B1, CD8, *Dnmt3a*$^{\Delta/\Delta}$ CLL, and *Dnmt3a*$^{\Delta/\Delta}$ PTCL. Methylation percentage for individual CpGs were annotated to the core promoter regions (-300bp to +150bp relative to the TSS, *left*) and long promoter regions (-300bp to +150bp relative to the TSS, *right*). Methylation percentages for all CpGs across the region were averaged to give a mean methylation value for each gene promoter. Hypomethylation is shown in yellow and hypermethylation in blue. **(f)** The number of genes with differentially methylated regions (DMRS) in their long promoters, **(g)** core promoters and **(h)** gene-bodies in *Dnmt3a*$^{\Delta/\Delta}$ PTCL (purple) relative to CD8 control, and *Dnmt3a*$^{\Delta/\Delta}$ CLL (blue) relative to B-1a control. The DMRS shared between tumor types are shown in black.

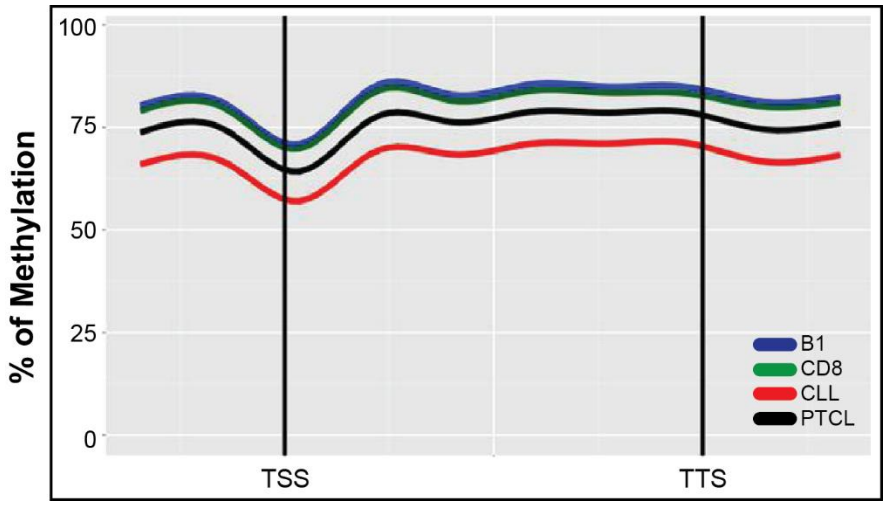


Figure 30. Average CpG methylation percentage for B-1a (blue), CD8 (green), *Dnmt3a*$^{\Delta/\Delta}$ CLL (red), and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (black) relative to their position within genes. Location of transcription start site (TSS) and the transcription termination site (TTS) are shown.

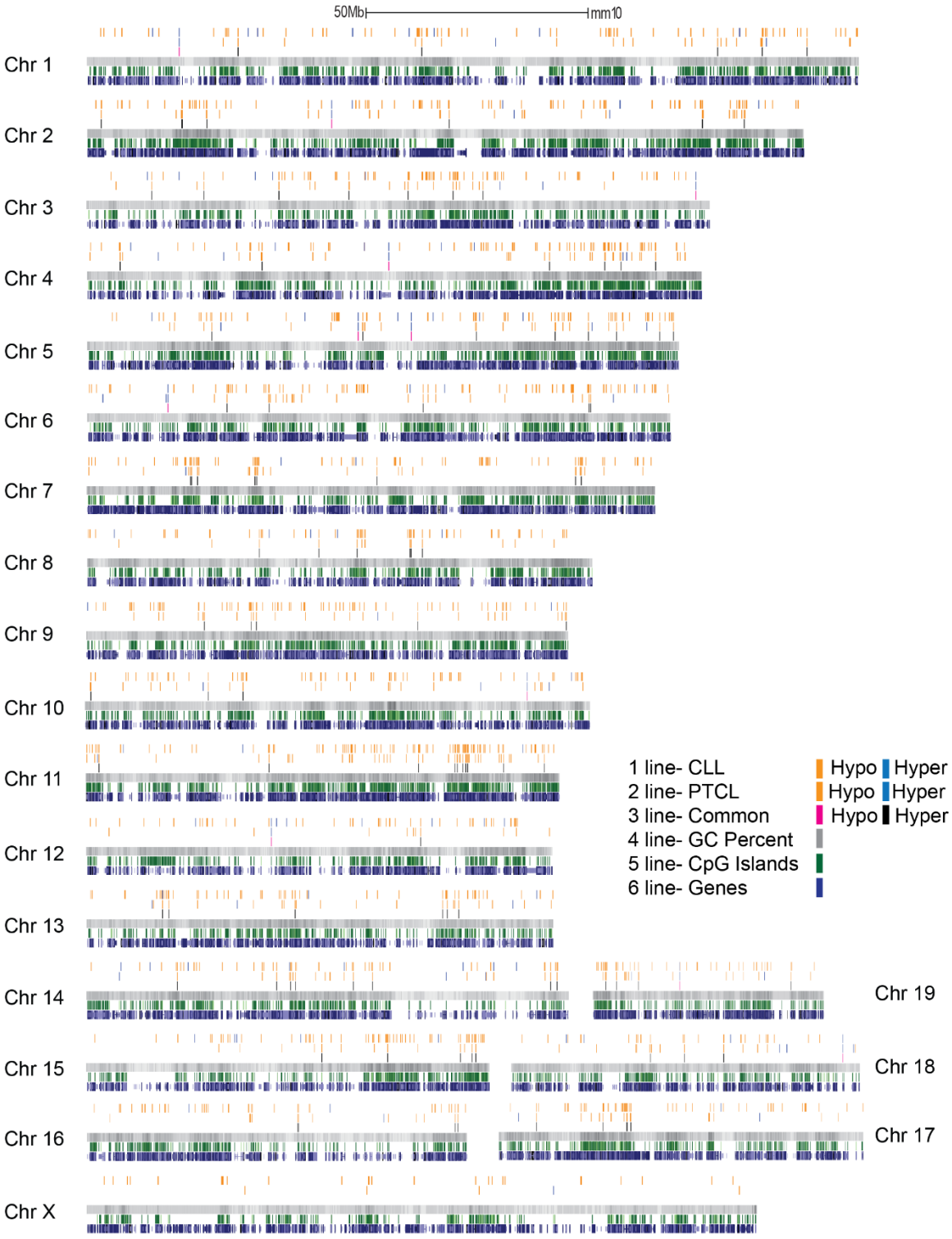


Figure 31. Loss of Dnmt3a results in unique methylation patterns in *Dnmt3a*$^{\Delta/\Delta}$ CLL and PTCL. Chromosome plot of differentially methylated regions annotated to the promoter region (-1,500 to +500bp relative to TSS) in *Dnmt3a*$^{\Delta/\Delta}$ CLL (1st row, orange bar) relative to B-1a and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (2nd row, purple bar) relative to CD8 control. Hypomethylated DMRS are indicated by orange lines and hypermethylated DMRS by blue. DMRS that are shared between the two tumor types are also shown (3rd row, red bar). The CG percentage (4th row, gray bar), location of CpG islands (5th row, green bar) and location of genes (6th row, blue bar) are also shown. 50Mb scale is shown to reference chromosome length.

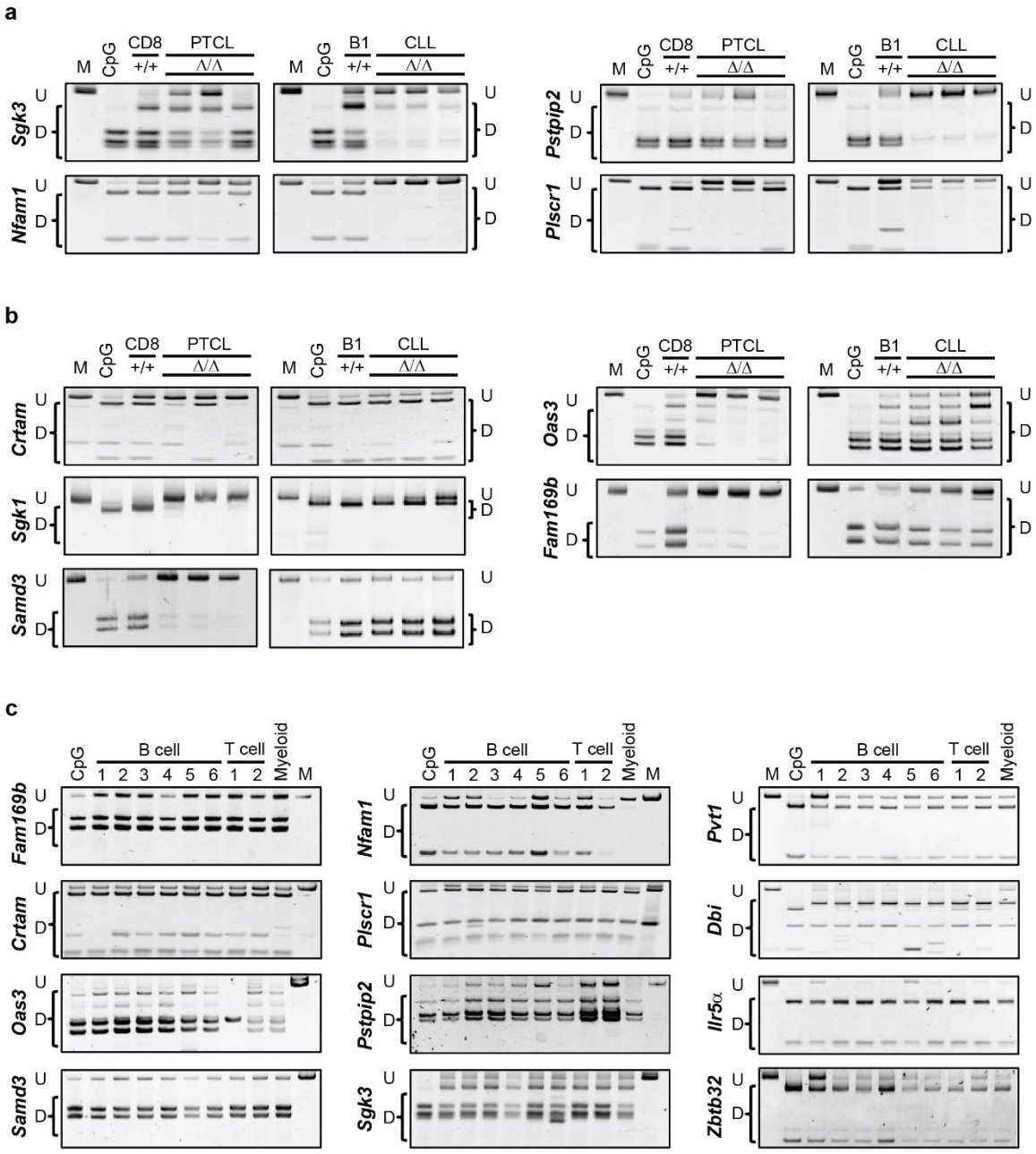


Figure 32. Promoter hypomethylation in *Dnmt3a*$^{\Delta/\Delta}$ CLL and PTCL is common in additional tumors. COBRA analysis of promoter methylation for **(a)** *Sgk3, Nfam1, Pstpip2, Plscr1* and **(b)** *Crtam, Sgk1, Samd3, Oas3 and Fam169b* in *wild-type* CD8+ T cells, *wild-type* B-1a cells, *Dnmt3a*$^{\Delta/\Delta}$ PTCL and *Dnmt3a*$^{\Delta/\Delta}$ CLL samples. PCR fragments amplified from bisulfite-treated genomic DNA were digested with the restriction enzymes *BstU*I, *Taq*1 or *Tai*I. Undigested (U) and digested (D) fragments correspond to unmethylated and methylated DNA, respectively. Control PCR fragments generated from fully methylated mouse genomic DNA that is undigested (M) or digested (CpG) are shown. **(c)** COBRA analysis of promoter methylation for *Fam169b, Crtam, Oas3, Samd3, Nfam1, Plscr1, Pstpip2, Shk3, Pvt1, Dbi, Il5ra,* and *Zbtb32.* The following samples were isolated from *wild-type* mice, B cell subsets: (1) splenic B-1a, (2) bone marrow immature B cells, (3) bone marrow mature B cells, (4) splenic CD19+ B cells, (5) splenic marginal zone B cells, (6) splenic follicular B cells, T cell subsets: (1) splenic CD4+ T cells, (2) splenic CD8 T cells and lastly CD11b+ (Myeloid) splenic cells.

***Dnmt3a*$^{\Delta/\Delta}$ CLL and *Dnmt3a*$^{\Delta/\Delta}$ PTCL has distinct transcriptomes.**

Next, we performed RNA-seq on *Dnmt3a*$^{\Delta/\Delta}$ PTCL and CLL to identify genes differentially expressed relative to CD8+ and B-1a cells, respectively. As with methylation, expression patterns in CLL and PTCL were dramatically different with only 235 upregulated (20%) and 194 downregulated (27%) genes common between the two tumor types (**Fig. 33a, 33b,** and data not shown). *Dnmt3a*$^{\Delta/\Delta}$ CLL had 555 upregulated and 204 downregulated genes that were not common to PTCL, while *Dnmt3a*$^{\Delta/\Delta}$ PTCL had 951 uniquely upregulated and 523 downregulated genes (**Fig. 33a, 33b,** and data not shown). Ingenuity pathway analysis (IPA) using all gene expression changes found in tumors did not provide a clear clue to the pathways promoting CLL or PTCL development resulting in many pathways being activated or inhibited in *Dnmt3a*$^{\Delta/\Delta}$ diseases (**Fig. 34**). Comparison of promoter methylation and gene expression revealed that 129 genes (8%) whose promoters were hypomethylated in CLL were associated with overexpression (**Fig. 33c, 35** referred to herein as HOC genes – Hypomethylated and overexpressed in CLL). A similar comparison of promoter methylation and gene expression revealed that 84 genes (17%) whose promoters were hypomethylated in PTCL were associated with overexpression (**Fig. 33c, 36,** referred to herein as HOT genes – Hypomethylated and overexpressed in TCL). In contrast, we detected only two genes for PTCL and one gene for CLL whose hypermethylation correlated with underexpression, suggesting that most of the cancer-specific hypermethylation has little effect on gene expression and tumor progression (**Fig. 33d**). Altogether, these data demonstrate that hypomethylation affects gene expression on a broader scale than hypermethylation in mouse *Dnmt3a*$^{\Delta/\Delta}$ CLL and PTCL.

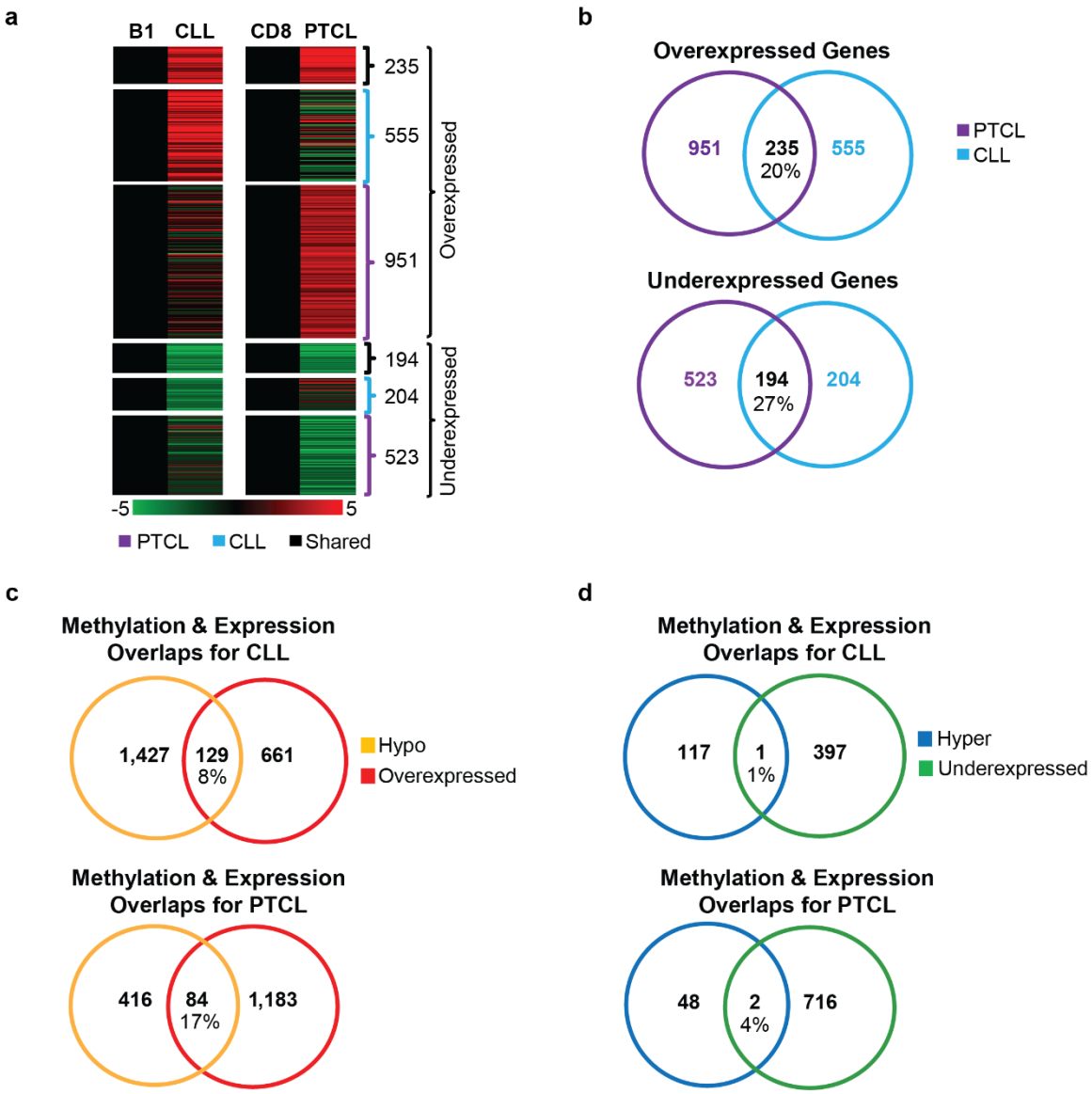


Figure 33. Gene expression is deregulated in a cell-specific manner in *Dnmt3a*$^{\Delta/\Delta}$ CLL and PTCL. (a) Heat maps derived from global expression profiling by RNA-seq. 555 Genes were overexpressed exclusively in *Dnmt3a*$^{\Delta/\Delta}$ CLL (N=8) relative to B-1a (B1) controls (N=2), 951 genes were overexpressed exclusively in *Dnmt3a*$^{\Delta/\Delta}$ PTCL (N=3) relative to CD8 controls (N=2), and 235 genes were overexpressed in both tumor types. 204 Genes were underexpressed exclusively in *Dnmt3a*$^{\Delta/\Delta}$ CLL relative to B1 controls, 523 genes were underexpressed exclusively in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8 controls, and 194 genes were underexpressed in both tumor types. A color bar is shown to reference fold change with upregulation in red and downregulation in green. Only genes with a fold change ≥2 and a q-value <0.05 are shown. **(b)** Venn diagram showing the number of unique and overlapping differentially expressed genes in CLL (blue) and PTCL (purple). **(c-d)** A graphical presentation of the number of genes differentially expressed (red; overexpression, green; underexpression) and the number of genes differentially methylated at the promoter region (yellow; hypomethylation, blue; hypermethylation) in *Dnmt3a*$^{\Delta/\Delta}$ CLL relative to B1 controls and in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8 controls. The number of genes with corresponding methylation and expression changes (Hypomethylated and overexpressed; top and hypermethylated and underexpressed; bottom) are shown in the overlapping portion of the circles.

PTCL	
Physiological System Development and Function	**# Genes**
Hematological System Development & Function	429
Tissue Morphology	399
Organismal Survival	489
Hematopoiesis	246
Lymphoid Tissue Structure & Development	243
Diseases and Disorders	**# Genes**
Immunological Disease	470
Endocrine System Disorders	307
Gastrointestinal Disease	117
Metabolic Disease	234
Cancer	1444
Inhibited Pathways	**# Genes**
TNFR1 Signaling	15
TNFR1 Signaling	11
iNOS signaling	13
RANK signaling in osteoclasts	18
CD18 Signaling in T Helper cells	21
Activated Pathways	**# Genes**
ATM signaling	22
Estrogen-mediated S-phase Entry	11
Cyclins and Cell Cycle Regulation	17
Apoptosis signaling	17
PDGF Signaling	14

CLL	
Physiological System Development and Function	**# Genes**
Hematological System Development & Function	346
Tissue Morphology	297
Immune Cell Trafficking	222
Hematopoiesis	179
Lymphoid Tissue Structure & Development	172
Diseases and Disorders	**# Genes**
Immunological Disease	340
Endocrine System Disorders	230
Gastrointestinal Disease	769
Metabolic Disease	177
Inflammatory Response	332
Inhibited Pathways	**# Genes**
iSOC-iCOSL Signaling in T Helper Cells	22
Calcium-induced T Lymphocyte Apoptosis	13
p53 signaling	14
G2/M DNA Damage checkpoint Regulation	9
Mouse Embryonic Stem Cell Pluripotency	12
Activated Pathways	**# Genes**
Pattern Recognition Receptors for Bacteria/Viruses	22
Toll-like Receptor Signaling	15
IL-6 Signaling	17
Estrogen-mediated S-phase Entry	6
Signaling by Rho Family GTPases	24

Figure 34. Summary of Ingenuity Pathway analysis (IPA) of all differentially expressed genes (FPKM ≥ 10) in *Dnmt3a*$^{\Delta/\Delta}$ PTCL and *Dnmt3a*$^{\Delta/\Delta}$ CLL relative to control samples. Genes with a fold change ≥ 2 and a q-value < 0.05 were used in the analysis. P-values were less than 0.05 for all categories.

Fold change	Mut % Meth	B1 % Meth	Gene
4.5	36	87	Prc1
8.2	26	78	Prkar2b
2.7	6	85	Pstpip2
2.1	19	71	Ptms
2.0	20.5	76.5	Ptp4a3
2.9	13	80	Pvt1
3.5	16	71	Racgap1
2.6	11	83.5	Rbm47
3.8	20	84	Rdh12
2.9	20	81	Rtn4ip1
5.3	3	67	S100a4
2.2	15	69	Sgk3
6.6	31	82	Sirt2
2.0	14	82	Slc37a2
3.1	20	90	Slc39a4
2.3	19	79	Slc7a7
2.2	13	76	Snx20
2.3	33	86	Soat1
2.5	10	71	Sp110
7.0	15	65	Spire1
4.7	9	65	Sspn
2.0	13	66	St3gal2
2.7	13	89	St3gal6
2.5	12	75	Stx7
2.8	7	59	Tagln2
2.1	17.5	91	Ticam1
4.6	2	63	Ticam2
2.9	9	62	Tlr2
7.7	5	69	Tmem106a
2.0	9	90	Tmem229b
5.3	32	90	Tnf
2.2	22	87	Tnfrsf13b
2.6	23	88	Tns1
2.3	8	82	Txn2
2.4	5	73	Ubash3b
5.2	2	59	Uhrf1
12.6	17	84	Upb1
2.3	4	50	Ypel2
22.2	9	58	Zbtb32
2.2	19.5	81.25	Zbtb38
5.6	35	91	Zcchc14
4.7	16	66	Zcchc18
3.3	15	70.5	Zeb2

Fold change	Mut % Meth	B1 % Meth	Gene
13.1	12	70	Gas7
379.0	14	82.5	Gnb3
32.1	21	78	Gpm6a
3.9	9	87	H6pd
4.8	13	77	Hepacam2
2.8	10	64	Hmga2-ps1
3.5	16	90	Hpse
15.1	19	75	Ifitm6
2.3	27	91	Ift27
2.6	9	69	Igsf8
4.0	11	81	Il5ra
3.0	20	86	Il9r
2.2	25	84	Inppl1
2.9	9	97	Irgm2
33.9	15	71	Itgam
2.7	15	83	Itgb7
2.7	31	91	Itpr1
3.1	27	82	Itsn1
2.1	11	84	Krt222
6.0	16	86	Lmna
2.8	11	91	Lsp1
2.8	9	85	Ly6c2
2.8	29	89	Man1c1
4.5	7	94	Mgmt
2.2	15.5	83.5	Mtss1
17.9	19	81	Myadm
2.0	8	71	N4bp3
2.5	18	85	Ncf4
4.1	26	96	Neurl1a
6.5	11	84.5	Nfam1
2.1	14	65	Nfkbiz
3.1	8	75	Nrp2
10.0	23	84	Ntng2
10.5	10	59	Ociad2
2433.2	9	67	Olfml2a
4.7	19	87.5	Pdcd1lg2
2.5	35	90	Pfkfb3
30.0	6	87	Pik3r6
2.9	27	89	Plec
4.4	11	76	Plscr1
4.0	17	80	Pon3
8.3	26	89	Ppfia4
2.8	9	60	Ppil1

Fold change	Mut % Meth	B1 % Meth	Gene
2.9	15.5	78.5	1810046K07Rik
8.6	19	87	2700054A10Rik
4.0	19	86	2810417H13Rik
2.4	12	65	Abi3
7.4	26	90	Adamtsl4
17.7	28	88	AF067061
6.7	22	91	Ahnak
105.9	10	72	AI427809
5.3	13	87	Aldh3b1
2.1	7	60	Anxa11
12.8	3	58	Anxa3
2.9	3	66	Apobec3
4.8	1	65	Arap3
3.5	20	81	Arhgap33
2.2	28	79	Arid3b
3.5	20	93	Aurkb
93.2	20	83	Avil
2.8	38	88	BC064078
2.4	27	80.5	Blvrb
8.5	21	74	C2
2.7	4	77	Capn2
18.3	10	88	Cd300ld
2.8	16	75.333	Cd80
4.8	18	85	Cdc42ep4
2.4	8	65	Cdca7
2.9	2	56	Cisd3
2.2	20	93	Cnp
5.3	13	78	Crip1
8.1	25	90	Cyp11a1
2.7	1	49	Dbi
8.6	5	61	Dgkg
3.4	10	82	Dse
3.7	10	58	E2f2
2.6	19	79	Ebi3
2.0	19	82	Ece1
2.4	32	88	Eps8
7.3	20	86	Espn
5.0	16	82	Etv5
2.7	24	86	Evi5
19782.5	18	76	Fabp7
2.2	35	95	Fam46c
2.5	12	77	Fdps
3.3	15.5	81	Fgd2

Figure 35. Genes Hypomethylated and Overexpressed in *Dnmt3a*$^{\Delta/\Delta}$ CLL (HOC genes). List of HOC genes. Percentage of promoter methylation in B-1a (B1 % meth; blue), and *Dnmt3a*$^{\Delta/\Delta}$ CLL (CLL % meth; yellow) is shown within boxes. Similarly, fold differences in gene expression between *Dnmt3a*$^{\Delta/\Delta}$ CLL relative to B-1a (B1 vs Mut; red) is shown within boxes.

Exp	PTCL % Meth	CD8 % Meth	Gene
5.4	12	87	Per3
31.7	27	74	Pif1
22.1	6	73	Plac8
2.8	10	71	Plekha8
6.0	4	55	Ppil1*
3.1	23	95	Pvt1*
2.0	13	68	Racgap1
2.0	22	82	Ran
3.4	14	63	Reep5
3.9	8	92	Rnasel
5.4	8	82	Rnf43
9.2	10	82	Samd3
3.2	37	85	Sh2d1a
2.0	22	90	Sh2d3c
3.3	19	85	Sla
2.2	37	90	Sntb2
8.6	14	85	Stat1
7.7	2	68	Tmem37
2.3	5	64	Tmpo
3.2	34	88	Trim14
17.6	7	74	Ttll11
2.1	21	87	Txk
2.0	8	75	Uba7
8.2	3	70	Wfikkn2
2.0	30	81	Wipf1
2356.8	17	75.5	Wnt8a
4.7	25	74	Xlr3b
3.3	11	70	Zfp808

Exp	PTCL % Meth	CD8 % Meth	Gene
4.9	11.5	77.5	Ifi47
3.5	8	83	Ifitm10
2.3	10	94	Ikzf3
3.1	26	85	Il18rap
2.8	10	88	Il2rb
3.3	10	74	Impa1
9867.8	21	84	Islr
5.4	21	85	Jdp2
3.2	31	86	Keap1
5.9	7	77	Klrc1
2.7	10	80	Klrd1
22.9	7.5	85	Klre1
10.7	18	85	Lpar5
4.6	5	53	Lyrm9
22.8	22	86	Mmp14
13.7	12	86	Ms4a4b
8.0	50	98	Ms4a4c
2.5	23	91	Myo1c
8.2	7	61	Myo6
4.6	34	91	N4bp1
3.1	4	61	Nfe2l2
2.4	31	96	Nkg7
2.6	13	76	Nod1
23.6	42	91	Oas2
12.0	14	82	Oas3
17.8	17	78	Osbpl3
2.1	30	84	Oxr1
5.0	19	86	P2rx7

Exp	PTCL % Meth	CD8 % Meth	Gene
7.6	20	79	1700025G04Rik
14.2	6	51	1700048O20Rik
10.4	20	77	4921525O09Rik
3.8	12	76	Abi3
3.4	9	69	Acot7
5.2	11	71	Alpk3
6.8	12	87	Amica1
11.3	7	70	Apobr
10.2	8.5	81.5	Arl4d
4.5	1	81	Atp8b4
4.7	11	78	AW112010
2.7	6	75	B4galt5
2.6	9	86	C920025E04Rik
28.9	2	72	Ccr2
2.4	28	83	Celsr1
3.6	7	70	Cln3
4.6	12.5	82	Coro2a
3.2	10	74	Crtam
6.6	9	81	Cxcr5
4.6	6	62	Fah
3.5	50	94	Fcgrt
12.2	7	88	G0s2
5.4	12	67	Gdf11
3.3	2	64	Gimap4
3.1	3	70	Gimap7
9.3	19	66	Gm19705
1437.6	4	71	Gpnmb
3.4	6	79	Gzmm

Figure 36. Genes Hypomethylated and Overexpressed in *Dnmt3a*$^{\Delta/\Delta}$ PTCL (HOT genes). List of HOT genes. Percentage of promoter methylation in CD8 (CD8 % meth; blue), and *Dnmt3a*$^{\Delta/\Delta}$ PTCL (PTCL % meth; yellow) is shown within boxes. Similarly, fold differences in gene expression between *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8 (CD8 vs Mut; red) is shown within boxes. Genes common between HOC and HOT datasets are shown in red*.

Identification of genes commonly overexpressed in mouse and human CLL and PTCL as potential drivers of disease development.

To determine the extent of similarity between mouse and human disease on the molecular level, we compared gene expression signatures obtained from mouse CLL and PTCL to the corresponding human disease. We utilized RNA-seq data obtained on a set of 10 human CLL samples [155], in which 1,605 genes were up- and 1,227 genes were down-regulated relative to normal CD19+ B cells (not shown). When we compared expression of these genes to those commonly deregulated in mouse *Dnmt3a*$^{\Delta/\Delta}$ CLL we found, we found a 21% overlap (139/659) between overexpressed genes and an 11% (40/352) overlap for underexpressed genes (**Fig. 37a, 37b,** and data not shown). The extent of overlap in up-and downregulated genes was significant for both comparisons ($P<0.00001$ and $P<0.00003$, respectively), suggesting that similar molecular events may drive CLL in both species.

Like for CLL, we were interested in determining the extent of overlap between genes deregulated in mouse and human PTCL. Therefore, we analyzed genes commonly up- and down-regulated in mouse *Dnmt3a*$^{\Delta/\Delta}$ PTCL and human PTCL. As a source of human gene expression we utilized data obtained by microarray on a set of 15 PTCL samples and 5 normal T cell samples [42, 156]. Our analysis revealed that 39% of genes (376/960) overexpressed in mouse PTCL were also overexpressed in human PTCL samples (**Fig. 37c, 37d,** and data not shown**)**. Similarly, 38% of genes (244/641) under-expressed in mouse PTCL were also under-expressed in human PTCL samples. Like in CLL, the extent of overlap in up-and down-regulated genes was significant ($P<0.00001$ for both comparisons), suggesting that similar molecular events may drive PTCL in both species.

To gain further insight into genes that may drive CLL and PTCL we considered the simplest explanation for the tumor suppressor function of Dnmt3a in the prevention of mouse CLL and PTCL – loss of methylase activity of Dnmt3a results in hypomethylation and overexpression

of subset of genes whose inappropriate expression contributes to cellular transformation. We identified 129 and 84 genes hypomethylated and overexpressed in CLL (HOC genes) and PTCL (HOT genes), respectively (**Fig. 35, 36**). In CLL, 23 HOC genes were overexpressed in human tumors, with 14 reported to be hypomethylated [67] (**Fig. 38a**). Similarly, 26 HOT genes were overexpressed in human CD8+ PTCL but their methylation status have not been reported yet (**Fig. 38b**). Using human CLL and PTCL samples, we confirmed the overexpression of two genes from the HOT and HOC signature. In particular, *Zbtb32* (Zinc Finger And BTB Domain Containing gene) was highly up-regulated in both human and mouse tumors relative to normal controls, suggesting that its increased expression is a common event in human and mouse CLL (**Fig. 38c**). Similarly, *Stat1* (Signal Transducer And Activator Of Transcription 1) was highly up-regulated in both human and mouse PTCL relative to normal CD3+ T cells, suggesting that its increased expression is a common event in human PTCL (**Fig. 38d**). Altogether, these data suggest that mouse and human CLL and PTCL could have common drivers of disease development and that the *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ model is a suitable model to study pathology of CLL and PTCL.

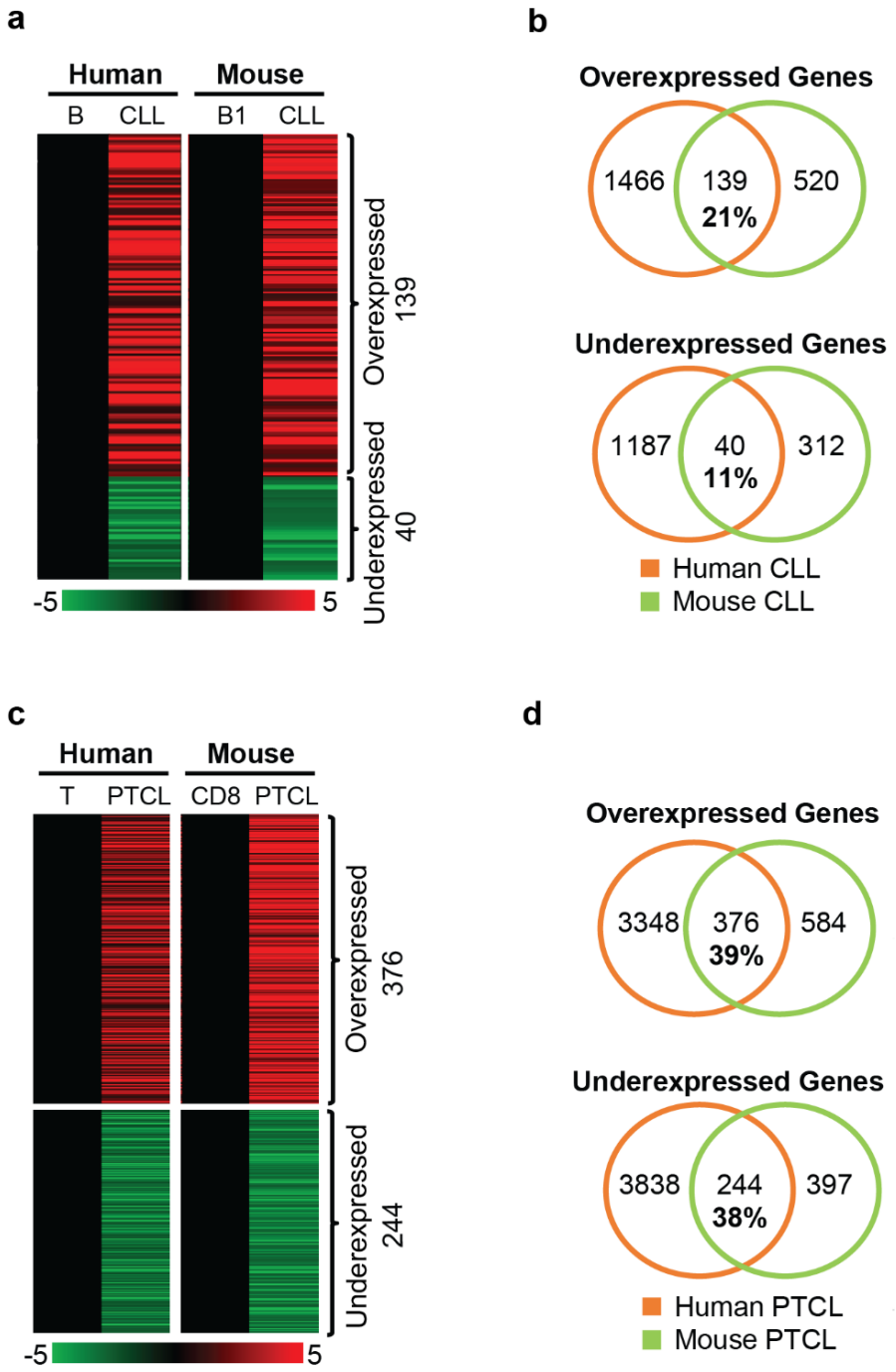


Figure 37. Gene expression is partially conserved between mouse and human CLL and PTCL. (a) Heat maps derived from global expression profiling by RNA-seq for human and mouse CLL. 139 genes were commonly overexpressed while 40 genes were commonly underexpressed in human CLL relative to CD19+ B cells and mouse *Dnmt3a*$^{\Delta/\Delta}$ CLL relative to B1-a controls. Only genes with a fold-change >2 and q-value <0.05 were considered differentially expressed. **(b)** Venn diagrams showing overlaps in gene expression between human and mouse CLL datasets, as determined in panel a**. (c)** Heat maps derived from global expression for human PTCL (determined by microarray) and mouse PTCL (determined by RNA-seq). 376 genes were commonly overexpressed, while 244 genes were commonly underexpressed between human PTCL relative to normal tonsil T cells and mouse *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8 controls. For microarray data, genes with a fold change >1.5 and a P-value <0.05 were considered significant. For RNA-seq data, genes with a fold change >2 and a q-value <0.05 were considered significant. **(d)** Venn diagrams showing overlaps in gene expression between human and mouse PTCL datasets, as determined in panel c.

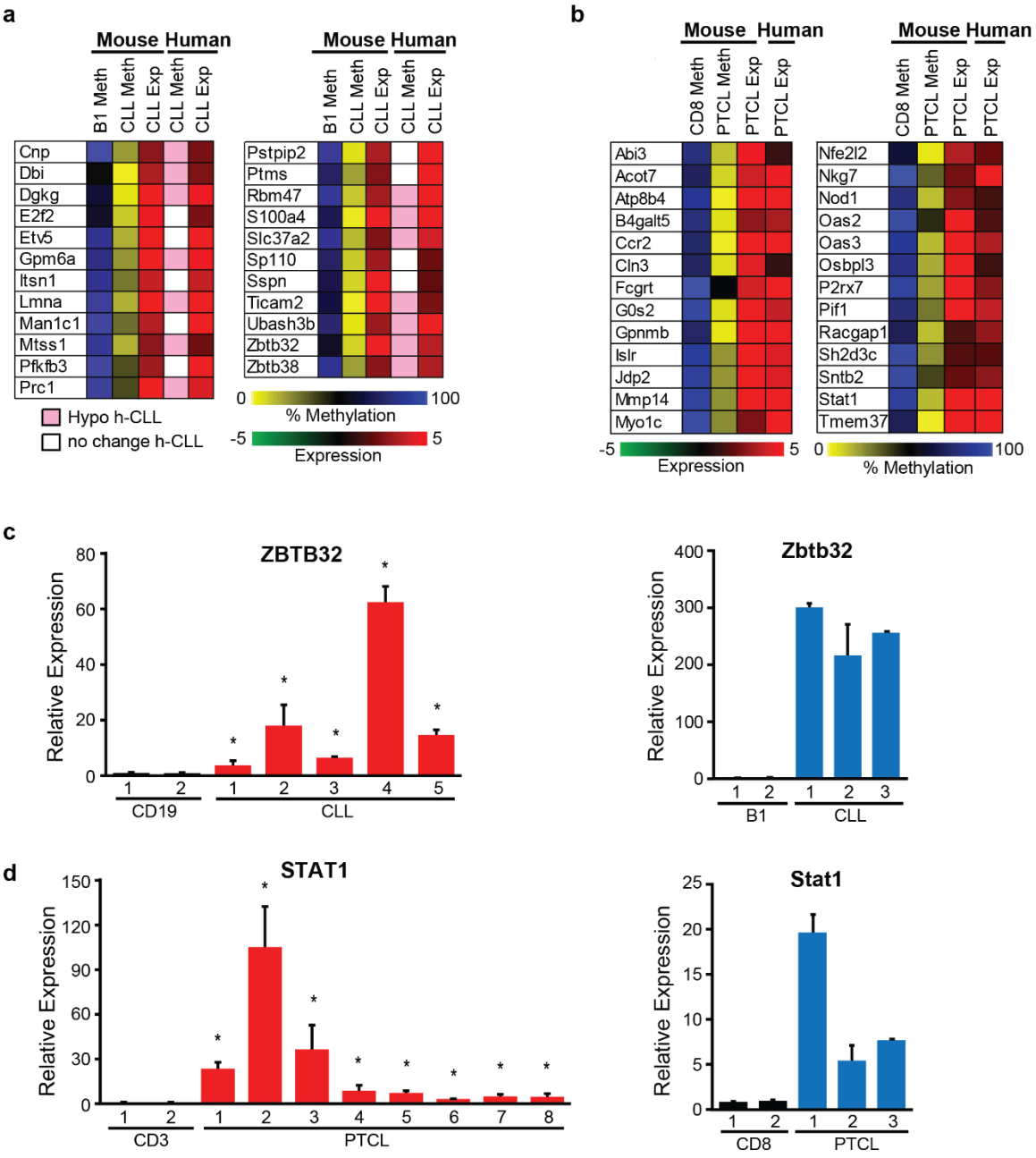


Figure 38. Identification of genes hypomethylated and overexpressed in mouse and human CLL and PTCL. (a) Heat map showing genes that were identified as hypomethylated and overexpressed in *Dnmt3a*$^{\Delta/\Delta}$ CLL relative to B1-a controls (B1) that were also overexpressed in human CLL samples, as determined by RNA-seq. Genes that were identified to be hypomethylated at their promoters in human CLL are shown in pink, while those with no-change in promoter methylation are shown in white. Hypomethylation is shown in yellow and hypermethylation in blue, while overexpression is shown in red. **(b)** Heat map showing genes that were identified as hypomethylated and overexpressed in *Dnmt3a*$^{\Delta/\Delta}$ PTCL relative to CD8 controls that were also overexpressed in human PTCL samples, as determined by microarray. Hypomethylation is shown in yellow and hypermethylation in blue, while overexpression is shown in red. **(c)** Real-time qRT-PCR analysis of ZBTB32 expression in human CD19 controls and CLL samples and analysis of STAT1 expression in human CD3 controls and PTCL samples. Averages of two independent experiments are presented. (*) denotes P<0.05 (two-tailed student's t-test) with error bars representing 1 Stdev. **(d)** Real-time qRT-PCR analysis of Zbtb32 expression in mouse B1 controls and *Dnmt3a*$^{\Delta/\Delta}$ CLL samples and analysis of Stat1 expression in mouse CD8+ T cell controls and *Dnmt3a*$^{\Delta/\Delta}$ PTCL samples. Averages of two independent experiments are presented. (*) denotes P<0.05 (two-tailed student's t-test) with error bars representing 1 Stdev.

Discussion

DNA methylation is an important mechanism of transcriptional regulation that plays a role in the proper differentiation of hematopoietic stem cells into differentiated lineages [25, 83]. In cancer, the key methylation event contributing to tumorigenesis is considered to be promoter hypermethylation, which results in the inactivation of a wide range of tumor suppressor genes such as VHL, p16 and MLH1 [84]. However, emerging evidence suggest that inactivation of methylase activity promotes the development of hematologic malignancies and is almost uniformly accompanied by large scale hypomethylation [79, 87-89, 93, 95, 96]. Such hypomethylation is not limited to non-coding areas of the genome but also affects a large number of promoters and gene bodies. As a result, the untimely expression of genes normally silenced may contribute to cellular transformation of normal cells. Thus, to begin to understand what molecules may be responsible for the transformation of normal cells, high resolution genome-wide approaches are needed to profile both normal and tumor cells to identify the nature of deregulated events. Here we used WGBS coupled with RNA-seq to determine the extent to which two malignancies that develop in the absence of Dnmt3a, CLL and PTCL, share deregulated molecular events. Analysis of promoter methylation revealed conservation between 90% of hypomethylated and 83% of hypermethylated promoters in B-1a B cells and CD8+ T cells. Similarly, gene expression was highly alike between these two cell types, with 91% and 95% of highly and lowly expressed genes overlapping, respectively. These results suggest a high level of similarities between normal methylomes and transcriptomes of B-1a and CD8+ T cells. These results are in stark contrast to malignant B-1a and CD8+ cells. Only 8% of hypomethylated promoters detected in CLL were also hypomethylated PTCL. Similarly, only 10% of hypermethylated promoters were shared between the two tumor types. Consistently, the transcriptomes of CLL and PTCL were substantially different, sharing only 20% of overexpressed and 27% of underexpressed genes. We further show that hypomethylation correlates with increased gene expression in 8% genes in CLL and 17% genes in PTCL. In contrast,

hypermethylation correlated with gene underexpression in 1% and 4% in CLL and PTCL, respectively. Finally, we show that 21% of overexpressed and 11% of underexpressed genes are conserved between mouse and human CLL. The degree of conservation is higher between mouse and human PTCL with 39% of up-regulated and 38% of down-regulated shared between mouse and human PTCL. These data along with downregulation of DNMT3A in human CLL and mutations detected in human PTCLs suggest that *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ can serve as a new mouse model to study CLL and PTCL in mice.

Several interesting observations can be made from data presented in this study. First, our data suggest that despite the similar methylomes and transcriptomes of normal B-1a and CD8 cells, Dnmt3a-deficient T- and B-cell malignancies are characterized by highly cell-type specific changes in both methylation and transcription. We found that the extent of promoter hypomethylation is ~3-fold higher in malignant B-1a cells than in malignant CD8 cells relative to their normal counterparts. This promoter hypomethylation either results from a failure to maintain established methylation patterns during tumor progression due to a lack of Dnmt3a or from an increased proliferation of cancer cells resulting from transformation but not necessarily linked to Dnmt3a maintenance activity. The reasons for cancer-specific patterns evolving in the absence of Dnmt3a remains unclear. It is possible that methylation of the same loci is controlled by different Dnmts or their complexes in B-1a and CD8 cells.

Relative abundance of accessory factors such Dnmt3L or acquisition of genetic alterations affecting evolution of the epigenome may also differentially affect locus- specific methylation. It is also possible that some hypomethylated loci in tumors are already present in normal *Dnmt3a*$^{\Delta/\Delta}$ B-1a and CD8 cells as a result of the lack of *de novo* methylation performed by Dnmt3a during normal differentiation. Finally, the ability of B-1a to self-renew may also contribute to accelerated hypomethylation and might be a reason why B-1a cells are preferentially transformed over other

hematopoietic cell types in the absence of Dnmt3a. Cell-type specific methylation patterns between acute myeloid leukemia (AML), myelodysplastic syndrome (MDS), and CD4+CD8+ T cell acute lymphoblastic leukemia (T-ALL) that develop in *Mx-1-Cre;Dnmt3a*$^{fl/fl}$ mice has been previously reported. Mayle and colleagues observed that while methylation changes that arose in AML and MDS overlapped significantly, T-ALL was characterized by unique methylation patterns. Interestingly, the major methylation change observed in *Mx-1-Cre;Dnmt3a*$^{fl/fl}$ mice with T-ALL was hypermethylation, with approximately 75% of DMRS showing a gain in methylation in tumors relative to controls. This is in strong contrast to our data in which hypomethylation was the dominant event in both PTCL and CLL induced by loss of Dnmt3a [87].

A second important question raised by this study is the extent to which *Dnmt3a*$^{\Delta/\Delta}$ mouse CLL and PTCL models are similar to human disease. The development of CLL in the absence of Dnmt3a is consistent with findings that Dnmt3a belongs to the top 1% of underexpressed genes in human CLL [79, 154]. The development of PTCL in *Dnmt3a*$^{\Delta/\Delta}$ is consistent with the presence of mutations in DNMT3A in human T cell malignancies [23]. Our analysis of overexpressed and underexpressed genes in these hematologic malignancies shows a significant overlap with deregulated molecular events detected in human disease with 21% overexpressed and 11% underexpressed genes conserved between mouse and human CLL and even higher overlap with 39% overexpressed and 38% underexpressed conserved between mouse and human PTCL. Given that Dnmt3a is a methyltransferase whose direct consequence of inactivation would be expected to affect DNA methylation patterns we next focused on genes whose up-regulation was also accompanied by promoter hypomethylation. We identified 129 and 84 genes hypomethylated and overexpressed in CLL (HOC genes) and PTCL (HOT genes), respectively. 23 HOC genes were not only overexpressed in human CLL, with many of them also reported to be hypomethylated in human CLL [67]. Similarly, 26 HOT genes were overexpressed in human CD8+ PTCL but their methylation status have not been reported yet. Altogether, these data suggest that *EμSRα-tTA;Teto-*

Cre;Dnmt3a$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ model is a suitable model to study pathology of CLL and PTCL.

Since our data suggest that HOC and HOT genes might be targets of Dnmt3a it is also possible that their increased and untimely expression contribute to disease development. In CLL, *Zbtb32* (Zinc Finger and BTB Domain Containing 32 gene) was recently identified as a gene whose increased expression predicts whether patients without disease will develop CLL later in life [157]. In PTCL one of the candidate oncogenes is STAT1 (Signal transducer and activator of transcription 1), whose inappropriate activation has been observed in a variety of malignant cells from breast cancer, head and neck squamous carcinoma, melanoma, lymphoma and leukemia, suggesting that STAT1 may under specific conditions contribute to malignant transformation[23]. Such idea is further supported by findings that *Stat1*$^{-/-}$ mice are partially protected from *v-abl*-induced leukemia development, suggesting that Stat1 functions as an oncogene [130]. However, only future functional studies will carefully dissect the potential contribution of these genes to the development of CLL and PTCL.

Lastly, recent studies using *Mx1-Cre* transgene to conditionally delete Dnmt3a in HSPCs followed by transplantation into lethally irradiated recipients showed that the vast majority of mice (69%) develop myeloid disorders such as MDS and AML with rare occurrences of CD4+CD8+ double positive T-ALL or B-ALL [87, 89]. Interestingly, we have not observed the development of myeloid malignancies in *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ model, despite the fact that Dnmt3a - like in *Mx1-Cre;Dnmt3a*$^{fl/fl}$ model - is conditionally inactivated in HSPCs and all hematopoietic lineages. The differences in observed phenotypes could come from different experimental approaches in the sense that in *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ deletion of Dnmt3a occurs in HSPCs of primary transgenic mice and bone marrow transplant of Dnmt3a-null HSPC into lethally irradiated recipients is not employed. Alternatively, we used FVB rather than C57BL/6 mice to inactivate Dnmt3a. As mouse

strains differ in levels of gene expression, the Dnmt3a phenotypes may be strain specific. To address these possibilities directly we used *Dnmt3a*$^{\Delta/\Delta}$ bone marrow isolated from *EμSRα-tTA;Teto-Cre;Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ mice for adoptive transfer into lethally irradiated recipients. Interestingly, we observed the development of CLL with no signs of myeloproliferation, thus suggesting that Dnmt3a tumor suppressor function in prevention of CLL is autonomous to the hematopoietic system. This result also suggests that loss of Dnmt3a in the hematopoietic compartment may have strain-specific effects with FVB mice preferentially developing CLL and C57BL/6 mice preferentially developing myeloid malignancies. Further studies using *EμSRα-tTA;Teto-Cre; Dnmt3a*$^{fl/fl}$*;Rosa26LOXP*$^{EGFP/EGFP}$ in C57BL/6 mouse strain will further clarify this issue.

CHAPTER 5:

DISCUSSION & FUTURE DIRECTIONS

General Discussion

The studies presented here describe the molecular, cellular, and macro-phenotypic consequences of monoallelic and biallelic inactivation of *Dnmt3a* in murine hematopoietic cells. Here we describe that in FVB mice conditional ablation of both *Dnmt3a* alleles in the hematopoietic compartment results in the presentation of chronic lymphoid malignancies of B and T-cell lineage resembling CLL and PTCL. Furthermore, we demonstrate in conventional heterozygous *Dnmt3a*$^{+/-}$ mice that deficiency of only one *Dnmt3a* allele is sufficient to provoke the development of CLL and PTCL, however at reduced penetrance and with delayed onset as compared to *Dnmt3a* null mice. Analysis of tumors from *Dnmt3a*$^{+/-}$ mice revealed no LOH, mutation, deletion, or silencing of the second allele indicating that DNMT3A is a *bona fide* haploinsufficient tumor suppressor. One additionally important disease phenotype observed in our studies is the exclusive presentation in *Dnmt3a*$^{+/-}$ mice of a chronic myeloproliferative neoplasm at low penetrance [95]. Thus on a cellular level, we collectively demonstrate that *Dnmt3a* haploinsufficiency provokes hematologic malignancy of B-lymphoid, T-lymphoid, and myeloid lineage in mice.

The latter two disease phenotypes are congruent with human *DNMT3A* mutational spectra observed in T-cell and myeloid malignancy. *DNMT3A* mutations in acute T-cell leukemia (i.e. T-ALL) are almost always homozygous, whereas heterozygous mutations predominate in PTCL and occur exclusively in myeloid neoplasms [46]. In regards to *Dnmt3a* gene dosage, *Dnmt3a*$^{\Delta/\Delta}$ PTCL is more similar to that of human T-ALL than it is to human PTCL, whereas murine *Dnmt3a*$^{+/-}$ PTCL accurately reflects the gene dosage of its human counterpart. Interestingly, when we compared expression of genes deregulated in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCL to those genes deregulated in human PTCL, we discovered that *Dnmt3a*$^{+/-}$ PTCL shared a greater absolute number of overexpressed and underexpressed genes with human PTCL than did *Dnmt3a*$^{\Delta/\Delta}$ PTCL (**Fig 13**). Thus, similar to the parallels observed in *Dnmt3a* gene dosage, the gene expression signature of *Dnmt3a*$^{+/-}$ PTCL more accurately recaptures the transcriptional profile of the human malignancy

than does *Dnmt3a*$^{\Delta/\Delta}$ PTCL. Though it was not performed in these studies, the question now arises if *Dnmt3a*$^{\Delta/\Delta}$ PTCL (despite a post-thymic immunophenotype) may transcriptionally most closely resemble human T-ALL – the only human neoplasm of the T-lymphoid lineage currently known to harbor biallelic loss of function mutations in *DNMT3A*. The Goodell group observed in C57BL/6 mice homozygous deficient for *Dnmt3a* the development of T-ALL contingent upon the co-occurrence of activating mutation of *Notch1* or *Flt3* [87, 158]. Consequently, it is likely that these mutations do not spontaneously occur in our model and/or the FVB background of our mice may contain SNPs that can affect the phenotype of neoplasm that develops [135].

One serendipitous observation discovered in our studies on was the functional attenuation of p53 protein and the p53 pathway in murine *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ PTCLs [112]. Interestingly, reduced p53 protein was observed in 9 month old "pre-tumorigenic" thymocytes of *Dnmt3a*$^{+/-}$ mice, but not in 6 week old unaffected *Dnmt3a*$^{+/-}$ mice. Moreover, in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ tumors, *Trp53* RNA expression was not downregulated despite loss of p53 protein (**Fig 16 and 17**). It is thought that loss of p53 through deletion or mutation is rare in human PTCL compared to other malignancies. However, it was recently reported that mutation and chromosomal deletion of *TP53* frequently occurs in Sézary Syndrome and among pathway members regulating p53 function, PTCLs are known to harbor constellations of mutations in such genes as *CDKN2A/B, CDKN1B, TP63, WWOX,* and *ANKRD11* [159-162]. The latter of the two mechanisms is likely to have occurred as a second hit within developing thymocytes of *Dnmt3a*-deficient mice – in particular lesion of *CDKN2A/B*, which is affected in both immature T-ALL and mature PTCL neoplasms [159, 161, 163]. If lesion of *Cdkn2a/b* were to have occurred in *Dnmt3a*$^{\Delta/\Delta}$ mice, this may explain the incongruence with the lack of T-ALL development as in *DNMT3A* homozygous mutated human T-ALL, co-mutation of *CDKN2A/B* is negatively correlated [46]. Additionally, a potentially interesting consequence of attenuated p53 function in developing thymocytes is that this event has the potential to bypass or weaken the restrictions of the β-selection checkpoint at the

DN4/ISP stage and allows cells to proceed through to the DP stage of development due to lack of p53 [34, 35].

Another serendipitous observation in our studies was the discovery that full length (i.e. functional) DNMT3B protein expression is highly repressed or lost in *Dnmt3a*$^{\Delta/\Delta}$ PTCLs (**Fig 2**) [150]. Our laboratory first reported the development of CD8+ PTCL in *Dnmt3a*$^{\Delta/\Delta}$; *Dnmt3b*$^{\Delta/\Delta}$ double KO mice, and, until we expanded our studies with *Dnmt3a*$^{\Delta/\Delta}$ mice, we previously believed only this genotype of mice to be affected by PTCL. Previous work in our lab demonstrated that *Dnmt3b* is a *bona fide* tumor suppressor of murine T-lymphoid neoplasms [92, 93]. Likewise, it has just recently been discovered that in Sézary Syndrome *DNMT3B* is also mutated and is mutated as frequently as *DNMT3A* or *TET2* – the canonical epigenetic drivers of PTCL [162]. This implies that concomitant repression of DNMTB protein in *Dnmt3a*-deficient mice may constitute an additional driver / factor contributing to tumorigenesis. As *Dnmt3b* mutations have only thus far been observed in human PTCL, it is tempting to speculate that perhaps loss of *Dnmt3b* may somehow re-direct cellular development – alone or in conjunction with *Dnmt3a* – towards the T-cell lineage in a similar manner as to how *IDH2*^{R172} and *RHOA* mutations reportedly direct a *DNMT3A* / *TET2* mutated HSCs towards a T-cell fate [46] [REF 47]. Mouse phenotypic data generated from previously published and unpublished work suggests that this may in fact be the case, as informative crosses between *Dnmt3a* and *Dnmt3b* mice resulted in variable penetrance of T-lymphoid neoplasms based on the gene dosage of either gene (**Fig 39**). Mouse genotypes for T-lymphoid neoplasia in order of least penetrant to most penetrant are: *Dnmt3a*$^{+/-}$*;Dnmt3b*$^{+/+}$, *Dnmt3a*$^{\Delta/\Delta}$*;Dnmt3b*$^{+/+}$, *Dnmt3a*$^{+/+}$*;Dnmt3b*$^{+/-}$, *Dnmt3a*$^{\Delta/\Delta}$*;Dnmt3b*$^{\Delta/\Delta}$, and *Dnmt3a*$^{+/-}$*;Dnmt3b*$^{+/-}$. Interestingly, during the course of these studies, we observed one case of DN T-ALL that arose in a *Dnmt3a*$^{\Delta/\Delta}$*;Dnmt3b*$^{\Delta/\Delta}$ mouse (**Fig 40**) among other cases of DP T-ALL (also called "thymic lymphoma") and rare CD4+ lymphoma in mice deficient for *Dnmt3b* (**Fig 39**). This implies that concomitant loss of *Dnmt3b* in *Dnmt3a*-deficient mice may direct neoplasia towards a T-cell phenotype and expand the spectrum of disease observed. Consequently, the next best step is to

perform genomic sequencing of these lymphomas to reveal definitively any additional molecular drivers of murine CD8+ PTCL beyond Cre-Lox recombination-induced ablation of *Dnmt3a* and attenuation of p53 and DNMT3B. Likewise, as DNMT3A possesses a methylation-independent repressor function and is known to associate with chromatin modifiers such as HDAC1 and EZH2 [164, 165], chromatin immunoprecipitation experiments are warranted to examine the how changes in chromatin dynamics may be affected by loss of *Dnmt3a* in PTCL.

In addition to the aforementioned need to discover additional drivers of disease, three major questions remain in regards to this work on *Dnmt3a*-deficient PTCL. Firstly, there is the irksome question as to why mice uniformly develop CD8+ lymphomas in contrast to those of human PTCLs which are predominantly CD4+. Chapter 4 of this work partially addresses this question, but further study is needed. Secondly, the cell of origin of these tumors remains yet to be determined definitively. In these studies we compared *Dnmt3a*-deficient CD8+ PTCLs to peripheral CD8+ T-cells as this was the closest immunophenotypic control available to us, however we did not rigorously establish that this is the cell of origin for PTCL as opposed to a thymic progenitor or bone marrow HSPC. Future studies restricting cre-recombination of $Dnmt3a^{Flox/Flox}$ mice at various stages of T-cell development are needed to adequately address this question. Thirdly, a looming problem for this epigenetic study and the field in general is an incomplete understanding of the mechanism of DNMT3A's tumor suppressor function. Functional studies are needed to dissect the methyltransferase-dependent and independent activities of DNMT3A in tumorigenesis.

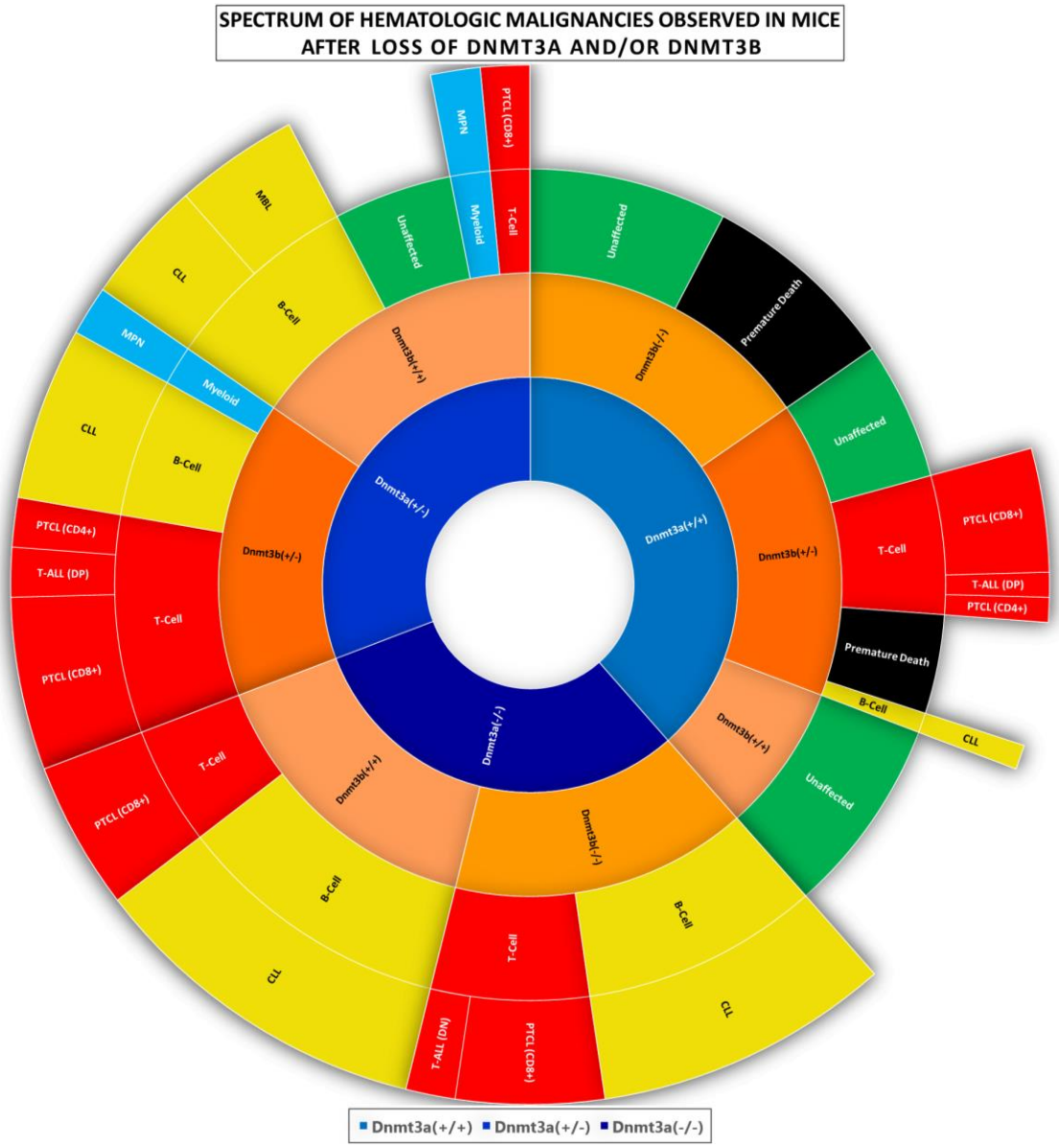


Figure 39. Spectrum of Hematologic Malignancies Observed in Mice After Loss of *Dnmt3a* and/or *Dnmt3b*. Sunburst diagram hierarchical clustering of hematologic malignancies observed in studies examining monoallelic or biallelic loss of *Dnmt3a* or *Dnmt3b* either alone or in combination with each other. Dark blue and orange of the inner two circles indicate allelic status of *Dnmt3a* and *Dnmt3b*, respectively, for experimental mice studied. Yellow, red, and light blue indicate observed neoplastic phenotypes of B lymphoid, T lymphoid, or myeloid origin, respectively, presenting in experimental mice. Green denotes unaffected mice surviving without a hematologic phenotype. Black indicates a premature death phenotype due to unknown causes observed in *Dnmt3b*$^{+/-}$ and *Dnmt3b*$^{-/-}$ mice.

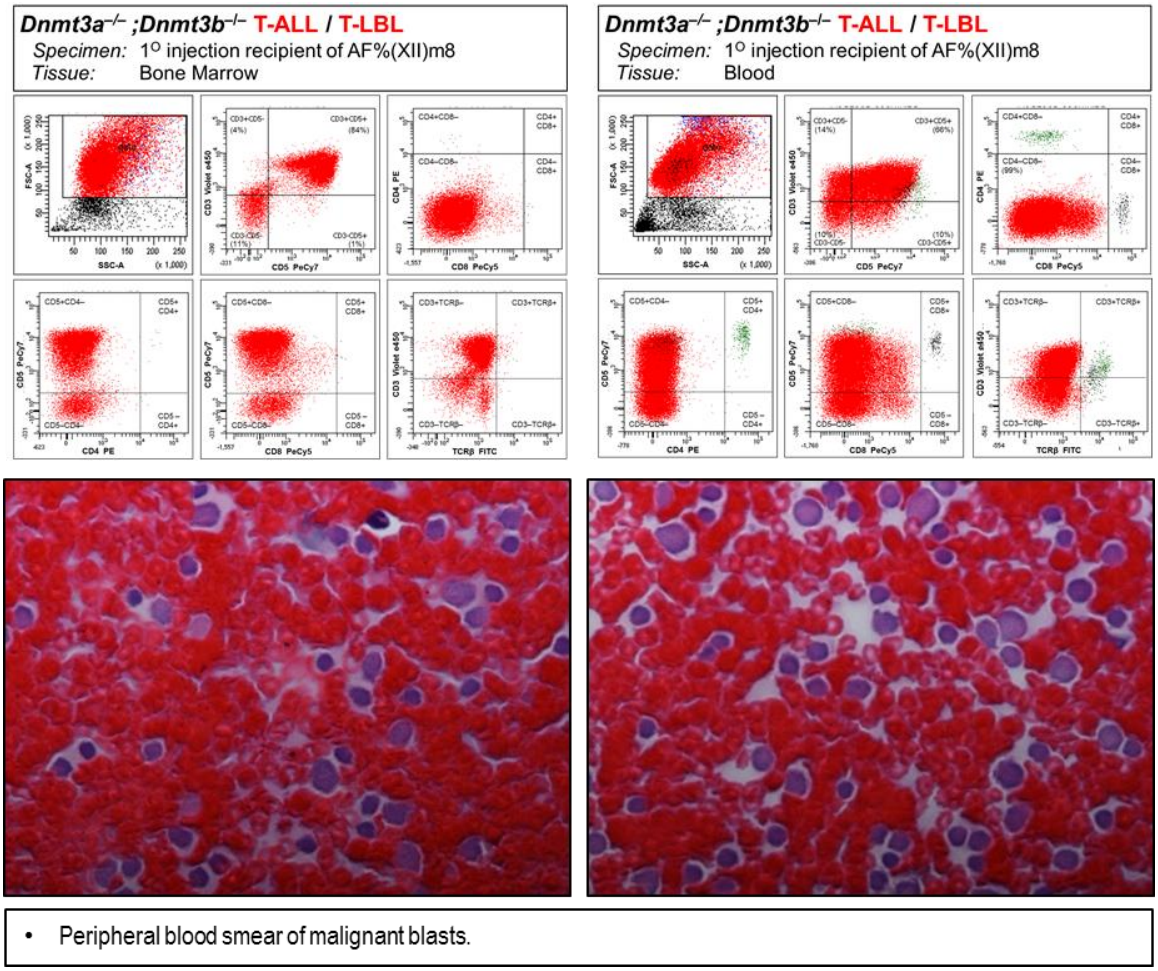


Figure 40. **A *Dnmt3a*$^{-/-}$*; Dnmt3b*$^{-/-}$ Mouse Presents with DN T-ALL / T-LBL.** Biallelic loss of the tumor suppressors *Dnmt3a* and *Dnmt3b* result in the rare presentation of double negative (CD4-CD8-) T-cell acute lymphoblastic leukemia / precursor T-cell lymphoblastic leukemia. Mouse presented with massive thymic enlargement, leukemia, bone marrow involvement, and lymphadenopathy. Leukemic precursor blasts are visible by peripheral blood smear (note: compare blasts to neutrophil or normal lymphocyte for reference to relative size of cells).

Murine *Dnmt3a*-Deficient PTCL: Cytotoxic vs Helper T-Cell Fate

Immunophenotypically however, murine CD8+ PTCL do not entirely reflect the majority of human mature T-cell lymphoma/leukemia as the neoplasms that present in the lymph nodes most frequently have a CD4+ immunophenotype [9, 161]. Unlike the tissue tropism observed in mice, the best classified CD8+ PTCL in humans have cutaneous involvement (e.g. primary cutaneous CD8+ aggressive epidermotropic cytotoxic T-cell lymphoma) and consequently do not accurately reflect the course of disease present in *Dnmt3a*-deficient mice [9]. There are however several caveats to potentially explain this discrepancy between the immunophenotypic presentation of mouse and human lymphoma.

Firstly, as cautionary note, no class of disease has been more hotly debated, split up, and re-classified more frequently than that of lymphomas – much less extremely rare forms of lymphoma such as PTCL. In fact approximately 50% of human PTCLs fall under the "wastebasket" classifier of PTCL-NOS as these tumors are not classifiable by current WHO definitions. In an international collaborative effort to pool samples and study these extremely rare human tumors, 372 PTCL-NOS tumors were analyzed by gene-expression profiling and re-classified based on *GATA3* and *TBX21* transcriptional signatures – transcription factors essential for T-cell polarization [166]. Importantly, this study identified a cytotoxic gene-signature within the *TBX21* subgroup that correlated with poor clinical outcome, indicating that though the majority of human PTCL-NOS may not express surface CD8, the transcriptome of these cells may however be more similar to a cytotoxic T-cell [166]. Consequently, *Dnmt3a*-deficient murine CD8+ PTCL may accurately model this subgroup of PTCL-NOS.

Secondly, in Chapter 4 of this work, we present data demonstrating that the methylomes of mature CD8+ cytotoxic T-cells are characterized by genome-wide hypermethylation of CpGs and extensive gene promoter hypermethylation (**Fig 22**) [150]. As more than 75% of all CpG dinucleotides in wild-type murine CD8+ T-cells are hypermethylated ($75\% < \beta \leq 100\%$), it is a

logical possibility that loss of *Dnmt3a* (or perhaps loss of methyltransferase activity from any DNMT) would have the largest phenotypic impact on a cell which under normal circumstances possesses an excess abundance of DNA methylation throughout its genome. Indeed, we observe that the epigenomes of both wild-type murine splenic B1a and CD8+ T-cells – the presumptive non-malignant counterparts of murine CLL and PTCL – are predominately characterized by genome-wide hypermethylation of CpGs and hypermethylation of promoters. This opens up the possibility that following ablation of *Dnmt3a*, subsequent loss of DNMT3A-associated maintenance methylation in cytotoxic T-cells results in genome-wide hypomethylation and eventually, transformation of CD8+ T-cells. However, this theory makes two crucial assumptions. It firstly presumes that DNMT3A's tumor suppressor activity is critically dependent on its methyltransferase activity and not a methyltransferase-independent function, such as association with chromatin modifiers like HDAC1 or EZH2. It secondly presumes that a mature CD8+ T-cell is the cell of origin for *Dnmt3a*-deficient PTCLs. Two fascinating alternative explanations exist if one or both of these assumptions do not prove to be true. In brief, if the first assumption is not met and methyltransferase activity is not essential for the tumor suppressive function of DNMT3A, then the rationale for suggesting that cells with extensively hypermethylated genomes become selectively transformed following loss of *Dnmt3a* is no longer valid. In this scenario, DNA hypomethylation observed in CD8+ PTCL would be a consequence of *Dnmt3a*-deficiency, but not a cause of disease. If the second assumption is not met and CD8+ T-cells are not the cell of origin of these neoplasms, then it is most likely that transformation occurs in an early T-lymphoid progenitor that retains the ability to mature into a cytotoxic T-cell (or simply express CD8). This explanation would be consistent with our discovery that p53 protein expression is attenuated in the thymuses of a 9 month old "pre-tumor" *Dnmt3a*$^{+/-}$ mice prior to the presentation of overt CD8+ PTCL. Interestingly, if the first assumption is confirmed, *in vivo* experiments determining the cell of origin of CD8+ PTCLs will also determine whether loss of DNMT3A maintenance methylation or *de novo* methylation is the mechanism of hypomethylation observed in tumors. If

hypomethylated CD8+ PTCLs develop in mice following ablation of *Dnmt3a* in mature CD8+ T-cells, then both assumptions are met and the original hypothesis is confirmed that loss of DNMT3A maintenance methylation is the mechanism of tumor genome-wide DNA hypomethylation. Alternatively, if inactivation of *Dnmt3a* in mature CD8+ T-cells does not result in hypomethylated CD8+ PTCL, but ablation of *Dnmt3a* in an immature progenitor cell does, it follows that lack of *de novo* methylation during thymopoiesis is responsible for the DNA hypomethylation observed in tumor genomes. Future studies must adequately address these two assumptions and will be discussed further in future directions.

Thirdly, it is also extremely interesting to note that the spectrum of lymphoid malignancies we observe in *Dnmt3a*-deficient mice are virtually identical to the neoplasms that develop in *Lck*^{pr}-*TCL1* and *Eμ-TCL1* transgenic mice: CD3+CD8+ T-cell leukemia/lymphoma and CD19+CD5+ CLL [79, 112, 167, 168]. *TCL1A* is a driver oncogene frequently found translocated to TCR loci in spontaneously arising neoplasms from ataxia telangiectasia patients such as T-cell prolymphocytic leukemia (T-PLL) and, less frequently, T-ALL [169, 170]. In addition to T-lymphoid neoplasms, *TCL1A* is highly over-expressed in B-lymphoid neoplasms such as CLL and its pre-leukemic precursor, MBL (although over-expression occurs by a mechanism independent of chromosomal translocation) [62, 68]. The redundant phenotypes observed in *Dnmt3a*-deficient and *TCL1* transgenic mice suggest both DNMT3A and TCL1 may function in a biologically related pathway in murine mature T and B-cell neoplasms. To this point, work in the *Eμ-TCL1* transgenic mouse model demonstrated that TCL1 can directly bind DNMT3A or DNMT3B, inhibit the methyltransferase activity of these enzymes, and induce DNA hypomethylation in pre-neoplastic and neoplastic B-cells [80]. This then begs the question if TCL1-mediated biochemical inhibition of DNMT3A/B is B-cell specific or common to both T and B-cell neoplasms. If TCL1 biochemical inhibition of DNMT3A/B also occurs in T-lymphoid neoplasms, then this may imply that the phenotypic spectrum of lymphoid disease observed in *Dnmt3a*-deficient mice is simply a reflection of that of TCL1 over-expression. However, as *Lck*^{pr}-*TCL1* mice were created with the intent to

model human T-PLL – a CD4+ malignancy – yet develop CD8+ leukemia/lymphoma in mice, this observation does not offer more than a correlative reason as to why *Dnmt3a*-deficient murine PTCL differ in immunophenotype from the majority of human PTCLs.

Lastly, cell surface CD8 expression may possibly be the biological consequence of loss of DNMT3A-mediated gene silencing in murine HSCs. Challen *et al* demonstrated by ChIP in murine HSCs that the *Runx1* (*AML1*) gene is directly bound by DNMT3A. Furthermore, in *Dnmt3a*$^{-/-}$ cells, the chromatin associated with the *Runx1* locus adopts a transcriptionally permissive state (H3K4me3) and consequently leads to *Runx1* over-expression [25]. Woolf *et al* reported that in mouse thymic medulla and cortex, *Runx3* and *Runx1* are highly expressed and function in the development of CD8+ T-cells during thymopoiesis. Similarly, CD2-restricted enforced expression of Runx2 results in an expanded CD8+ ISP population of thymocytes [171]. *Runx3*$^{-/-}$ mice displayed impaired CD8+ T-cell development in the thymus and anomalous expression of CD4 during T-cell lineage decisions. Compound *Runx3*$^{-/-}$;*Runx1*$^{+/-}$ mice displayed anomalous CD4 expression in all peripheral CD8+ T-cells of the spleen [172]. It was later discovered by Setoguchi *et al* that expression of the transcription factor Th-POK is necessary for murine CD4+CD8+ DP thymocytes to differentiate into CD4+ T-helper lineage cells. In the absence of Th-POK, DP cells will commit instead to the CD8+ cytotoxic T-cell lineage as a default fate [173]. The authors found that mouse cytotoxic T-cell differentiation is dependent on the active inhibition of Th-POK transcription by binding of a RUNX transcription factor complex to a silencer located within the Th-POK locus. Loss of RUNX transcription factor complex binding to the silencer sequence was able to restore Th-POK expression and CD4-lineage commitment [173]. Later studies further revealed that RUNX binds both the CD8 enhancer and CD4 silencer to activate and inhibit expression of these genes, respectively. To oppose RUNX, Th-POK can directly repress the CD8 locus and interrupt the physical interaction of RUNX with the CD4 silencer [38]. Consequently, we speculate that loss of DNMT3A-mediated gene repression may lead to the unopposed activity

of RUNX family transcription factors in developing DP thymocytes and bias T-cell differentiation towards a cytotoxic phenotype through transcriptional inhibition of Th-POK and CD4.

However, this mechanism would have to be reconciled with our previous *in vivo* observation that in young *Dnmt3a*-deficient mice (21 days old), loss of *Dnmt3a* did not affect development of hematopoietic lineages [79]. Two observations may reconcile these incongruences. Firstly, as *Dnmt3a*-deficient mice age, there is a slow steady expansion of CD8+ T-cells in peripheral blood, followed by inversion of the CD8:CD4 ratio of T-lymphocytes, and finally a leukemic phase of disease prior to presentation of T-cell neoplasms in the lymph nodes and spleen (Upchurch, unpublished observation). Secondly, *in vitro* tissue culture of the thymus from a 6 week old *Dnmt3a*-deficient mouse with recombinant IL-2 and anti-CD3 mAb stimulation resulted in a cellular population >90% CD3+CD8+CD4– within less than 48 hours (Upchurch, unpublished data). One interpretation of these data is that as HSCs or thymic progenitors proliferate in the absence of *Dnmt3a*, cells progressively become more hypomethylated due to lack of DNMT3A maintenance methylation until such time as the *Runx1/3* loci become hypomethylated in these cells and additional epimutations / lesions occur to synergize with unopposed RUNX (or other lineage biasing factor) in DP thymocytes. Almost serendipitously, we discovered RUNX1 (AML1) transcription factor binding site enrichment in the sequences of hypomethylated promoters in *Dnmt3a*-deficient PTCLs (**Fig 6F**), further implicating HOT genes in the pathogenesis of these lymphomas. Importantly, RUNX family transcription factors do not differ in the recognition of DNA consensus sequences but rather typically differ in their temporal-spatial expression. Consequently, expression of *Runx1/2/3* could be functionally redundant if aberrantly expressed. As such, one candidate epimutation that might synergize with RUNX could be the HOT gene *Il2rb*, which encodes the shared beta subunit of the IL-2 and IL-15 receptors. IL-2 is the primary growth factor for T-cells and IL-15, in conjunction with IL-7, specifies the cytotoxic lineage fate of CD8+ SP T-cells in the thymus [174]. Furthermore, IL-15 signaling through IL2Rb is essential for NKT and memory phenotype TCR-α/β CD8+ T-cells. Consequently IL2Rb signaling likely influences

other HOT genes, including *Klrc1*, *Klre1*, and *Klrd1*, which encode activating receptors of the CD94/NKG2 receptor family that are expressed in NK cells and activated cytotoxic T-cells [175]. As cytotoxic T-cells are known to express NK-cell activation receptors after persistent stimulation, the expression of these three genes may not necessarily imply that they bias development towards a cytotoxic T-cell lineage, but rather suggests that lymphomas are experiencing (or perhaps "perceive" that they are experiencing) chronic antigenic stimulation [30]. Interestingly, the developmental relatedness of cytotoxic T-cells and NK cells appears to be so close that they are biologically separated by the function of only one transcription factor: Bcl11b [30, 176]. This point is highlighted by the fact that KO of *Bcl11b* in murine T-cells induced T-cells to acquire NK-cell properties and function as NK cells [176]. Overall, this suggests that RUNX complex transcription factors may bias lineage determination and synergize with HOT genes to reinforce a cytotoxic lymphoma phenotype. Furthermore, the expression of NK-cell activation receptors may imply that these lymphomas are driven, in part, by chronic antigenic stimulation.

Consistent with this theory, using RNA-seq data we transcriptionally profiled wild-type CD8+ T-cells, 6 week old pre-malignant *Dnmt3a-/-* CD8 T-cells, and *Dnmt3a-/-* and +/- PTCLs collected from terminal mice and observe that T-cell lineage biasing transcription factors that favored a cytotoxic fate were over-expressed PTCLs relative to wild-type and pre-malignant CD8 control cells (**Fig 41a – d**). Though *Runx1* was not increased, these data show transcriptional upregulation of already highly expressed *Runx2/3* and *Cbfb* coincident with a massive absolute increase in *CD8a* expression over physiological baseline expression in mature CD8+ wild-type and pre-malignant *Dnmt3a-/-* T-cells. Consistent with the role of RUNX3 as a transcriptional antagonist of *Th-POK* and *CD4*, expression was uniformly low and virtually unchanged. Expression of *Bcl11b* was significantly increased in PTCL samples compared to controls and though this is a known tumor suppressor gene, it also acts to repress expression of *Th-POK*. *Th-Pok* transcriptional agonist *Myb* was significantly repressed whereas agonists *Tox*, and *Gata3* (that is positively regulated by MYB) were transcriptionally unchanged at a moderate-low level. Similar

to gene expression profiling performed in human CD4+ PTCL samples that found expression of the canonical transcription factors governing T-cell polarization correlated with the ontogeny of lymphomas [166], analysis of the canonical cytotoxic T-cell transcription factors *Stat1, Tbx21/Tbet,* and *Eomes* reveal massive over-expression of these genes (**Fig 41d**). Interestingly STAT1 is a transcriptional agonist of *Tbx21* and RUNX3 is an agonist of *Eomes.* *Stat1* is a HOT gene as reported in our results. Analysis of WGBS data revealed no overt changes in *Runx3* locus DNA methylation; whereas there was significant hypomethylation of the *Tbx21* locus in promoter and gene body regions and gene body hypomethylation in *Eomes.*

Unfortunately, this theory is limited by the fact that in these studies we only transcriptionally profiled mature CD8+ T-cells and CD8+ PTCL, but did not profile HSCs, early thymic progenitors, or DP thymocytes. At this time we lack expression data from the progenitor cells of wild-type and *Dnmt3a*-deficient mice that is necessary to reveal temporal dysregulation of *Runx1/3* expression (or that of any other lineage biasing factor) during thymopoiesis prior to the SP stage. Consequently, this theory remains speculative at this time until future experiments examine the transcriptomes and methylomes of these progenitor cells from both young and aged mice. These studies will determine if the observed transcriptional changes in lineage-specifying transcription factors are cause or consequence of a cytotoxic T-cell phenotype.

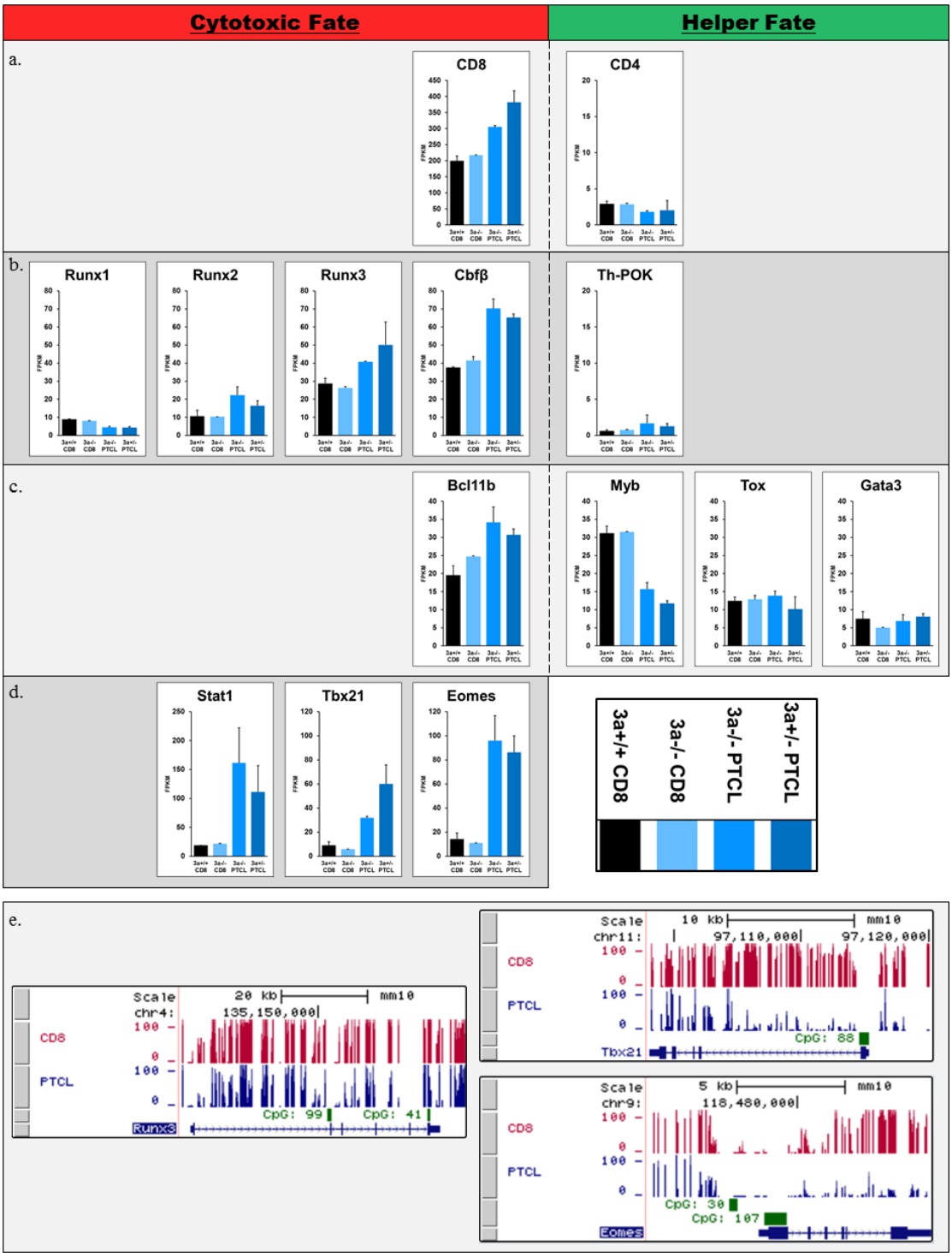


Figure 41. Cytotoxic T-Cell Lineage-Biasing Transcription Factors are Over-Expressed in *Dnmt3a*-deficient CD8+ PTCL. (a. – d.) RNA-seq data comparing transcription factors known to influence T-cell fate determination. Increasing expression values to the left of the dashed line favor CD8 SP T-cell development, whereas increasing expression values to the right of the dashed line favor CD4 SP T-cell development. (a.) T-cell co-receptors: CD8a vs. CD4. (b.) Transcriptional drivers of CD8a & CD4: RUNX family TFs enhance CD8a gene transcription which is opposed by Th-POK. (c.) Transcriptional agonists/antagonists of RUNX3 and Th-POK: (Left) BCL11B: agonist of *Runx3* / antagonist of *Th-POK*. (Right) MYB is an agonist of GATA3 which is an agonist of *Th-POK*. TOX: agonist of *Th-POK*. (d.) Core TF profile of cytotoxic T-cell: STAT1 is a transcriptional agonist of Tbx21/Tbet. (e.) WGBS data from *Dnmt3a*+/+ (CD8) control and *Dnmt3a-/- PTCL* showing DNA methylation at the mouse Runx3, Tbx21, and Eomes loci.

CD8+ PTCL: Testing the Cell of Origin

One of the most important questions in the field of cancer biology is that of a tumor's cell of origin. It known that in both myeloid and lymphoid neoplasia (as well as in other organ systems) that tumors often arise from transformation of a progenitor population rather than differentiated cells and can then manifest as a proliferation of mature cells [5, 15, 33, 46]. As it is known that many human lymphomas harboring activating mutations downstream of the TCR signaling pathway can allow these cells to bypass the β-checkpoint [33], it is necessary to assay at which point during T-cell development transformation occurs in *Dnmt3a*-deficient mice. Several lines of Cre mice are available from repositories such as Jackson Labs that can adequately test this question by breeding (*T-lineage Driver)-Cre;Dnmt3a*$^{FL/FL}$ mice. Based on this approach, one has only to exchange the "Driver Cre" and observe mice for the development of CD8+ PTCL. Tg(CD2-Cre) (Jax #008520) can be used to determine if the cell of origin resides downstream of lymphoid progenitor commitment to the T and B-cell lineage. Development of PTCL neoplasms from this stage may imply that loss of *Dnmt3a* can occur in early lymphoid progenitor cells. Tg(Lck-cre) (Jax #003802) under the control of the proximal Lck promoter is lowly expressed at the DN1-DN2 stages, but highly expressed at DN3 stages and thereafter in thymocytes and mature T-cells. If PTCL were to arise from crosses to this Cre, this would imply an ETP/DN thymocyte cell of origin. Tg(CD4-Cre) (Jax #017336) is expressed at the late DP stage of thymocyte development and can be used to determine if DP thymocytes are the cell of origin for CD8+ PTCL. IL2-Cre (Jax #029619) can be used to conditionally delete *Dnmt3a* in mature CD4+ T-cells, to see if these cells are the cell of origin. If this were the case, this would be a curious result and warrant further exploration of the mechanism of loss of CD4 expression and gain of CD8 expression. As no Cre is specific for post-thymic CD8+ T-cells, *Rosa26-Cre*^{ERT2}*;Dnmt3a*$^{FL/FL}$ mice can be bred, and CD8+ cells sorted, transplanted into recipient mice, and then afterwards induced with tamoxifen to ablate

Dnmt3a expression. However, this transplantation approach may be problematic based due to a limited number of cells available in mice. In addition to these experiments, *Dnmt3a*-deficient mice can also be crossed to Rag1-KO (Jax # 002216) mice to determine if TCR rearrangement is necessary or extraneous for lymphomagenesis. (Likewise, this Rag1-KO mice can be used to produce IGHV un-mutated CLL in *Dnmt3a*-deficient mice.)

The recently described role of *DNMT3A* mutations in LT-HSCs as an initiating lesion of CHIP and myeloid neoplasia has shed light on the cellular origins of these diseases [5, 22, 46, 158, 177]. Likewise in T-ALL the cell of origin is also believed to be a LT-HSC [46]. It was recently demonstrated in a mouse model of *Dnmt3a*-deficient T-ALL driven by NOTCH1 activation that the cell of origin was in fact an HSPC [178]. However, it remains to be determined if this is the case for PTCL. As is the case for angioimmunoblastic T-cell lymphoma (AITL), evidence suggests that these neoplasms arise from transformation of follicular T helper cells [41]. The possibility also exists, and may rather be likely, that *DNMT3A* inactivation must occur in self-renewing progenitor populations and only later during developmental maturation are secondary lesions acquired in thymocytes that fully transform cells.

Mechanism of DNMT3A Tumor Suppressor Function

DNMT3A is predominantly known by its *de facto* title a "*de novo* DNA methyltransferase" despite the fact that it also possesses a maintenance methylation function and methylation-independent repressor activity. It is precisely the dissection of these individual activities that is necessary for determining with precision the mechanistic actions of DNMT3A in hematopoiesis and hematologic malignancy. In order to determine the methyltransferase-dependent and methyltransferase-independent tumor suppressor functions of DNMT3A, we are creating novel knock-in mice harboring a point mutation in the catalytic cysteine of DNMT3A that ablates all intrinsic catalytically activity without impairing other methyltransferase-independent enzymatic functions or potentially gaining functions (as may be the case with the R882 mutation) [102, 179]. These *Dnmt3a*-KI mice will be used first to determine if they develop malignancy and these malignancies fully recapture the phenotypic spectrum of disease present in *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ mice. If this result occurs, this will conclusively demonstrate *in vivo* that methyltransferase activity is essential for DNMT3A's tumor suppressor function. This result is most likely and is consistent with the hypothesis presented in our studies that hypomethylation of promoters induced by inactivation of *Dnmt3a* is a driver of lymphoid neoplasms. To further functionally test the role of hypomethylation – specifically promoter hypomethylation – in the transformation of murine cells, we have created a library of lentiviruses containing the top 5 most promising candidate HOT genes discovered from our studies in Chapter 3 (**Fig 10**). HSCs from 6 week-old *Dnmt3a*$^{+/+}$, *Dnmt3a*$^{+/-}$, and *Dnmt3a*$^{\Delta/\Delta}$ mice will be transduced with pooled HOT gene lentivirus or lentiviral vector control and subsequently injected into lethally irradiated syngeneic recipient mice where they will be observed for the development of PTCL. Comparison of HOT gene transduced WT HSC mice with those of *Dnmt3a*$^{+/-}$ and *Dnmt3a*$^{\Delta/\Delta}$ mice will determine if HOT genes are in and of themselves sufficient to provoke the development of PTCL or if other

uncontrolled events engendered by loss of *Dnmt3a* (e.g. genomic instability) is necessary in addition to HOT gene over-expression to sufficiently transform cells. Collectively these studies would show that methyltransferase activity is essential for tumor suppression and DNMT3A specifically functions to repress oncogene expression through promoter methylation.

In contrast, if no tumors were to develop in *Dnmt3a*-KI mice, this result would instead argue that loss of methyltransferase activity is not sufficient to predispose cells to transformation, but rather DNMT3A's methyltransferase-independent function is the crucial component necessary for DNMT3A's tumor suppressor function. This result would not be too surprising as EZH2, the chromatin modifier and known associate of DNMT3A, is documented to harbor recurrent loss of function mutations in human neoplasms of the T-lymphoid lineage as well as in MDS [24, 169, 180-182] – both diseases canonically affected by loss of *Dnmt3a* – and partially relies on DNMT family members to epigenetically silence chromatin, implying a certain degree of functional redundancy between these two proteins [165]. Consequently, chromatin immunoprecipitation studies would be warranted in *Dnmt3a*-deficient PTCL samples compared to *Dnmt3a*-KI and WT cells.

The need to determination of the essential mechanism of DNMT3A's tumor suppressor function is currently a fundamental problem in the field of cancer epigenetics and its resolution promises to be of both academic and clinical utility.

Conclusions

In the work presented in this dissertation, we describe the molecular, cellular, and macro-phenotypic consequences of monoallelic and biallelic inactivation of *Dnmt3a* in murine hematopoietic cells. On a macro-phenotypic level, we demonstrate that loss of one or two *Dnmt3a* alleles in murine HSPCs is able to provoke hematologic malignancy if allowed sufficient latency for disease to develop in mice. On a cellular level, we observe both common and distinct hematologic presentations that are *Dnmt3a* gene dosage-dependent. We demonstrate that ablation of two *Dnmt3a* alleles in murine HSPCs results in the presentation of CLL or PTCL. Likewise, we demonstrate that ablation of only one *Dnmt3a* allele in mice also results in the presentation of CLL and PTCL – albeit at reduced penetrance – in addition to the presentation of a MPN. In our previous studies published on *Dnmt3a*-deficient CLL and in the work presented in this dissertation on *Dnmt3a*-deficient PTCL, we demonstrate on the molecular level that tumors from $Dnmt3a^{+/-}$ mice do not undergo mutation, deletion, LOH, or silencing of the second *Dnmt3a* allele [95]. Furthermore, as we profiled both $Dnmt3a^{+/-}$ CLL and $Dnmt3a^{\Delta/\Delta}$ CLL cells by whole genome bisulfite sequencing, we show that there exists a DNMT3A dose-dependent decrease in the CLL genome's DNA methylation, indicating that in heterozygous tumors the residual DNMT3A enzyme expressed from the remaining wild type allele is functional and not inhibited [95]. Consequently, we collectively demonstrate that *Dnmt3a* is a haploinsufficient tumor suppressor whose loss provokes hematologic malignancy of B-lymphoid, T-lymphoid, and myeloid lineage in mice.

This work provides *in vivo* evidence that loss of *Dnmt3a* – or perhaps more generally, loss of DNA methylation – can be an initiating step on the road eventually leading to cellular transformation. This concept is in agreement with the recently described condition called clonal hematopoiesis, in which pre-neoplastic mutations occur in LT-HSCs – of which, *DNMT3A* is the most frequently occurring [6]. As *DNMT3A* mutations are often found at the highest allelic frequency in human T-lymphoid and myeloid malignancies, the current model of CHIP as a

precursor to neoplasms of these two lineages is well supported by observations in humans as well as in our mouse model. Unlike myeloid and T-lymphoid neoplasms, spontaneous *DNMT3A* mutations have not been documented in human CLL or any other human neoplasms of the B-lymphoid lineage. In spite of the lack of *DNMT3A* mutation, epigentic data from human CLL samples reveals that DNA hypomethylation is a major molecular phenotype of this disease [67, 68]. Interestingly, the *TCL1A* oncogene – which encodes for a protein that is direct biochemical inhibitor of DNMT3A and DNMT3B methyltransferase activity – is one of the most highly expressed genes in human CLL beginning from the MBL stage of disease development [62, 68, 80]. The transgenic *Eμ-Tcl1* mouse model demonstrates the sufficiency of the *Tcl1* oncogene to induce a CLL-like disease in mice. Consequently, the fact that *Dnmt3a*-deficient mice present with CLL likely indicates that inhibition of DNMT3A methyltransferase activity is an essential (if not causal) mechanism of TCL1A-induced oncogenesis in murine, and perhaps, human CLL.

In regards to *Dnmt3a*-deficient PTCL, we present in this work two additional novel findings that are likely driver events: attenuation of p53 and full-length (functional) DNMT3B proteins. Interestingly, it has just recently been discovered that lesion of *DNMT3B, TP53,* and *TP53* related regulatory pathway genes frequently occur in Sézary Syndrome [159-162]. Our findings in mice and the recent genomic profiling in this one type of T-cell lymphoma/leukemia suggest that these genes might be affected in T-cell neoplasms more commonly than previously appreciated. Additionally, one other interesting observation is that DNMT3B inactivation appears to correlate with a T-lymphoid phenotype, suggesting not only that further epigenetic dysregulation occurs, but also that loss of *Dnmt3b* may influence the lineage of the resulting neoplasm.

These findings are relevant to human disease and demonstrate that epigenetic dysregulation of DNA methylation is fundamental to the pathogenesis of hematologic malignancies stemming from any one branch of the hematopoietic tree.

BIBLIOGRAPHY

[1-20, 22, 23, 25-40, 42, 46, 51, 58, 63, 67, 79, 80, 82-99, 102-149, 151-157, 161, 166-168, 172-176, 183-185]

1. Yan, H., et al., *IDH1 and IDH2 mutations in gliomas.* N Engl J Med, 2009. **360**(8): p. 765-73.

2. Versteege, I., et al., *Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer.* Nature, 1998. **394**(6689): p. 203-6.

3. Feinberg, A.P., M.A. Koldobskiy, and A. Gondor, *Epigenetic modulators, modifiers and mediators in cancer aetiology and progression.* Nat Rev Genet, 2016. **17**(5): p. 284-99.

4. Cazzola, M., M.G. Della Porta, and L. Malcovati, *The genetic basis of myelodysplasia and its clinical relevance.* Blood, 2013. **122**(25): p. 4021-34.

5. Shlush, L.I., et al., *Identification of pre-leukaemic haematopoietic stem cells in acute leukaemia.* Nature, 2014. **506**(7488): p. 328-33.

6. Xie, M., et al., *Age-related mutations associated with clonal hematopoietic expansion and malignancies.* Nat Med, 2014. **20**(12): p. 1472-8.

7. Flajnik, M.F. and M. Kasahara, *Origin and evolution of the adaptive immune system: genetic events and selective pressures.* Nat Rev Genet, 2010. **11**(1): p. 47-59.

8. Arber, D.A., et al., *The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia.* Blood, 2016. **127**(20): p. 2391-405.

9. Swerdlow, S.H., et al., *The 2016 revision of the World Health Organization classification of lymphoid neoplasms.* Blood, 2016. **127**(20): p. 2375-90.

10. Ribatti, D. and E. Crivellato, *Mast cell ontogeny: an historical overview.* Immunol Lett, 2014. **159**(1-2): p. 11-4.

11. Rosenbauer, F. and D.G. Tenen, *Transcription factors in myeloid development: balancing differentiation with transformation.* Nat Rev Immunol, 2007. **7**(2): p. 105-17.

12. Pabst, T., et al., *Dominant-negative mutations of CEBPA, encoding CCAAT/enhancer binding protein-alpha (C/EBPalpha), in acute myeloid leukemia.* Nat Genet, 2001. **27**(3): p. 263-70.

13. Fasan, A., et al., *The role of different genetic subtypes of CEBPA mutated AML.* Leukemia, 2014. **28**(4): p. 794-803.

14. Vainchenker, W. and R. Kralovics, *Genetic basis and molecular pathophysiology of classical myeloproliferative neoplasms.* Blood, 2017. **129**(6): p. 667-679.

15. Mead, A.J. and A. Mullally, *Myeloproliferative neoplasm stem cells.* Blood, 2017. **129**(12): p. 1607-1616.

16. Reilly, J.T., et al., *Guideline for the diagnosis and management of myelofibrosis.* Br J Haematol, 2012. **158**(4): p. 453-71.

17. Kelly, L.M. and D.G. Gilliland, *Genetics of myeloid leukemias.* Annu Rev Genomics Hum Genet, 2002. **3**: p. 179-98.

18. Nolte, F. and W.K. Hofmann, *Molecular mechanisms involved in the progression of myelodysplastic syndrome.* Future Oncol, 2010. **6**(3): p. 445-55.

19. Perrotti, D., et al., *Chronic myeloid leukemia: mechanisms of blastic transformation.* J Clin Invest, 2010. **120**(7): p. 2254-64.

20. Zhang, S.J., et al., *Gain-of-function mutation of GATA-2 in acute myeloid transformation of chronic myeloid leukemia.* Proc Natl Acad Sci U S A, 2008. **105**(6): p. 2076-81.

21. Shih, A.H., et al., *The role of mutations in epigenetic regulators in myeloid malignancies.* Nat Rev Cancer, 2012. **12**(9): p. 599-612.

22. Ley, T.J., et al., *DNMT3A mutations in acute myeloid leukemia.* N Engl J Med, 2010. **363**(25): p. 2424-33.

23. Couronne, L., C. Bastard, and O.A. Bernard, *TET2 and DNMT3A mutations in human T-cell lymphoma.* N Engl J Med, 2012. **366**(1): p. 95-6.

24. Ganguly, B.B. and N.N. Kadam, *Mutations of myelodysplastic syndromes (MDS): An update.* Mutat Res Rev Mutat Res, 2016. **769**: p. 47-62.

25. Challen, G.A., et al., *Dnmt3a is essential for hematopoietic stem cell differentiation.* Nat Genet, 2011. **44**(1): p. 23-31.

26. Challen, G.A., et al., *Dnmt3a and Dnmt3b have overlapping and distinct functions in hematopoietic stem cells.* Cell Stem Cell, 2014. **15**(3): p. 350-64.

27. DuPage, M. and J.A. Bluestone, *Harnessing the plasticity of CD4(+) T cells to treat immune-mediated disease.* Nat Rev Immunol, 2016. **16**(3): p. 149-63.

28. Fisher, S.G. and R.I. Fisher, *The epidemiology of non-Hodgkin's lymphoma.* Oncogene, 2004. **23**(38): p. 6524-34.

29. Aifantis, I., E. Raetz, and S. Buonamici, *Molecular pathogenesis of T-cell leukaemia and lymphoma.* Nat Rev Immunol, 2008. **8**(5): p. 380-90.

30. Paul, W.E., *Fundamental immunology*. 7th ed. 2013, Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins. xviii, 1283 p.

31. Staal, F.J., et al., *The functional relationship between hematopoietic stem cells and developing T lymphocytes.* Ann N Y Acad Sci, 2016. **1370**(1): p. 36-44.

32. Paul, W.E., *Fundamental immunology*. 6th ed. 2008, Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins. xviii, 1603 p., 16 p. of plates.

33. Schmedt, C. and A. Tarakhovsky, *Autonomous maturation of alpha/beta T lineage cells in the absence of COOH-terminal Src kinase (Csk).* J Exp Med, 2001. **193**(7): p. 815-26.

34. Okazuka, K., et al., *p53 prevents maturation of T cell development to the immature CD4-CD8+ stage in Bcl11b-/- mice.* Biochem Biophys Res Commun, 2005. **328**(2): p. 545-9.

35. Jiang, D., M.J. Lenardo, and J.C. Zuniga-Pflucker, *p53 prevents maturation to the CD4+CD8+ stage of thymocyte differentiation in the absence of T cell receptor rearrangement.* J Exp Med, 1996. **183**(4): p. 1923-8.

36. Conroy, L.A. and D.R. Alexander, *The role of intracellular signalling pathways regulating thymocyte and leukemic T cell apoptosis.* Leukemia, 1996. **10**(9): p. 1422-35.

37. Punt, J.A., et al., *Negative selection of CD4+CD8+ thymocytes by T cell receptor-induced apoptosis requires a costimulatory signal that can be provided by CD28.* J Exp Med, 1994. **179**(2): p. 709-13.

38. Taniuchi, I., *Views on helper/cytotoxic lineage choice from a bottom-up approach.* Immunol Rev, 2016. **271**(1): p. 98-113.

39. Collins, A., D.R. Littman, and I. Taniuchi, *RUNX proteins in transcription factor networks that regulate T-cell lineage choice.* Nat Rev Immunol, 2009. **9**(2): p. 106-15.

40. Wang, L. and R. Bosselut, *CD4-CD8 lineage differentiation: Thpok-ing into the nucleus.* J Immunol, 2009. **183**(5): p. 2903-10.

41. Cortes, J.R. and T. Palomero, *The curious origins of angioimmunoblastic T-cell lymphoma.* Curr Opin Hematol, 2016. **23**(4): p. 434-43.

42. Wang, C., et al., *IDH2R172 mutations define a unique subgroup of patients with angioimmunoblastic T-cell lymphoma.* Blood, 2015. **126**(15): p. 1741-52.

43. Rohr, J., et al., *Recurrent activating mutations of CD28 in peripheral T-cell lymphomas.* Leukemia, 2016. **30**(5): p. 1062-70.

44. Palomero, T., et al., *Recurrent mutations in epigenetic regulators, RHOA and FYN kinase in peripheral T cell lymphomas.* Nat Genet, 2014. **46**(2): p. 166-70.

45. Wang, M., et al., *Angioimmunoblastic T cell lymphoma: novel molecular insights by mutation profiling.* Oncotarget, 2017. **8**(11): p. 17763-17770.

46. Yang, L., R. Rau, and M.A. Goodell, *DNMT3A in haematological malignancies.* Nat Rev Cancer, 2015. **15**(3): p. 152-65.

47. Lemonnier, F., et al., *The IDH2 R172K mutation associated with angioimmunoblastic T-cell lymphoma produces 2HG in T cells and impacts lymphoid development.* Proc Natl Acad Sci U S A, 2016. **113**(52): p. 15084-15089.

48. Xu, W., et al., *Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases.* Cancer Cell, 2011. **19**(1): p. 17-30.

49. Pileri, S.A., *Follicular helper T-cell-related lymphomas.* Blood, 2015. **126**(15): p. 1733-4.

50. Conway O'Brien, E., S. Prideaux, and T. Chevassut, *The epigenetic landscape of acute myeloid leukemia.* Adv Hematol, 2014. **2014**: p. 103175.

51. Jeong, M., et al., *Large conserved domains of low DNA methylation maintained by Dnmt3a.* Nat Genet, 2014. **46**(1): p. 17-23.

52. Mourad, R. and O. Cuvier, *Predicting the spatial organization of chromosomes using epigenetic data.* Genome Biol, 2015. **16**: p. 182.

53. Fortin, J.P. and K.D. Hansen, *Reconstructing A/B compartments as revealed by Hi-C using long-range correlations in epigenetic data.* Genome Biol, 2015. **16**: p. 180.

54. Lunning, M.A. and J.M. Vose, *Angioimmunoblastic T-cell lymphoma: the many-faced lymphoma.* Blood, 2017. **129**(9): p. 1095-1102.

55. Vallois, D., et al., *Activating mutations in genes related to TCR signaling in angioimmunoblastic and other follicular helper T-cell-derived lymphomas.* Blood, 2016. **128**(11): p. 1490-502.

56. Fernandez-Piqueras, J., *New mutations for nodal lymphomas of TFH origin.* Blood, 2016. **128**(11): p. 1446-7.

57. Yu, E.M., A. Kittai, and I.A. Tabbara, *Chronic Lymphocytic Leukemia: Current Concepts.* Anticancer Res, 2015. **35**(10): p. 5149-65.

58. Fabbri, G. and R. Dalla-Favera, *The molecular pathogenesis of chronic lymphocytic leukaemia.* Nat Rev Cancer, 2016. **16**(3): p. 145-62.

59. Nieto, W.G., et al., *Increased frequency (12%) of circulating chronic lymphocytic leukemia-like B-cell clones in healthy subjects using a highly sensitive multicolor flow cytometry approach.* Blood, 2009. **114**(1): p. 33-7.

60. Landgren, O., et al., *B-cell clones as early markers for chronic lymphocytic leukemia.* N Engl J Med, 2009. **360**(7): p. 659-67.

61. Karube, K., et al., *Monoclonal B cell lymphocytosis and "in situ" lymphoma.* Semin Cancer Biol, 2014. **24**: p. 3-14.

62. Lanasa, M.C., et al., *Immunophenotypic and gene expression analysis of monoclonal B-cell lymphocytosis shows biologic characteristics associated with good prognosis CLL.* Leukemia, 2011. **25**(9): p. 1459-66.

63. Strati, P. and T.D. Shanafelt, *Monoclonal B-cell lymphocytosis and early-stage chronic lymphocytic leukemia: diagnosis, natural history, and risk stratification.* Blood, 2015. **126**(4): p. 454-62.

64. Xochelli, A., D. Oscier, and K. Stamatopoulos, *Clonal B-cell lymphocytosis of marginal zone origin.* Best Pract Res Clin Haematol, 2017. **30**(1-2): p. 77-83.

65. Guieze, R. and C.J. Wu, *Genomic and epigenomic heterogeneity in chronic lymphocytic leukemia.* Blood, 2015. **126**(4): p. 445-53.

66. Landau, D.A., et al., *Mutations driving CLL and their evolution in progression and relapse.* Nature, 2015. **526**(7574): p. 525-30.

67. Kulis, M., et al., *Epigenomic analysis detects widespread gene-body DNA hypomethylation in chronic lymphocytic leukemia.* Nat Genet, 2012. **44**(11): p. 1236-42.

68. Oakes, C.C., et al., *DNA methylation dynamics during B cell maturation underlie a continuum of disease phenotypes in chronic lymphocytic leukemia.* Nat Genet, 2016. **48**(3): p. 253-64.

69. Wang, L., et al., *Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia.* Cancer Cell, 2016. **30**(5): p. 750-763.

70. Barrio, S., et al., *Genomic characterization of high-count MBL cases indicates that early detection of driver mutations and subclonal expansion are predictors of adverse clinical outcome.* Leukemia, 2017. **31**(1): p. 170-176.

71. Puiggros, A., G. Blanco, and B. Espinet, *Genetic abnormalities in chronic lymphocytic leukemia: where we are and where we go.* Biomed Res Int, 2014. **2014**: p. 435983.

72. Malek, S., *Advances in chronic lymphocytic leukemia*. Advances in experimental medicine and biology,. 2013, New York: Springer. xi, 334 pages.

73. Pekarsky, Y., N. Zanesi, and C.M. Croce, *Molecular basis of CLL.* Semin Cancer Biol, 2010. **20**(6): p. 370-6.

74. Palamarchuk, A., et al., *13q14 deletions in CLL involve cooperating tumor suppressors.* Blood, 2010. **115**(19): p. 3916-22.

75. Darman, R.B., et al., *Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point.* Cell Rep, 2015. **13**(5): p. 1033-45.

76. Alsafadi, S., et al., *Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage.* Nat Commun, 2016. **7**: p. 10615.

77. Jeelall, Y.S. and K. Horikawa, *Oncogenic MYD88 mutation drives Toll pathway to lymphoma.* Immunol Cell Biol, 2011. **89**(6): p. 659-60.

78. Wang, L., et al., *Somatic mutation as a mechanism of Wnt/beta-catenin pathway activation in CLL.* Blood, 2014. **124**(7): p. 1089-98.

79. Peters, S.L., et al., *Tumor suppressor functions of Dnmt3a and Dnmt3b in the prevention of malignant mouse lymphopoiesis.* Leukemia, 2014. **28**(5): p. 1138-42.

80. Palamarchuk, A., et al., *Tcl1 protein functions as an inhibitor of de novo DNA methylation in B-cell chronic lymphocytic leukemia (CLL).* Proc Natl Acad Sci U S A, 2012. **109**(7): p. 2555-60.

81. Upchurch, G.M., S.L. Haney, and R. Opavsky, *Aberrant Promoter Hypomethylation in CLL: Does It Matter for Disease Development?* Front Oncol, 2016. **6**: p. 182.

82. McCabe, M.T., J.C. Brandes, and P.M. Vertino, *Cancer DNA methylation: molecular mechanisms and clinical implications.* Clin Cancer Res, 2009. **15**(12): p. 3927-37.

83. Celik, H., A. Kramer, and G.A. Challen, *DNA methylation in normal and malignant hematopoiesis.* Int J Hematol, 2016. **103**(6): p. 617-26.

84. Herman, J.G. and S.B. Baylin, *Gene silencing in cancer in association with promoter hypermethylation.* N Engl J Med, 2003. **349**(21): p. 2042-54.

85. Okano, M., et al., *DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development.* Cell, 1999. **99**(3): p. 247-57.

86. Robert, M.F., et al., *DNMT1 is required to maintain CpG methylation and aberrant gene silencing in human cancer cells.* Nat Genet, 2003. **33**(1): p. 61-5.

87. Mayle, A., et al., *Dnmt3a loss predisposes murine hematopoietic stem cells to malignant transformation.* Blood, 2015. **125**(4): p. 629-38.

88. Guryanova, O.A., et al., *Dnmt3a regulates myeloproliferation and liver-specific expansion of hematopoietic stem and progenitor cells.* Leukemia, 2016. **30**(5): p. 1133-42.

89. Celik, H., et al., *Enforced differentiation of Dnmt3a-null bone marrow leads to failure with c-Kit mutations driving leukemic transformation.* Blood, 2015. **125**(4): p. 619-28.

90. Chang, Y.I., et al., *Loss of Dnmt3a and endogenous Kras(G12D/+) cooperate to regulate hematopoietic stem and progenitor cell functions in leukemogenesis.* Leukemia, 2015. **29**(9): p. 1847-56.

91. Gao, X.N., et al., *AML1/ETO cooperates with HIF1alpha to promote leukemogenesis through DNMT3a transactivation.* Leukemia, 2015. **29**(8): p. 1730-40.

92. Haney, S.L., et al., *Methylation-independent repression of Dnmt3b contributes to oncogenic activity of Dnmt3a in mouse MYC-induced T-cell lymphomagenesis.* Oncogene, 2015. **34**(43): p. 5436-46.

93. Hlady, R.A., et al., *Loss of Dnmt3b function upregulates the tumor modifier Ment and accelerates mouse lymphomagenesis.* J Clin Invest, 2012. **122**(1): p. 163-77.

94. Vasanthakumar, A., et al., *Dnmt3b is a haploinsufficient tumor suppressor gene in Myc-induced lymphomagenesis.* Blood, 2013. **121**(11): p. 2059-63.

95. Haney, S.L., et al., *Promoter Hypomethylation and Expression Is Conserved in Mouse Chronic Lymphocytic Leukemia Induced by Decreased or Inactivated Dnmt3a.* Cell Rep, 2016. **15**(6): p. 1190-201.

96. Peters, S.L., et al., *Essential role for Dnmt1 in the prevention and maintenance of MYC-induced T-cell lymphomas.* Mol Cell Biol, 2013. **33**(21): p. 4321-33.

97. Trowbridge, J.J., et al., *Haploinsufficiency of Dnmt1 impairs leukemia stem cell function through derepression of bivalent chromatin domains.* Genes Dev, 2012. **26**(4): p. 344-9.

98. Broske, A.M., et al., *DNA methylation protects hematopoietic stem cell multipotency from myeloerythroid restriction.* Nat Genet, 2009. **41**(11): p. 1207-15.

99. Gaudet, F., et al., *Induction of tumors in mice by genomic hypomethylation.* Science, 2003. **300**(5618): p. 489-92.

100. Russler-Germain, D.A., et al., *The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers.* Cancer Cell, 2014. **25**(4): p. 442-54.

101. Grossmann, V., et al., *The molecular profile of adult T-cell acute lymphoblastic leukemia: mutations in RUNX1 and DNMT3A are associated with poor prognosis in T-ALL.* Genes Chromosomes Cancer, 2013. **52**(4): p. 410-22.

102. Guryanova, O.A., et al., *DNMT3A mutations promote anthracycline resistance in acute myeloid leukemia via impaired nucleosome remodeling.* Nat Med, 2016. **22**(12): p. 1488-1495.

103. Tseng, Y.Y., et al., *PVT1 dependence in cancer with MYC copy-number increase.* Nature, 2014. **512**(7512): p. 82-6.

104. Ansell, S.M., et al., *PD-1 blockade with nivolumab in relapsed or refractory Hodgkin's lymphoma.* N Engl J Med, 2015. **372**(4): p. 311-9.

105. McClanahan, F., et al., *PD-L1 checkpoint blockade prevents immune dysfunction and leukemia development in a mouse model of chronic lymphocytic leukemia.* Blood, 2015. **126**(2): p. 203-11.

106. Ding, W., et al., *Pembrolizumab in patients with chronic lymphocytic leukemia with Richter's transformation and relapsed CLL.* Blood, 2017.

107. Martinez-Trillos, A., et al., *Mutations in TLR/MYD88 pathway identify a subset of young chronic lymphocytic leukemia patients with favorable outcome.* Blood, 2014. **123**(24): p. 3790-6.

108. Mackay, F. and P. Schneider, *Cracking the BAFF code.* Nat Rev Immunol, 2009. **9**(7): p. 491-502.

109. Planelles, L., et al., *APRIL promotes B-1 cell-associated neoplasm.* Cancer Cell, 2004. **6**(4): p. 399-408.

110. Enzler, T., et al., *Chronic lymphocytic leukemia of Emu-TCL1 transgenic mice undergoes rapid cell turnover that can be offset by extrinsic CD257 to accelerate disease progression.* Blood, 2009. **114**(20): p. 4469-76.

111. Lascano, V., et al., *Chronic lymphocytic leukemia disease progression is accelerated by APRIL-TACI interaction in the TCL1 transgenic mouse model.* Blood, 2013. **122**(24): p. 3960-3.

112. Haney, S.L., et al., *Dnmt3a Is a Haploinsufficient Tumor Suppressor in CD8+ Peripheral T Cell Lymphoma.* PLoS Genet, 2016. **12**(9): p. e1006334.

113. Pradhan, S., et al., *Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation.* J Biol Chem, 1999. **274**(46): p. 33002-10.

114. Krueger, F. and S.R. Andrews, *Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications.* Bioinformatics, 2011. **27**(11): p. 1571-2.

115. Assenov, Y., et al., *Comprehensive analysis of DNA methylation data with RnBeads.* Nat Methods, 2014. **11**(11): p. 1138-40.

116. Wu, H., et al., *Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates.* Nucleic Acids Res, 2015. **43**(21): p. e141.

117. Krzywinski, M., et al., *Circos: an information aesthetic for comparative genomics.* Genome Res, 2009. **19**(9): p. 1639-45.

118. Lara-Astiaso, D., et al., *Immunogenetics. Chromatin state dynamics during blood formation.* Science, 2014. **345**(6199): p. 943-9.

119. Quinlan, A.R. and I.M. Hall, *BEDTools: a flexible suite of utilities for comparing genomic features.* Bioinformatics, 2010. **26**(6): p. 841-2.

120. Daily, K., et al., *MotifMap: integrative genome-wide maps of regulatory motif sites for model species.* BMC Bioinformatics, 2011. **12**: p. 495.

121. Xie, X., P. Rigor, and P. Baldi, *MotifMap: a human genome-wide map of candidate regulatory motif sites.* Bioinformatics, 2009. **25**(2): p. 167-74.

122. Sun, Z., et al., *SAAP-RRBS: streamlined analysis and annotation pipeline for reduced representation bisulfite sequencing.* Bioinformatics, 2012. **28**(16): p. 2180-1.

123. Kim, D., et al., *TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions.* Genome Biol, 2013. **14**(4): p. R36.

124. Trapnell, C., et al., *Differential analysis of gene regulation at transcript resolution with RNA-seq.* Nat Biotechnol, 2013. **31**(1): p. 46-53.

125. Talos, F., et al., *Mitochondrially targeted p53 has tumor suppressor activities in vivo.* Cancer Res, 2005. **65**(21): p. 9971-81.

126. Hirst, M. and M.A. Marra, *Next generation sequencing based approaches to epigenomics.* Brief Funct Genomics, 2010. **9**(5-6): p. 455-65.

127. van der Weyden, L., et al., *Jdp2 downregulates Trp53 transcription to promote leukaemogenesis in the context of Trp53 heterozygosity.* Oncogene, 2013. **32**(3): p. 397-402.

128. Wang, L., et al., *Genomic profiling of Sezary syndrome identifies alterations of key T cell signaling and differentiation genes.* Nat Genet, 2015. **47**(12): p. 1426-34.

129. Piccaluga, P.P., et al., *Platelet-derived growth factor alpha mediates the proliferation of peripheral T-cell lymphoma cells via an autocrine regulatory pathway.* Leukemia, 2014. **28**(8): p. 1687-97.

130. Kovacic, B., et al., *STAT1 acts as a tumor promoter for leukemia development.* Cancer Cell, 2006. **10**(1): p. 77-87.

131. Su, X., et al., *Overexpression of TRIM14 promotes tongue squamous cell carcinoma aggressiveness by activating the NF-kappaB signaling pathway.* Oncotarget, 2016. **7**(9): p. 9939-50.

132. Piu, F., et al., *AP-1 repressor protein JDP-2: inhibition of UV-mediated apoptosis through p53 down-regulation.* Mol Cell Biol, 2001. **21**(9): p. 3012-24.

133. Bitton-Worms, K., E. Pikarsky, and A. Aronheim, *The AP-1 repressor protein, JDP2, potentiates hepatocellular carcinoma in mice.* Mol Cancer, 2010. **9**: p. 54.

134. Hwang, H.C., et al., *Identification of oncogenes collaborating with p27Kip1 loss by insertional mutagenesis and high-throughput insertion site analysis.* Proc Natl Acad Sci U S A, 2002. **99**(17): p. 11293-8.

135. Donehower, L.A., et al., *Effects of genetic background on tumorigenesis in p53-deficient mice.* Mol Carcinog, 1995. **14**(1): p. 16-22.

136. Kato, K., et al., *Molecular genetic and cytogenetic analysis of a primary cutaneous CD8-positive aggressive epidermotropic cytotoxic T-cell lymphoma.* Int J Hematol, 2016. **103**(2): p. 196-201.

137. Hollander, M.C., et al., *Gadd45a acts as a modifier locus for lymphoblastic lymphoma.* Leukemia, 2005. **19**(5): p. 847-50.

138. Hodson, D.J., et al., *Deletion of the RNA-binding proteins ZFP36L1 and ZFP36L2 leads to perturbed thymic development and T lymphoblastic leukemia.* Nat Immunol, 2010. **11**(8): p. 717-24.

139. Malik, D., et al., *miR-2909-mediated regulation of KLF4: a novel molecular mechanism for differentiating between B-cell and T-cell pediatric acute lymphoblastic leukemias.* Mol Cancer, 2014. **13**: p. 175.

140. Li, W., et al., *Genome-wide analyses identify KLF4 as an important negative regulator in T-cell acute lymphoblastic leukemia through directly inhibiting T-cell associated genes.* Mol Cancer, 2015. **14**: p. 26.

141. Roman-Gomez, J., et al., *5' CpG island hypermethylation is associated with transcriptional silencing of the p21(CIP1/WAF1/SDI1) gene and confers poor prognosis in acute lymphoblastic leukemia.* Blood, 2002. **99**(7): p. 2291-6.

142. Herman, J.G., et al., *Distinct patterns of inactivation of p15INK4B and p16INK4A characterize the major types of hematological malignancies.* Cancer Res, 1997. **57**(5): p. 837-41.

143. Toujani, S., et al., *High resolution genome-wide analysis of chromosomal alterations in Burkitt's lymphoma.* PLoS One, 2009. **4**(9): p. e7089.

144. Fong, L.Y., et al., *Muir-Torre-like syndrome in Fhit-deficient mice.* Proc Natl Acad Sci U S A, 2000. **97**(9): p. 4742-7.

145. Gilfillan, S., et al., *DNAM-1 promotes activation of cytotoxic lymphocytes by nonprofessional antigen-presenting cells and tumors.* J Exp Med, 2008. **205**(13): p. 2965-73.

146. Aries, I.M., et al., *EMP1, a novel poor prognostic factor in pediatric leukemia regulates prednisolone resistance, cell proliferation, migration and adhesion.* Leukemia, 2014. **28**(9): p. 1828-37.

147. Xie, L. and P.L. Green, *Envelope is a major viral determinant of the distinct in vitro cellular transformation tropism of human T-cell leukemia virus type 1 (HTLV-1) and HTLV-2.* J Virol, 2005. **79**(23): p. 14536-45.

148. Ye, J., L. Xie, and P.L. Green, *Tax and overlapping rex sequences do not confer the distinct transformation tropisms of human T-cell leukemia virus types 1 and 2*. J Virol, 2003. **77**(14): p. 7728-35.

149. Hervouet, E., F.M. Vallette, and P.F. Cartron, *Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation.* Epigenetics, 2009. **4**(7): p. 487-99.

150. Haney, S.L., et al., *Loss of Dnmt3a induces CLL and PTCL with distinct methylomes and transcriptomes in mice.* Sci Rep, 2016. **6**: p. 34222.

151. Ellin, F., et al., *Real-world data on prognostic factors and treatment in peripheral T-cell lymphomas: a study from the Swedish Lymphoma Registry*. Blood, 2014. **124**(10): p. 1570-7.

152. Abramson, J.S., et al., *Peripheral T-cell lymphomas in a large US multicenter cohort: prognostication in the modern era including impact of frontline therapy.* Ann Oncol, 2014. **25**(11): p. 2211-7.

153. Quesada, V., et al., *Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia.* Nat Genet, 2011. **44**(1): p. 47-52.

154. Haferlach, T., et al., *Clinical utility of microarray-based gene expression profiling in the diagnosis and subclassification of leukemia: report from the International Microarray Innovations in Leukemia Study Group.* J Clin Oncol, 2010. **28**(15): p. 2529-37.

155. Liao, W., et al., *Gene expression and splicing alterations analyzed by high throughput RNA sequencing of chronic lymphocytic leukemia specimens.* BMC Cancer, 2015. **15**: p. 714.

156. Iqbal, J., et al., *Molecular signatures to improve diagnosis in peripheral T-cell lymphoma and prognostication in angioimmunoblastic T-cell lymphoma*. Blood, 2010. **115**(5): p. 1026-36.

157. Chadeau-Hyam, M., et al., *Prediagnostic transcriptomic markers of Chronic lymphocytic leukemia reveal perturbations 10 years before diagnosis.* Ann Oncol, 2014. **25**(5): p. 1065-72.

158. Yang, L., et al., *DNMT3A Loss Drives Enhancer Hypomethylation in FLT3-ITD-Associated Leukemias.* Cancer Cell, 2016. **29**(6): p. 922-34.

159. Vasmatzis, G., et al., *Genome-wide analysis reveals recurrent structural abnormalities of TP63 and other p53-related genes in peripheral T-cell lymphomas.* Blood, 2012. **120**(11): p. 2280-9.

160. da Silva Almeida, A.C., et al., *The mutational landscape of cutaneous T cell lymphoma and Sezary syndrome.* Nat Genet, 2015. **47**(12): p. 1465-70.

161. Sandell, R.F., R.L. Boddicker, and A.L. Feldman, *Genetic Landscape and Classification of Peripheral T Cell Lymphomas.* Curr Oncol Rep, 2017. **19**(4): p. 28.

162. Kiel, M.J., et al., *Genomic analyses reveal recurrent mutations in epigenetic modifiers and the JAK-STAT pathway in Sezary syndrome.* Nat Commun, 2015. **6**: p. 8470.

163. Girardi, T., et al., *The genetics and molecular biology of T-ALL.* Blood, 2017. **129**(9): p. 1113-1123.

164. Fuks, F., et al., *Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription.* EMBO J, 2001. **20**(10): p. 2536-44.

165. Vire, E., et al., *The Polycomb group protein EZH2 directly controls DNA methylation.* Nature, 2006. **439**(7078): p. 871-4.

166. Iqbal, J., et al., *Gene expression signatures delineate biological and prognostic subgroups in peripheral T-cell lymphoma.* Blood, 2014. **123**(19): p. 2915-23.

167. Virgilio, L., et al., *Deregulated expression of TCL1 causes T cell leukemia in mice.* Proc Natl Acad Sci U S A, 1998. **95**(7): p. 3885-9.

168. Bichi, R., et al., *Human chronic lymphocytic leukemia modeled in mouse by targeted TCL1 expression.* Proc Natl Acad Sci U S A, 2002. **99**(10): p. 6955-60.

169. Kiel, M.J., et al., *Integrated genomic sequencing reveals mutational landscape of T-cell prolymphocytic leukemia.* Blood, 2014. **124**(9): p. 1460-72.

170. Roos, J., et al., *Expression of TCL1 in hematologic disorders.* Pathobiology, 2001. **69**(2): p. 59-66.

171. Vaillant, F., et al., *Enforced expression of Runx2 perturbs T cell development at a stage coincident with beta-selection.* J Immunol, 2002. **169**(6): p. 2866-74.

172. Woolf, E., et al., *Runx3 and Runx1 are required for CD8 T cell development during thymopoiesis.* Proc Natl Acad Sci U S A, 2003. **100**(13): p. 7731-6.

173. Setoguchi, R., et al., *Repression of the transcription factor Th-POK by Runx complexes in cytotoxic T cell development.* Science, 2008. **319**(5864): p. 822-5.

174. McCaughtry, T.M., et al., *Conditional deletion of cytokine receptor chains reveals that IL-7 and IL-15 specify CD8 cytotoxic lineage fate in the thymus.* J Exp Med, 2012. **209**(12): p. 2263-76.

175. Borrego, F., et al., *The CD94/NKG2 family of receptors: from molecules and cells to clinical relevance.* Immunol Res, 2006. **35**(3): p. 263-78.

176. Li, P., et al., *Reprogramming of T cells to natural killer-like cells upon Bcl11b deletion.* Science, 2010. **329**(5987): p. 85-9.

177. Spencer, D.H., et al., *CpG Island Hypermethylation Mediated by DNMT3A Is a Consequence of AML Progression.* Cell, 2017. **168**(5): p. 801-816 e13.

178. Kramer, A.C., et al., *Dnmt3a regulates T-cell development and suppresses T-ALL transformation.* Leukemia, 2017.

179. Koya, J., et al., *DNMT3A R882 mutants interact with polycomb proteins to block haematopoietic stem and leukaemic cell differentiation.* Nat Commun, 2016. **7**: p. 10924.

180. Ntziachristos, P., et al., *Genetic inactivation of the polycomb repressive complex 2 in T cell acute lymphoblastic leukemia.* Nat Med, 2012. **18**(2): p. 298-301.

181. Zhang, J., et al., *The genetic basis of early T-cell precursor acute lymphoblastic leukaemia.* Nature, 2012. **481**(7380): p. 157-63.

182. Thol, F., et al., *Acute myeloid leukemia derived from lympho-myeloid clonal hematopoiesis.* Leukemia, 2017.

183. Meyer, S.E., et al., *DNMT3A Haploinsufficiency Transforms FLT3ITD Myeloproliferative Disease into a Rapid, Spontaneous, and Fully Penetrant Acute Myeloid Leukemia.* Cancer Discov, 2016. **6**(5): p. 501-15.

184. Roy, D., et al., *Tumor suppressor genes FHIT and WWOX are deleted in primary effusion lymphoma (PEL) cell lines.* Blood, 2011. **118**(7): p. e32-9.

185. Haura, E.B., J. Turkson, and R. Jove, *Mechanisms of disease: Insights into the emerging role of signal transducers and activators of transcription in cancer.* Nat Clin Pract Oncol, 2005. **2**(6): p. 315-24.