The role of exon 1B-containing DNMT3B in embryonal tumors with multilayered rosettes

Michael Neirinck
260308116

Department of Experimental Medicine - McGill University

A thesis submitted in partial fulfillment of the requirements for the degree of
Master of Science

Dr. Cindy Goodyer - Chairperson of Supervisory Committee
Dr. Nada Jabado - Supervisor

April 2016

Abstract

Embryonal tumors of the central nervous system are the most frequent malignant solid cancers found within the pediatric population. These are currently being researched around the world due to their complex genetic origins, unusual aggressiveness and tendency to be treatment-resistant. Recent work has reclassified a subset of these cancers as embryonal tumors with multilayered rosettes or ETMRs. It has been demonstrated that almost all ETMR patients have one specific epigenetic alteration in common, a chromosomal fusion between the promoter of the tweety family member 1 (TTYH1) gene (which encodes a chloride ion channel expressed in the brain) and chromosome 19 microRNA cluster (C19MC), the largest microRNA cluster within the human genome. It was shown that, as a result of this fusion, overexpression of a specific isoform of DNA methyltransferase 3 beta (DNMT3B), exon 1B-containing DNMT3B, occurs specifically in these tumors. Due to the *de novo* ability of DNMT3B to change the DNA methylation landscape within cells, it is not surprising that all ETMR patients share a common DNA methylation pattern when analyzed. As exon 1B-containing DNMT3B is exclusively significantly overexpressed in ETMRs, it would be of great interest to determine the carcinogenic properties of this potential oncoprotein. We set out to overexpress exon 1B-containing DNMT3B (using nucleofection and lentiviral transduction) in PFSK-1, a primitive neuroectodermal tumor (PNET) cell line that would be likely to mimic ETMRs following exon 1B-containing DNMT3B overexpression. Demonstration of a change in DNA methylation (which would hopefully cluster with ETMR patients), cell morphology and cell behavior following overexpression of this unique DNMT3B variant would be a step forward in elucidating the genetic origins of ETMRs and aid in the future treatment of ETMR patients through the potential development of DNMT3B inhibitors.

Abstract (French)

Les tumeurs embryonnaires du système nerveux central sont les cancers solides malignes les plus fréquemment trouvées dans la population pédiatrique. Ceux-ci sont actuellement l'objet de recherches actives dans le monde entier en raison de leurs origines génétiques complexes, l'agressivité et la tendance à être résistantes aux traitements usuels. Des publications récentes ont reclassés un sous-ensemble de ces cancers comes étants des tumeurs embryonnaires avec rosettes ou ETMRs multicouches. Il a été démontré que les patients ayant l'ETMR ont tous une modification épigénétique spécifique en commun, une fusion chromosomique entre le promoteur du membre de la famille de tweety 1 (TTYH1), un gene qui code pour un canal de chlorure ionique exprimé dans le cerveau, et le chromosome 19 pres du microARN (C19MC), le plus grand groupe de microARN dans le génome humain. Il a été démontré que, à la suite de cette fusion, la surexpression d'un isoforme spécifique de l'ADN méthyltransférase 3 bêta (DNMT3B), le DNMT3B exon 1B, par ces tumeurs. En raison de la capacité de novo de DNMT3B de changer le repertoire des genes methylés dans les cellules, il est pas surprenant que tous patients ayant l'ETMR partagent un motif de méthylation de l'ADN commun lors de l'analyse. Par exemple, exon 1B contenant DNMT3B est exclusivement significativement surexprimé dans ETMR. Il serait d'un grand intérêt de déterminer les propriétés cancérigènes de cette oncoprotéine potentiel. Nous avont debuté par la surexpression de l'exon 1B contenant le DNMT3B (en utilisant la nucléofection et transduction par lentivirus) dans la lignée cellulaire, PFSK-1, qui represente une tumeur primitive neuroectodermique (PNET) et qui serait susceptible d'imiter ETMRs suivant la sur-expression de l'exon 1B contenant le DNMT3B. La démonstration d'un changement dans la méthylation de l'ADN (ce qui, on l'espère regrouper avec les patients de ETMR), la morphologie cellulaire et le comportement cellulaire après surexpression de cette variante unique DNMT3B serait un pas en avant dans la compréhension des origines génétiques de ETMRs. Ceci sans doute aiderait dans le traitement futur des patients ETMR suite au développement potentiel des inhibiteurs DNMT3B.

Table of Contents **Page**

List of Figures

List of Tables

List of Abbreviations

MB – Medulloblastoma

AT/RT – Atypical teratoid/rhabdoid tumor

CNS – Central nervous system

PNET – Primitive neuroectodermal tumor

ETMR – Embryonal tumor with multilayered rosettes

Rb – Retinoblastoma

RB1 – Retinoblastoma 1

RBL1 – Retinoblastoma-like protein 1

RBL2 – Retinoblastoma-like protein 2

DNMT1 – DNA methyltransferase 1

DNMT3A – DNA methyltransferase 3A

DNMT3B – DNA methyltransferase 3 beta

TTYH1 – Tweety family member 1

C19MC – Chromosome 19 microRNA cluster

PTMs – Post-translational modifications

hNSC – Human neural stem cell

Acknowledgements

I would like to begin by thanking my supervisor, Dr. Nada Jabado, for giving me the opportunity to carry out research in her lab and for funding my graduate studies. It was amazing to work with someone who is both extremely knowledgeable and passionate about the work they do. Despite the many technical problems I encountered, Dr. Jabado always kept a positive attitude and focused on solving problems instead of dwelling on them. The critical thinking and research skills I have gained since being under her guidance are invaluable assets I will carry with me for the rest of my career. I also want to thank Dr. Jabado for allowing me to attend the University of Toronto's 2014 Symposium on Epigenetic Mechanisms in Cancer with her. Hearing from some of the top scientists in the field of epigenetic cancer research was a unique experience that I will not forget.

In addition, I want to thank everyone who was a member of the Jabado lab during my time there: Dr. Adam Fontebasso, Yelu Zhang, Dr. Noha Gerges, Dr. Tenzin Gayden, Dr. Denise Bechet, Nisreen Ibrahim, Andrea Bajic, Byron Bellaiche, Sima Khazaei, Dr. Leonie Mikael, Dr. Ashot Harutyunyan, Dr. Alexander Weil and Dr. Eef Harmsen. I greatly appreciate the chance to work with such hard working and intelligent individuals who helped to create a supportive lab environment that facilitated productivity and the ability to learn.

I want to specifically thank Damien Faury and Caterina Russo for always being there when I had a question that needed to be answered. I came into this lab with little proficiency in cell culture and the laboratory techniques that revolved around it, but thanks to Damien and Cathy, I now know how to carry out a multitude of new experiments. I cannot even count the number of times I came to them with a problem and left with a solution; thank you for all your help during my time here.

I would also like to give a special thanks to my academic advisor Dr. Cindy Goodyer and my thesis committee members Dr. Kolja Eppert and Dr. Marc Fabian, for offering advice when needed and helping to facilitate the troubleshooting that occurred throughout my Master's.

Lastly, I want to thank my family and friends for providing me with the support I needed to carry out completion of this degree. Their daily support kept me motivated to continue pushing forward and working towards my goals, without them, this would not have been possible.

Chapter 1 – Introduction

1.1 – Objective of This Thesis

The objective of my thesis is to study the role of DMMT3B in the development of embryonal brain cancer. DMNT3B is a *de novo* DNA methyltransferase, and catalyzes the addition of a methyl group (-CH_3) to carbon 5 on cytosine residues in the context of CpG dinucleotides[1]. In a recent study from our lab, a novel isoform of DNMT3B, whose expression is usually restricted to weeks 8-10 of embryogenesis in fetal brain, was found to be highly upregulated in a cohort of patients with ETMRs[2]. Results from our lab indicted that an amplification of a gene fusion between the promoter of *TTYH1* (which encodes a transmembrane chloride ion channel specific to the brain) and the microRNA (miRNA) cluster C19MC was found within 100% of patient samples analyzed[2]. This fusion caused an increased transcription of miRNAs within this cluster, with a concomitant decrease in expression of mRNA for the tumor suppressor retinoblastoma-like 2 (RBL2)[2]. This, in turn, causes overexpression of a unique DNMT3B splice variant (exon 1B-containing DNMT3B), thereby changing the epigenetic landscape of neuronal cells in ETMR patients through the establishment of a unique DNA methylation profile in these cells[2].

In order to investigate this, in Chapter 1, I will describe the epidemiology and genetic basis of the four major groups of pediatric embryonal brain cancer (medulloblastomas, ATRTs, CNS PNETs and ETMRs), followed by a specific focus on ETMRs, and the roles of miRNAs in oncogenesis. The association of the retinoblastoma family of tumor suppressors in cancer, specifically RBL2 and its ability to suppress expression of DNMT3B, has also been examined. The epigenetic modification of DNA methylation, its role in development and its role in the development of cancer are reviewed in detail, with a special focus on the involvement of DNMT3B and exon 1B-containing DNMT3B in embryogenesis and cancer. Chapter 1 will conclude with a rationalization of why it is believed that exon 1B-containing DNMT3B is responsible for the formation of ETMRs in children, and a description of how we will set out to prove this. Chapter 2 will describe the experimental data I have collected, while Chapter 3 outlines the conclusions related to my work, in the context of the reviewed literature.

1.2 - Epidemiology and Genetic Alterations in Central Nervous System Embryonal Tumors

1.2.1 - Brief Introduction

Central nervous system (CNS) and miscellaneous intracranial and intraspinal neoplasms comprise the largest group of malignant solid cancers within the pediatric population, making up over 17% of all childhood cancers[3]. Tumors of embryonal origin are by far the leading group of malignant brain tumors in children, and carry with them a devastatingly high mortality rate and long-term morbidity for survivors[4]. These cancers are known for their undifferentiated cellular nature, and originate from neuroepithelial cells that share a predisposition to disseminate through the neuraxis[5]. In recent years, our understanding of these aggressive neoplasms has increased substantially, with many studies focused on determining the underlying genetic alterations that cause these tumors to form. These experiments have led to the division of these cancers into numerous subgroups, classified by various characteristics including histopathology, aberrant protein signaling, genetic structure, gene expression and clinical statistics. There are four major CNS embryonal tumor types affecting children, all of which are grade IV due to their metastatic potential and primitive cellular nature; medulloblastomas, atypical teratoid/rhabdoid tumors (AT/RT), CNS primitive neuroectodermal tumors (PNETs) and embryonal tumors with multilayered rosettes (ETMRs)[5].

1.2.2 - Medulloblastoma

Medulloblastoma (MB) is a highly malignant infratentorial primitive neuroectodermal tumor that arises mainly in the cerebellum, a region of the brain responsible for motor control[6]. Within the major groups of embryonal brain tumors, medulloblastoma is the most common and is the most prevalent malignant pediatric brain cancer, representing 18-28% of all brain tumors that affect children between the ages of 0 and 14 years[7,8]. Medulloblastoma has a bimodal distribution for occurrence, peaking at the ages of 3-4 and again between the ages of 8-9. It has a skewed gender distribution, with males 1.7 times more likely to develop the cancer than females[8]. It is currently classified by the WHO as a grade IV malignant embryonal neoplasm, and was previously divided into five histological varieties; classical, anaplastic tumors with enlarged nuclei, large cell, the desmoplastic/nodular variant and medulloblastoma with extensive nodularity[4,9]. Although there were originally five histological varieties, medulloblastoma is now

known to comprise four distinct non-overlapping groups based on transcriptional profile and genetic alterations (WNT, SHH, Group 3 and Group 4) that correlate with clinical parameters, histology and metastatic status[10]. This set the precedent for the more recent reclassification of several histologically different embryonal tumors into a single entity now known as ETMRs. Furthermore, a recent study that performed integrative genomic and DNA methylation analysis of primary medulloblastoma samples, and their corresponding metastatic compartments, demonstrated that these subgroup specific molecular aberrations remain stable over time[11]. As there are four molecular subgroups of medulloblastoma, it is not surprising that the oncogenic changes taking place in these primitive embryonal cells are unique to each subgroup.

Irregular activation of the sonic hedgehog (SHH) pathway is present in roughly 25% of medulloblastoma cases[12,13]. The hedgehog (Hh) signaling pathway begins with a family of three precursor proteins: sonic hedgehog (SHH), Desert hedgehog (DHH) and Indian hedgehog (IHH), whose cleaved and post-translationally modified downstream products play essential roles in embryogenesis, cellular proliferation, tissue development and stem cell maintenance[14]. These proteins bind to the transmembrane receptor PATCH (*PTCH1*) localized at the base of the primary cilium of cells, causing the inhibition of the downstream G-protein coupled receptor Smoothened (SMO) to be released[15]. This allows SMO to translocate from an intracellular compartment to the cell membrane and begin signal transduction[16]. Following the activation of SMO, a family of zinc-finger transcription factors called GLI (GLI1, GLI2 and GLI3) are stabilized and activated (through several intermediate kinases) in the primary cilium, allowing them to translocate to the nucleus and regulate the expression of downstream targets[17]. The development of the cerebellum occurs during postnatal days 3-9, and requires extensive proliferation of granule neuron progenitor (GNP) cells in the outer layer of the cerebellum as a result of SHH signaling[18]. Therefore, constitutive activation of this signaling pathway can lead to oncogenesis, and in some cases, the development of medulloblastoma.

It is known that inactivating mutations in the *PTCH1* gene (encoding the Hh receptor PATCH) are present within a subset of SHH-MB patients, and transgenic mice containing only one copy of the *PTCH1* gene have been shown to develop medulloblastoma[19,20]. One study of particular interest demonstrated through sequencing a cohort of SHH-MB patients that infant (< 4 years old), children (4-17 years old) and adult (> 18 years old) SHH-MB do not contain the same frequency of genetic mutations[21,22]. Mutations in *PTCH1* were found in all age groups, but

mutations in *SUFU* (a suppressor of the SHH pathway) and *SMO* were found disproportionately in infants and adults, respectively. Mutations in the *SMO* gene have also been detected in PNET patients. Given the recent trend towards personalized medicine, it is important to consider that mutations in genes upstream in the SHH pathway (*SHH*/*SHH* analogue amplification, *PTCH1* and *SMO*) are susceptible to the currently available SMO inhibitors, while mutations downstream (such as those found in *GLI2*) would be resistant to SMO inhibitors[22]. Mutations in the SHH pathway are not the only oncogenic drivers of SHH-MB, as mutations in the tumor suppressor *TP53* correlate with five-year survival (41% to 81% in patients with and without *TP53* mutations, respectively)[23]. This further demonstrates the complexity of oncogenic drivers within one subgroup of a single embryonal brain cancer.

WNT subgroup medulloblastoma patients carry the most favorable prognosis within the four MB subgroups, and are associated with a highly favorable survival outcome, with 95% 5- and 10-year overall survival in pediatric patients[12]. This favorable prognosis is likely due to WNT subgroup MB patients have fewer genomic alterations when compared to the other three subgroups[24]. This WNT subgroup is also the smallest subgroup of MB, with only 11% of MB patients[25].

The WNT protein family is comprised of 19 members of secreted glycoproteins with a highly conserved cysteine-rich region. In the absence of these proteins, cytoplasmic β-catenin is phosphorylated at three amino acid residues by the trimeric destruction complex (consisting of APC, AXIN and GSK3β) following phosphorylation at Serine 45 by CK1-α[26,27]. Phosphorylated β-catenin is then recognized by a E3 ubiquitin ligase (β-Trcp) and marked for degradation by the proteasome, keeping β-catenin at low levels in the cytoplasm. In the canonical WNT signaling pathway, WNT family proteins bind to the extracellular domains of transmembrane proteins (Frizzled and Lrp5/6) to induce phosphorylation of Dishevelled (Dvl), which in turn phosphorylates and inactivates GSK3β[27,28]. This allows β-catenin to accumulate in the cytoplasm, translocate to the nucleus and form a complex with the TCF family of transcription factors and coactivators (such as CBP and p300) to drive expression of target genes, which include c-Myc and cyclin D1[29]. The resulting effects on transcription are important for proliferation and survival, making the canonical WNT pathway oncogenic when overactivated[30]. A second, non-canonical, WNT pathway (which is mainly β-catenin independent) is involved in differentiation, cell polarity and cell migration[30]. The interaction between the canonical and noncanonical WNT

pathways allow complex processes to be coordinated, and is required for embryonic development, stem cell maintenance and adult neurogenesis[29,31].

Genomic analysis in sporadic WNT-MB patients has shown frequent activating mutations in the gene encoding β-catenin (*CTNNB1*)[32]. Activating germline mutations in the gene encoding APC (an inhibitor of the WNT pathway) have also been detected in medulloblastoma patients with Turcot syndrome[33]. Less common mutations found in sporadic WNT-MB include those found in the genes encoding the RNA helicase DDX3X (which allows transactivation of a TCF promoter by mutant β-catenin) and AXIN1/2[34,35]. WNT-MB subgroup patients have also been shown to have a distinct microRNA (miRNA) profile, with several plausible oncogenic miRNAs being overexpressed in this group, including miR-193a-3p, miR-224, miR-148a, miR-23b, and miR-365[36]. This is of significance, as ETMRs, a separate group of embryonal brain cancer is also driven by the amplification of a miRNA cluster on chromosome 19, C19MC (causing decreased expression of different tumor suppressors)[2].

Dysregulation of the Hedgehog and Wingless pathways are not the only drivers of medulloblastoma. Group 3 and Group 4 medulloblastoma patients carry the worst survival outcome and little is known about their molecular pathogenesis, although they make up 28% and 34% of all medulloblastomas[13]. The main difference between the two groups is that overexpression/amplification of the oncogene *MYC* occurs mainly in Group 3 tumors and is associated with the worst prognosis, while Group 4 tumors have an overrepresentation of signaling in pathways related to neuronal development[13]. Group 3 and Group 4 medulloblastoma patients tend to have overlapping gene expression profiles. This is likely due to both groups having the most frequent structural aberration in medulloblastoma, isochromosome 17q (i17q), although this is more frequently seen in Group 4. This is currently believed to arise due to non-allelic homologous recombination between repeat sequences in the pericentric region of 17q or the 'Smith-Magenis' region[37]. This causes one of the copies of chromosome 17 to become a chromosome with two centromeres and two long 'q' arms (with the short 'p' arm of chromosome 17 being deleted in the process) resulting in the formation of aneuploid cells that are favored in oncogenic selection[4]. Group 3 and 4 MB have also been shown to have mutations in EZH2 (a histone trimethyltransferase part of the PRC2 complex) and KDM6A (a histone demethylase), demonstrating the involvement of epigenetic dysregulation in the development of medulloblastoma[38].

1.2.3 - Atypical Teratoid/Rhabdoid Tumors

Although atypical teratoid/rhabdoid tumors (AT/RTs) are somewhat histologically similar to medulloblastoma, they are a distinct clinical entity with unique immunohistology and gene expression[8]. AT/RTs occur primarily in childhood, with the median age of onset being 11 months; the median time of survival following diagnosis is around 10 months[39,40]. Treatment with chemotherapy, radiotherapy or in combination has demonstrated to extend survival to 14-18 months[39]. AT/RT is the central nervous system component of the malignant rhabdoid tumor (MRT) family, and tumors are generally located infratentorial (most commonly in the cerebellum or at the margin of the cerebellum and pons)[41]. AT/RT patients harbor inactivating mutations in the *Snf5/Ini1/SMARCB1* gene roughly 70% of the time[42,43].

The *SMARCB1* gene encodes the Snf5/INI1 protein, which is a key component of the SWI/SNF core chromatin remodeling complex[44]. The majority of AT/RTs show loss of heterozygosity (LOH) at the *Ini1* gene locus; these AT/RT biopsies lack immunohistochemical staining of INI1 and deletions of *Ini1* can be detected through fluorescence in situ hybridization (FISH)[43]. The relation between INI1 expression and oncogenesis is currently not fully understood, although studies have demonstrated that it indirectly regulates the expression of several cell cycle regulators (such as cyclin D1, p16INK4A and pRb) through the ATP-dependent chromatin-remodeling SWI/SNF complex[45]. Irregular expression of these important cell cycle proteins could lead to uncontrolled proliferation and development of AT/RT, as it has been demonstrated that INI1 activates G2-M phase transition through the $p16^{INK4A}$-cyclinD/CDK4-pRb-E2F pathway[46]. The main genetic mutation detected in AT/RT, being in a gene that encodes a component of a chromatin remodeling complex, highlights the fact that a large proportion of embryonal brain cancers develop due to dysregulated epigenetics.

1.2.4 - Central Nervous System Primitive Neuroectodermal Tumors

According to the 2007 WHO classification of tumors of the central nervous system (CNS), there were five subtypes of primitive neuroectodermal tumors (PNETs); CNS PNET (not otherwise classified), CNS neuroblastoma, CNS ganglioneuroblastoma, medulloepithelioma and ependymoblastoma, with the latter two having been more recently classified under the umbrella term ETMR[9,47]. When CNS PNETs displayed features of embryonal neural tube formation, they were referred to as medulloepithelioma, and those with ependymoblastic rosettes were termed

ependymoblastoma[7]. However, recent genetic profiling of these two tumor subgroups revealed that they share a conserved and distinct genetic profile with another subgroup of embryonal brain cancer, causing them to be reclassified. CNS PNETs are rare embryonal tumors that account for approximately 5% of cerebral hemispheric tumors and 3% of all brain tumors in children[48]. Event-free survival at 5 years is 75% for patients with average risk and 60% for patients with high-risk disease[5].

The term CNS PNET generally refers to supratentorial (above the cerebellar tentorium) PNETs; medulloblastomas were previously classified as PNETs, but due to recent transcriptional profiling, clinical data, and different anatomical location (infratentorial), they are now described as a unique clinical entity[9]. CNS PNETs with only neuronal differentiation are termed CNS neuroblastomas, while those with neuronal differentiation and the presence of neoplastic ganglion cells are deemed CNS ganglioneuroblastomas[9].

Very few genes are found recurrently mutated in patients with CNS neuroblastoma, but one study determined that recurrent alterations in the telomerase reverse transcriptase (*TERT*) gene were found in 31% of patients with high-risk neuroblastoma and were mutually exclusive to *MYCN* amplification and mutations in *ATRX* (which encodes a histone chaperone)[49]. They also identified that alternative lengthening of telomeres (ALT) was present in neuroblastoma without *TERT* rearrangement or *MYCN* amplification, demonstrating that telomere lengthening is likely a key mechanism for the progression of CNS neuroblastoma in some cases[49].

A study by Picard et al. identified through an integrative genomic analysis on a large cohort of hemispheric CNS PNETs that there are 3 molecular subgroups of CNS PNETs with distinct clinical and molecular features[50]. Although the study ruled out misdiagnosis of AT/RT through immunoreactivity of INI1, it was not known at the time that ETMRs were their own clinical entity, and one of the three subgroups was characterized by high expression of LIN28 and amplification of C19MC (hallmarks of ETMRs)[50]. This group, deemed Group 1, also showed decreased OLIG2 (a transcription factor in the CNS) expression and increased WNT and SHH signaling. Group 2 was found to have decreased WNT and SHH signaling, increased OLIG2 expression and decreased LIN28 expression; while the highly metastatic Group 3 had decreased expression of both LIN28 and OLIG2 and increased signaling in other known oncogenic pathways (those involving TGFβ and PTEN)[50]. This demonstrates the complex heterogeneity within even a single embryonal brain cancer subgroup.

Roughly 10-15% of PNETs harbor mutations in the *PTCH1* tumor suppressor gene, which encodes a component of the SHH signaling pathway[51]. High cyclin D1 expression has also been detected in CNS PNETs, likely aiding in the ability of the cancer to mimic stem cells, as high cyclin D1 expression is found developing fetal neuroblasts[52]. A large genetic screening on a cohort of CNS-PNETs and pediatric glioblastomas (GBM, another grade IV brain cancer) revealed a common mutation in a gene encoding a variant of histone H3 (resulting in H3.3 G34R) within the two groups[53]. Although in this study only 11% of CNS PNETs (4 of 33) had mutations in the histone encoding gene, the wide prevalence of this mutation in pediatric GBM gives insight into the epigenetic involvement within pediatric brain cancer. The presence of H3.3 G34R mutations in CNS PNETs could lead to a misdiagnosis of GBM, however due to the difference in cellularity of GBMs and CNS PNETs, careful examination of the histopathology would prevent this from occurring[53]. Given the recent reclassification of some subgroups of CNS PNETs and the wide variety of genetic alterations that take place within the remaining groups, it is likely that further reclassification will occur in the future as more research on the biology causing these cancers to form is performed.

1.2.5 - Embryonal Tumors with Abundant Neuropil and True Rosettes

Prior to being reclassified as a single entity, there were three known histological varieties of embryonal rosette-forming neuroepithelial brain tumors. Two of these were recognized as a subdivision of CNS PNETs, according to the 2007 WHO classification of tumors of the CNS, medulloepithelioma (MEPL) and ependymoblastoma (EBL)[9]. In addition, a rare histologically distinct PNET subtype, known as embryonal tumors with abundant neuropil and true rosettes (ETANTR) was first reported by Eberhart et al. in 2000[54]. These three variants (MEPL, EBL, ETANTR) all predominantly affect infants under the age of 4 years[47], and are associated with a very poor clinical outcome. MEPL is a rare, nonsmall cell neoplasm that resembles primitive medullary epithelial cells, and EBL is a rare embryonal neoplasm with cells committed to ependymal differentiation in association with small cells[55]. ETANTR used to be recognized as a subtype within the category of CNS PNETs (having mixed histological features of ependymoblastoma and neuroblastoma), and it had occurred more frequently than MEPL and EBL[56,57]. The prognosis of ETANTR is poor, with 76% of of children dying of disease at a median survival of only 9 months[58].

Elucidation of the molecular biology behind ETANTR began with the identification of high-level amplification of a microRNA cluster on chromosome 19 at position 19q13.42 (which includes the oncomirs miRNA-372 and miRNA-373) in a single patient[59]. A subsequent study determined that focal genomic amplification at 19q13.42 occurred in 37 of 40 ETANTR/EBL samples (93%), with the 3 outliers likely being misdiagnoses, hinting at the relation between EBL and ETANTR[60]. Following the reclassification of ETANTR into the ETMR subgroup, the same group determined that LIN28A (which encodes a protein that binds small RNA molecules, and is implicated in a number of cellular processes including cellular metabolism and tumorigenesis) immunoexpression was detected in 100% of ETMR samples analyzed (n = 37)[61].

Korshunov et al. recommended to reclassify MEPL, EBL and ETANTR as a single cancer (ETMR) due to the identification of uniform molecular signatures in a cohort of 97 cases, regardless of the unique cellular histology present in the three types of cancer[47]. It was recommended that overexpression of LIN28A and amplification of the 19q13.42 gene locus is sufficient for the diagnosis of ETMR. This will be further discussed in the following section[47]. Having only been recognized as a unique clinical entity by Korshunov et al. in 2014, the available literature on ETMRs is limited with few studies focused on the underlying molecular genetics[47]. However, recent work has determined the presence of a plausible oncogenic driver (a unique isoform of DNMT3B) that is a overexpressed as a result of the 19q13.42 gene locus amplification and a unique chromosomal fusion.

1.3 - Embryonal Tumors with Multilayered Rosettes

1.3.1 - Prognosis, Diagnostic Markers and Cytogenetic

Embryonal tumors with multilayered rosettes (ETMRs) are extremely rare cancers; their prevalence is not currently accurately known, although it has been estimated that fewer than 100 cases have occurred since its first identification in the literature in 2000 (not including patients that were misdiagnosed)[61]. For those who do receive the diagnosis of ETMR (almost exclusively infants), the prognosis is bleak, with the vast majority of affected infants dying within 24 months of diagnosis despite the intense therapeutic combination of radiation and chemotherapy[2,47,61].

ETMRs are characterized by several diagnostic protein markers, including the presence of lin-28 homolog A and B (LIN28A/B), and the absence of oligodendrocyte lineage

transcription factor 2 (OLIG2)[62]. As several other types of brain cancers frequently harbor decreased or altered expression of OLIG2, including CNS PNETs, it is not helpful to look at OLIG2 histology when classifying a specific grade IV embryonal brain cancer as an ETMR[50]. LIN28A immunostaining, however, is found to be recurrently and exclusively present in ETMRs and not in other types of pediatric CNS malignant tumors[61].

However, as discussed before, overexpression of LIN28A is integral in the diagnosis of ETMRs. LIN28A is part of a pair of Lin-28 homologues (LIN28A and LIN28B) which are capable of repressing *let-7* miRNA and influencing mRNA translation (indirectly regulating the self-renewal of mammalian embryonic stem cells)[63]. The RNA-binding protein Lin-28 was first identified in the nematode *C. elegans* while screening for lineage-modifying genes, along with the first identified microRNAs *let-7 and lin-4*[64]. Both the *let-7* and *lin-4* miRNAs directly repress *lin-28* to suppress cellular differentiation in different cellular lineages[64]. In addition to this occurring in *C. elegans*, mammalian homologues of Lin28 were found to be highly expressed in mouse embryonic stem cells (ESCs) and expression was shown to decrease upon cellular differentiation[63,65]. Several studies followed, and they demonstrated the ability of Lin28 to inhibit *let-7* maturation in embryonic stem cells (allowing embryonic stem cell self-renewal to occur)[61]. It does so through binding to pre-*let-7* RNA, inducing uridylation at the pre-miRNAs 3' end and preventing processing by the ribonuclease Dicer in ESCs (causing the pre-*let-7* RNA to be degraded over time)[66]. LIN28B is localized primarily in the nucleolus, preventing processing of primary-*let-7* while LIN28A is mainly in the cytoplasm where it can recruit Tut4 to oligo-uridyylate pre-*let*-7 and prevent processing by Dicer[67]. This two way inhibition demonstrates the importance of a switch in LIN28A expression when going from an immature stem-cell state to a more differentiated lineage.

Overexpression of LIN28A is theorized to be oncogenic through repression of *let-7* function, allowing upregulated translation of *let-7* target mRNA (mRNA), which includes known oncogenes such as *MYCN*, *RAS, CDK6* and *HMGA2*[68]. It was shown by Korshunov et al. in 2012 that there is a strong association between expression levels of *LIN28A,* and *MYCN, NRAS, CDK6* and *HMGA2* in a set of 13 ETMR samples[61]. Another group who established a novel ETMR cell line in 2013 (known as BT183), has also demonstrated that LIN28A expression is positively correlated with the activation of the oncogenic mTOR pathway and inversely correlated to *let-7* expression [69]. Furthermore, tumors with dispersed and high levels of LIN28A immunoexpression

correlate with very poor overall survival (an estimated 5-year overall survival of 0%), while other malignant pediatric brain cancers with absence or focal positivity of LIN28A have a 5-year overall survival of 68%[61]. LIN28B is also highly expressed within ETMR samples, but it's diagnostic capability is limited due to the fact that it is overexpressed in other embryonal brain cancers including medulloblastoma[70]. Further investigation in the unique role LIN28A plays in the formation of ETMRs is necessary for understanding the biogenesis of this aggressive cancer.

Cytogenetics have been key in the recent progress that has resulted in defining ETMRs as a unique clinical entity. Amplification of 19q13.42 gene locus was initially detected using two-color interphase fluorescence in situ hybridization (FISH) in a set of patient samples that would now be classified as ETMR patients[60]. This hallmark cytogenetic aberration is now known to occur in almost all ETMR samples, and gives clinicians the unique opportunity to classify a malignant brain cancer solely based on cytogenetics. Amplification of 19q13.42 is associated with the upregulation of two miRNA clusters, C19MC and miR-371-373[71,72]. Following detection of the amplification of 19q13.42 gene locus with FISH in a subgroup of CNS PNETs, one group used differential PCR to detect the amplification of specific miRNAs; *miR-371*, *miR-372* and *miR-373* from the *miR-371-373* locus, and two miRNAs with oncogenic potential from C19MC (*miR-517c* and *miR-520g*) present in this gene locus[71]. Constant overexpression of *miR-517c* and *miR-520g* has been shown to promote oncogenesis both *in vitro* and *in vivo* and cause significant dysregulation of the WNT pathway[72]. Although cytogenetics is key in the classification of ETMR patients, it was not known until recently that a specific gene fusion between one of microRNA clusters in the 19q13.42 gene locus (C19MC) and a relatively unknown brain specific chloride ion channel gene (*TTYH1*) is in part responsible for the enhanced expression of oncomirs within this cluster[2].

1.3.2 - The TTYH1-C19MC fusion

In 2013, Kleinman et al. made one of the most breakthrough discoveries in the field of embryonal brain cancer research, specifically for the subgroup of ETMRs, by identifying a recurrent gene fusion between the promoter of Tweety family member 1 (*TTYH1*), a relatively unstudied neural tissue specific chloride-ion channel, and C19MC in 100% of a set of ETMR samples analyzed (n = 12)[2]. The breakpoints for this fusion were unique to each sample, although all breakpoints fell 5.5-16kb upstream of the C19MC cluster and downstream of the *TTYH1*

promoter, forming a new chimeric gene in which the *TTYH1* promoter drives expression of miRNAs within C19MC[2]. This was further confirmed through demonstrating that the expression levels of C19MC miRNAs were significantly increased in these ETMR samples. Small RNA sequencing allowed the group to determine that expression levels of C19MC miRNAs were 150-1,000 times higher in ETMRs than in control CNS PNET samples[2]. As *TTYH1* expression was found to be high in embryonic stem cells and in the adult brain, it would make sense that canonical activation of its promoter would drive irregular overexpression of miRNAs from C19MC, whose expression is highly regulated and generally restricted to early embryonal development. This not only gives insight into the oncogenesis of ETMRs, but why they occur specifically in the brain.

Following this, the 450K DNA methylation array was used (as miRNAs are known to have effects on epigenetics and can alter DNA methylation patterns) to analyze this cohort of ETMRs, and it was determined that ETMRs harbor a highly specific global DNA methylation pattern when compared to other CNS PNETs and non-cancerous brain samples from various ages[2]. Because of this unique DNA methylation profile, the transcriptional profile of genes affecting DNA methylation were subsequently looked at; it was identified that significant overexpression of *DNMT3B*, a *de novo* (capable of establishing a new methylation profile within cells) DNA methyltransferase, is present in ETMRs when compared to CNS PNETs and other pediatric brain tumors[2]. Of great significance to this study, it was found that transcription of *DNMT3B* occurred not only at its endogenous promoter, but also at a previously uncharacterized alternative promoter (resulting in the inclusion of exon 1B, and the addition of 11 amino acids to the N-terminus of the protein)[2]. This promoter ("promoter 1B") is normally silenced through DNA methylation following fetal development. Inversely correlating with exon 1B-containing DNMT3B expression; demethylation of this promoter locus was found to occur specifically in ETMRs and fetal brain, with endogenous expression occurring in a specific window during embryogenesis (peaking at 8 weeks post-conception)[2].

In an attempt to find the link between C19MC overexpression (as a result of the *TTYH1*-C19MC fusion and amplification) and overexpression of exon 1B-containing DNMT3B, targets of miRNAs from C19MC were examined. It was determined through bioinformatics algorithms that RBL2 (a protein known to repress expression of DNMT3B) is likely to be targeted by several miRNAs in C19MC with identical miRNA seed sequences to miRs found in the miR-

290-295 miRNA cluster, that have previously been shown to target *RBL2* mRNA[2]. Therefore, *RBL2* and *DNMT3B* expression levels were examined in ETMR and control PNET samples through RNA-seq and qPCR; protein expression was determined by immunohistochemical staining. *RBL2* transcript levels were found to be significantly lower in ETMRs compared to PNETs, and protein expression was completely absent (while being detected in all PNETs examined)[2]. Stable overexpression of several miRNAs from C19MC in PFSK-1 (a control PNET cell line) led to the identification of a couple plausibly oncogenic miRNAs (miR-519a and miR-512-3p) by causing a decreased expression of RBL2 and an increase in DNMT3B expression[2].Overexpression of exon 1B-containing DNMT3B was not observed in this specific experiment, which might be due to the cell line that was used, because this cell line (PFSK-1) harbors methylation at promoter 1B of DNMT3B). These findings led to the hypothesis that overexpression of miRNAs from the C19MC miRNA cluster, as a result of the 19q13.42 gene locus amplification and *TTYH1*-C19MC gene fusion, causes a targeted downregulation of RBL2 and in turn causes the usage of an alternative DNMT3B promoter (promoter 1B) in neonatal development. The overexpression of this unique DNMT3B isoform (exon 1B-containing DNMT3B) is believed to cause epigenetic reprogramming in neural development and lead to the formation of ETMRs.

1.3.2.1 - The C19MC miRNA Cluster

The chromosome 19 microRNA cluster (C19MC) is the largest miRNA cluster within the mammalian genome, spanning roughly 100kb and containing 46 miRNA genes[73]. The miRNAs present within this cluster are specific to primates, and conserved within humans[74]. This cluster is highly regulated in terms of expression, as it is located within imprinted genes and is only expressed from the paternally inherited chromosome (the CpG rich promoter region on the maternal chromosome is methylated, preventing transcription from occurring)[74]. The expression of C19MC is generally restricted to the reproductive system and placenta, peaking at 8-10 weeks after conception, at which point it begins to be silenced through DNA methylation (with some C19MC members retaining expression in human embryonic stem cells)[2,74,75].

The oncogenic effects of two miRNAs present within C19MC (miR-520g and miR-517c) have been looked at both *in vitro* and *in vivo*, and it was shown that stable expression of these miRs leads to enhanced cell growth and transformation in both CNS PNET and medulloblastoma

cell lines[72]. In addition, it was demonstrated through Illumina BeadChip arrays that miR-520g expression correlates with irregular WNT pathway signaling in human neural stem cells[72]. Although the functions of C19MC are for the most part unknown, expression of miR-520g and four other C19MC miRNAs decreases with human embryonal stem cell differentiation, showing that expression of miRs from this cluster is likely required to maintain a stem cell-like state[76].

Amplification of C19MC, including the TTYH1-C19MC chromosome 19 fusion leads to a marked overexpression of miRNAs within this cluster in ETMR patients. Several of the miRNAs within this cluster have roles in cellular development that can drive proliferation, promote cell survival and increase the propensity of these cells to drive tumorigenesis[72]. The hypothesized role of exon 1B-containing *DNMT3B* in the formation of ETMRs is further cemented by the fact that the expression of *TTYH1*, C19MC and exon 1B-containing *DNMT3B* peaks at 8-10 weeks following conception[2]. As fetal development (which starts at this time) relies on the use of undifferentiated stem cells, it is striking that overexpression of this tightly regulated miRNA cluster and tightly regulated *DNMT3B* isoform would both be present in a highly specific embryonal brain cancer with stem cell-like characteristics. Thus, dysregulated overexpression of C19MC (which is usually restricted to fetal development) is likely capable of epigenetic reprogramming in neuronal cells (mediated through exon 1B-containing DNMT3B), causing a "re-awakening" of an early embryonic state and oncogenesis.

1.4 - microRNAs and Cancer

1.4.1 - Oncogenic miRNAs (OncomiRs)

MicroRNAs (miRNAs) are 18-22 nucleotide-long noncoding strands of RNA that are capable of regulating the expression of genes encoding proteins[77]. To date, there are more than 2000 individual mature miRNAs identified in the human genome, each capable of altering the expression of several target proteins[78]. Each mature miRNA is believed to target up to several hundred genes through its "seed" sequences (nucleotides 2-7 in the mature strand) that are complementary to sequences which are predominantly in the 3' untranslated region (3'UTR) of the target mRNA; base pairing is also known to occur in the central region (bases ~9-12)[79,80]. Target mRNA is then degraded by the RISC complex, preventing translation of the target protein from occurring.

Oncogenic miRNAs are those that tend to be overexpressed in malignancies, and generally cause enhanced signaling in known oncogenic pathways through inhibition of tumor suppressor genes. MiR-21 is a well known oncogenic miRNA, it is overexpressed in the majority of human cancers, and is strongly overexpressed in glioblastomas multiforme[81,82]. It causes an inhibition of apoptosis, targeting mRNA that encodes proteins in both the intrinsic and extrinsic apoptotic pathway which would normally lead to caspase activation[82]. MiR-21 targets several tumor suppressor genes such as *PTEN*, programmed cell death 4 (*PDCD4*) and tropomyosin 1 (*TPM1*)[78]. As previously discussed, the C19MC microRNA cluster contains hypothesized oncomiRs (specifically miR-519a and miR-512-3p) which are currently believed to be the cause of ETMR formation in infants through their targeted downregulation of RBL2 and subsequent increased expression of exon 1B-containing DNMT3B[2].

There are several oncogenic miRNAs implicated in breast cancer, including miR-21. Others frequently overexpressed in breast cancer include miR-155 (a target of *caspase-3* and *FOXO3a*, promoting cellular survival), miR-182 (an inducer of epithelial to mesenchymal transition through targeting *RECK*), miR-10b (a downregulator of the *HOXD10* tumor suppressor pathway, contributing to invasion and metastasis), miR-27a (an inhibitor of apoptosis by targeting *FOXO1*) and miR-9 (a negative regulator of E-cadherin and inducer of *VEGF*, aiding in metastasis and angiogenesis, respectively)[83]. miR-10b is of particular note, as it is specifically overexpressed in cancer cells that have the ability to metastasize[84]. It is regulated by the transcription factor Twist, which promotes cell mobilization and tissue reorganization during the early stages of embryogenesis[84,85]. Lentiviral transduction and immunohistochemical analysis has demonstrated that miR-10b is capable of promoting invasion and metastasis *in vivo*[100].

Another well studied group of oncogenic miRNAs includes the miR-17-92 cluster (miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1 and miR-92-1), also designated oncomiR-1, which are highly expressed in a number of solid and hematological malignancies, including medulloblastoma[84,86]. The expression of this cluster is driven by the overexpression of c-Myc, and it has been demonstrated that downstream targets of this cluster induces expression of c-Myc, causing an oncogenic positive feedback loop[87]. Targets of these OncomiRs include *P21* (inhibiting cell cycle arrest that would normally be triggered by DNA damage), the proapoptotic gene *BCL2L11* and the E2F family of transcription factors (which are capable of inducing apoptosis)[87]. A specific chromosomal translocation in lymphoma cells is responsible for both

overactivation of oncomiR-1 and increased Notch1 signaling (a transmembrane protein that can promote signaling in the anti-apoptotic Pi3K and pro-proliferative ErbB2 pathways)[84]. OncomiR-1 is also known to protect lymphocytes from apoptosis following the depletion of growth factors, facilitating oncogenesis[84].

The oncomiR miR-155 is overexpressed in a breast, lung, colon, pancreatic and thyroid malignancies[88]. It targets the tumor suppressor gene, suppressor of cytokine signaling 1 (*SOCS1*)[84]. Interestingly, a point mutation in the 3'-UTR region of *SOCS-1* seen in breast neoplasms has been found to allow *SOCS1* to avoid repression from miR-155; in addition, knockdown of *SOCS-1* parallels the oncogenic effects of miR-155, further solidifying it's role as an oncomiR[84].

Lastly, miR-372 and miR-373 have been identified as oncomiRs in testicular germ cell tumors. They promote oncogenesis in germ cells that specifically harbor overexpression of the oncogenic RAS protein and wild-type p53[89]. These miRNAs are thought to disable the p53 pathway through inhibition of the cell cycle regulator CDK2, likely through inhibition of the large tumor suppressor homologue 2 (*LATS2*), causing the cells to be resistant to the tumor suppressive properties of p53[90].

1.4.2 - Tumor Suppressor miRNAs (Anti-oncomiRs)

Loss of miRNA expression in cancer is well documented, signifying the tumor suppressive functions of many miRNAs in humans, and their ability to target oncogenes. The first link between miRNAs and cancer was found in chronic lymphocytic leukemia (CLL) cells which had a marked downregulation of miR-15a/miR-16-1[91]. In one study, reduced expression of these miRNAs was present in 68% of CLL cases analyzed[91]. This is caused by a single nucleotide polymorphism (SNP) in the flaking sequence of the 3' region of miR-16-1 (likely indicating the presence of a nearby enhancer)[92]. These two miRNAs target oncogenes involved in cell cycle regulation (*CCND1*), apoptosis (*BCL2*) and others, which causes cells to proliferate and have extended survival when the expression of these tumor suppressors is lost[93].

The human *let-7* miRNA family (which is comprised of 12 closely related miRs) is the most well studied miRNA in development and cancer; downregulation of these miRNAs is associated with elevated RAS expression in lung cancer[94]. They directly downregulate the expression of K-RAS and N-RAS, known proto-oncoproteins, that are capable of stimulating

uncontrolled proliferation in cells when overexpressed[95]. *Let-7* is also antagonistic towards the well studied oncogene *MYC*, and other cell cycle regulatory proteins such as *CDK6*, *CCND1*, *CCND2* and *CDC25A* (loss of this miRNA will thus lead to uncontrolled G1 to S transition in the cell cycle)[96]. There are also a large number of tumor suppressor miRNAs implicated in human breast cancer; the most notable being the *let-7* family, which canonically targets *HRAS* and *HMGA2*, maintaining the stem cell-like properties of breast cancer when lost[83]. *Let-7* miRNA also targets several genes in the actin cytoskeleton pathway, aiding in invasion and metastasis when absent[83]. Other miRNAs implicated in breast cancer include miR-145 (a suppressor of RTKN, a protein that normally inhibits cell growth and induces apoptosis), miR-205 (a negative regulator of epithelial to mesenchymal transition by targeting *ZEB1* and *ZEB2*), miR-335 (which targets the known oncogene *BRCA1*) and miR-19a (a miR that is downregulated in tumor associated macrophages, allowing them to transition from a pro-immune phenotype to an immunosuppressive phenotype)[83].

The miR-29 miRNA family is a family of anti-oncomiRs that have been shown to reduce cell growth and induce apoptosis in primary acute myeloid leukemia (AML) cells; this is done through targeting of the anti-apoptotic *Mcl-1* gene[97]. This family is found to be downregulated in in several cancers, specifically hematological malignancies including chronic lymphocytic leukemia (CLL) and AML[98]. Another group of microRNAs found to be deleted or genetically inactivated in both hematopoietic and solid cancers is the C14MC microRNA cluster, suggesting it is the largest tumor suppressor cluster in the mammalian genome[74].

The miR-34 family of miRNAs (miR-34a, miR-34b and miR-34c) are found at abnormally low levels in neuroblastoma, and reintroduction of these members into neuroblastoma cell lines has shown to cause a significant decrease in cellular proliferation (through triggering of the intrinsic apoptotic pathway)[99]. Regulation of miRNA expression by transcription factors plays a key role in their ability to transform cells. The well known tumor suppressor p53 directly recognizes the promoters, and activates transcription, of the miR-34 miRNA family, which themselves control p53-meidated cell cycle arrest and apoptosis[100]. In addition, miR-200c is also controlled by p53, and it's expression has been shown to reduce stemness properties in breast cancer by targeting *BMI1*[101].

The RISC associated protein Dicer (and its recruitment protein TRBP) are responsible for processing pre-miRNA into a mature miRNA duplex. TRBP is a phosphoprotein, and the

required phosphorylation for it's activation is dependent on kinases within the MAPK pathway, a signal transduction pathway that promotes cell division[84,102]. Activation of TRBP through the MAPK pathway enhances processing of pre-miRNA molecules; pro-growth and proliferation miRNA profiles have been demonstrated in cells with enhanced MAPK signaling[110]. In this case, the anti-oncomiR *let-7* family is downregulated, while the oncomiRs miR-17, miR-20a and miR-92a are upregulated[110]. This demonstrates that changes in the microRNA processing pathway can lead to downstream effects (altering the production of both oncomiRs and anti-oncomiRs) that cause oncogenesis, sugguesting that a balance between these two types of miRs is required for cellular homeostasis.

1.5 - RBL2

1.5.1 - The RB Family of Tumor Suppressors

The retinoblastoma (Rb) family of proteins is made up of three proteins; retinoblastoma 1 (RB1/p105), retinoblastoma-like 1 (RBL1/p107) and retinoblastoma-like 2 (RBL2/p130), all of which are structurally similar, but function differently depending on cellular context[103]. These proteins are known for their roles in cell cycle regulation. They first unknowingly came to attention in 1971 when Alfred Knudson proposed his "two-hit" hypothesis on the formation of retinoblastoma. In hereditary retinoblastoma, one mutation is inherited in an autosomal dominant manner through germline cells, while the second occurs in somatic cells (affecting the other allele of the mutated gene)[104]. However, it was later discovered that the gene being inactivated in hereditary retinoblastoma was p105, now known as RB1, which at the time was only known to be a nuclear phosphoprotein involved in cell growth regulation[105]. RB1 is now known to be dysregulated in numerous human cancers (including SCLC and breast cancer), either through inactivating mutations or through altered function of upstream regulatory proteins[103,106,107].

The Rb family of proteins regulate the cell cycle by inhibition of the G1-S phase transition through their ability to interact with the E2F family of transcription factors[103]. RB1 binds the transcriptional activators E2F1, E2F2 and E2F3a, inhibiting them directly, while RBL1 and RBL2 recruit the transcriptional repressors E2F4 and E2F5 (respectively) to the nucleus to inhibit transcription indirectly[108]. Rb proteins are themselves regulated through their state of phosphorylation, which is controlled through cyclin-dependent kinase (CDK)-cyclin

complexes[109]. Phosphorylation renders Rb proteins incapable of binding E2F transcription factors, causing progression of the cell cycle[103]. When cells are in a quiescent state, hypophosphorylated Rb associates with E2F to inhibit transcription of target genes[110]. When mitogenic signaling occurs, sequestration of CDK4 and CDK6 by INK4 proteins (encoded by $p16^{INK4A}$, $p15^{INK4B}$, $p18^{INK4C}$ and $p19^{INK4D}$) is released through phosphorylation of these cyclin depending kinases on their T-loop by CDK activating kinases (in this case, CDK7)[111,112]. This allows CDK4/6 to interact with Cyclin D, allowing them to phosphorylate RB1, RBL1 and RBL2, causing E2F transcription factors to express genes required for G1 to S phase transition and DNA synthesis[113].

The human RB1 protein is 928 amino acids and contain three main domains, the central domain is responsible for binding viral oncoproteins like adenovirus E1A, SV40 large T antigen and human papilloma virus E7; it is nicknamed the "pocket" domain[114]. The pocket domain also contains a conserved region known as the 'LxCxE' binding cleft, which serves as a docking point for chromatin-remodelling and histone-modifying enzymes[114]. It has two subdomains, A and B, which hold the domain together through their non-covalent interactions[114]. The N-terminal domain (RBN) resembles the central pocket domain, save for a few subtle differences[115]. The C-terminal domain (RBC) consists of the final 150 residues, and it lacks a tertiary protein structure[114].

RB1 shares approximately 25% of its sequence with both RBL1/p107 (107kDa) and RBL2/p130 (130kDa), while the two retinoblastoma-like proteins share approximately 54% sequence identity with each other[114]. The structural data on both RBL1 and RBL2 is limited, but sequencing has revealed that both proteins contain an RBN-like domain (107N and 130N), pocket domains, and carboxy-terminal domains (107C and 130C)[114]. The pocket domain of all of these proteins serves as a way to interact with E2F transcription factors; the conserved LxCxE binding site in this domain not only binds viral proteins, but also cellular adhesion proteins like Cadherin-1[114].

RB1 is expressed throughout the cell cycle, while RBL1 and RBL2 are expressed during different phases of the cell cycle; RBL2 expression is higher in the G0 phase, while RBL1 expression is highest during the S phase[113,116]. Expression of these proteins is also specific to the type of cell, as RBL2 is primarily expressed in neurons and skeletal muscle and RBL1 expression is significantly high in breast and prostate epithelial cells[113]. RB1 is universally

expressed in all normal tissues[113]. RBL2 is also known to act as a direct inhibitor of CDK2, mediated by its unique amino acid sequence present in its spacer region[117]. In addition to cell cycle regulation, RB proteins are known to aid in preservation of chromosomal stability, regulate apoptosis, induce angiogenesis, affect cellular differentiation and cause cellular senescence[103,118–120].

1.5.2 - Biological Function and Clinical Significance

As stated previously, the primary role of the RB family of proteins is to regulate cell cycle progression at G1-S phase transition. This is carried out through their ability to either repress the transcription of E2F family target genes that are pro-proliferative, or activate transcription of anti-proliferative E2F family target genes. It is no surprise that when these proteins are unable to do their job, cells can proliferate at a concerning rate leading to carcinogenesis. The tumor suppressive roles of RB1, RBL1 and RBL2 are further cemented by their roles in regulating genomic stability, their ability to cause cellular senescence, their interaction with p53, their antiapoptotic properties, their ability to promote cellular differentiation and their antimetastatic capabilities[114].

RB1 loss in wild type adult murine hepatocytes has been shown to increase chromosomal number and cause cells to re-enter the cell cycle[121]. Triple knockout (TKO) of the three Rb family of proteins in mouse embryonic fibroblasts (MEFs) under serum starvation results in completion of S phase but arrest in G2; these TKO MEFs accumulate DNA damage while arrested in G2 that is not completely repaired following the addition of mitogens and continuation of proliferation[122]. Rb proteins are also required for the maintenance of constitutive heterochromatin (localized at pericentric regions and telomeres) through maintenance of the histone 4 lysine 20 tri-methylation (H4K20me3) post-translational modification[123]. The absence of these proteins leads to defects in telomere elongation and chromosomal segregation during mitosis, leading to chromosomal instability (CIN).

In senescent cells, RB1 exists only in its hypophosphorylated (active) form, and re-expression of RB1 in p53 and RB1-deficient breast carcinoma cell lines has been shown to cause growth arrest, inhibition of telomerase, and expression of the senescence marker SA-β-gal[124,125]. The tumor suppressor p16 (encoded by *CDKN2A*) is a positive regulator of RB1. It inhibits the binding between cyclin D and CDK4/6 preventing the phosphorylation (inactivation) of RB1[126].

Aberrant expression of p16 in human prostate cancer cell lines can induce senescence in cells with wild type RB1, but not RB1 null cells[127].

DNA damage induced senescence in multiple p53-expressing human tumor cell lines and fibroblasts relies on RBL2, as an increase in hypophosphorylated RBL2 (dependent on functional p53 and p21) but a decrease in RB1 and RBL1 was seen in senescent cells[103,128]. It was also demonstrated that in order to prevent S phase progression, RBL2 induces senescence through binding to the promoter of the gene that encodes cyclin A, inhibiting its transcription[129]. RBL2 also plays a primary role in senescence of breast cancer cells in response to DNA damage through inhibition of cyclin B and CDK1 in a similar manner, with RB1 and RBL1 compensating in its absence (demonstrated through *RBL2* knockdown in ZR-75 cells)[130]. Furthermore, RBL2 has an oncosuppressive role *in vivo* by mediating senescence in response to loss of RB1; this was proven by the fact that loss of both copies of *RB1* in cells of the developing retina does not cause immediate progression to retinoblastoma, but forms a benign precursor lesion that expresses high levels of p16 and RBL2[103,131].

Inactivation of RB1 through mutation, viral oncoprotein sequestration or caspase-dependent degradation correlates with higher levels of apoptosis (programmed cell death)[103,132]. Dephosphorylation of RB1 carried out by phosphatase 1α (PP1α), is connected to triggering apoptosis in leukemia cell lines[133]. However, the role of RB proteins in apoptosis does vary depending on cellular context, as *RB1*-deficient prostate cancer cells are resistant to radiation induced apoptosis (while wild-type *RB1* cells underwent apoptosis) and RBL2 represses the transcription of the antiapoptotic protein Survivin through promoter interaction[134,135].

The role of RB1 in promoting cellular differentiation throughout embryonic development and in adult cells is well defined in mouse models and *in vitro*[120]. It promotes differentiation in several cell lineages by binding to tissue-specific transcription factors, general transcription factors, chromatin remodelling proteins and inhibitors of differentiation such as ID2 and EID1[120,136]. Sporadic retinoblastomas with loss or reduced expression of RBL2 demonstrate a lower degree of differentiation[137]. The presence of RBL1 and RBL2 is also associated with cyclin D3 expression in muscle fibers, which is a protein known to be upregulated during myogenesis, facilitating cell cycle exit and muscle specific gene expression[138] .

Angiogenesis is the acquisition of vasculature by a neoplasm in order to supply required nutrients and facilitate metastatic invasion. RB1 can inhibit angiogenesis through its prevention

of transcription of E2F1 target genes which include the proangiogenic genes *FGF2* and *VEGFB*[139]. In addition, RBL2 overexpression causes a reduction of vestigial endothelial growth factor (VEGF) mRNA and protein levels both *in vivo* and *in vitro*[140].

RB1 and its family members RBL1 and RBL2, alongside the Cip/Kip family of CDK inhibitors (which are known to regulate the timing of neuronal progenitor cell cycle exit) also help regulate neuronal migration and differentiation during CNS formation[116]. As the CNS develops, neuronal progenitor cells both migrate and differentiate, once this process is complete they exit the cell cycle and enter the post-mitotic G0 phase, losing their proliferative potential[116]. The RB family is required for regulating the entrance of differentiated neurons into the post-mitotic G0 phase, while the Cip/Kip family of CDK inhibitors is responsible for timing progenitor cell cycle exit and initiation of neuronal differentiation (in progenitor cells of the retina and cerebral cortex)[116]. This is clinically significant, as neuronal progenitor cells of the cerebral cortex and retina that harbor mutations in RB family proteins initiate differentiation without exiting the cell cycle, resulting in tumor formation[116].

Many oncoproteins from small DNA viruses are known to interact with RB1, disrupting its interaction with E2F, causing an increase in cellular proliferation[141]. This includes oncoproteins from human papillomavirus E7 (HPV-E7), Adenovirus EIA (Ad-EIA), and Simian virus 40 large T antigen (SV40-LTa)[141]. These all share the conserved LxCxE motif that binds to the pocket domain of RB1, RBL1 and RBL2, aiding in their inactivation[141]. This is clinically relevant as the development of molecules that bind to LxCxE (inhibiting viral interaction) could be important in preventing viral oncogenesis.

In order for cancer cells to spread throughout the body, they must be able to breach through the basement membrane that encapsulates the primary tumor (in the case of solid cancers), survive migration in the blood stream, invade target tissue and thrive in their new cellular environment; this process is known as metastasis[142]. RB1 expression has been shown to be absent in nearly half of all metastatic lesions in a study done on hepatocellular carcinoma patients, with expression levels in the primary samples being markedly higher[143]. Further evidence was provided in a mouse model study through a triple conditional knockout of RB1, RBL2 and p53. This study demonstrated that knocking out these three tumor suppressive cell cycle regulators resulted in a high incidence of small cell lung cancer (SCLC) metastasis to the liver, kidney and pulmonary lymph nodes[144]. This is particularly interesting as this level of

metastasis was not present in double RB1/p53 knock-out mice, implying that RB1 and RBL2 work in tandem to suppress metastasis[144].

1.5.2.1 - Suppression of RBL2 by miRNAs

The tumor suppressor RBL2 is inhibited by a number of oncogenic miRNAs in cancer, and by regulatory miRNAs in healthy somatic cells. The most significant finding related to embryonal brain cancer was that miRs present within the C19MC miRNA cluster are now validated as targets of RBL2[2]. As previously discussed, ETMRs are hypothesized to be caused by the amplification of miRNAs present within C19MC (specifically miR-519a, and miR-512-3p) through its fusion and amplification with the promoter of *TTYH1*, which then target *RBL2* (decreasing its translation) and subsequently leading to a global change in DNA methylation due to the ability of *RBL2* to alter expression levels of *DNMT3A* and *DNMT3B*[2]. However, suppression of *RBL2* by miRs is not limited to those present in C19MC. Expression levels of the microRNA miR-106b-5p have also been demonstrated to be significantly upregulated (20-fold) in glioma tumor samples compared to healthy brain[145]. MiR-106b-5p targets both *RBL1*, *RBL2*, and the proapoptotic gene *CASP8* through their 3'-UTR at the posttranscriptional level, causing irregular proliferation and inhibition of apoptosis, thereby significantly contributing to gliomagenesis[145]. The upregulation of miR-106b-5p has been demonstrated to dysregulated glioma cell proliferation and apoptosis both *in vitro* and *in vivo*[145].

The well studied miR-17-92 miRNA cluster is overexpressed during adipocyte differentiation, allowing the differentiation of these fat cells to occur at a more accelerated level (this was demonstrated through hormone induction in stably transfected 3T3L1 cells; differentiation occurred much faster in cells transfected with miR-17-92 compared to vector only cells)[146]. Overexpression of the miR-17-19 cluster in 3T3L1 cells (through stable transfection) has also revealed its role in targeting RBL2[146]. RBL2 protein levels were nearly undetectable 24h following transfection, and RBL2 mRNA levels were 2.0 fold lower in cells transfected with the miR-17-19 cluster compared to control cells[146]. This highlights the fluid role of RBL2 on cellular differentiation depending on cellular context.

Kaposi's sarcoma-associated herpesvirus (KSHV) encodes a cluster of 12 miRNAs encoded by a single transcript[147]. One of these viral miRNAs, miR-K12-4-5p targets RBL2;

irregular expression of miR-K12-4-5p has been shown to reduce RBL2 protein expression and increase levels of DNMT1, DNMT3A and DNMT3B mRNA relative to control cells[147].

The tumor suppressor p53 acts as a transcription factor, in addition to being a cell cycle regulator protein. One of the genes it binds to and activates is the *panthothenate kinase 1* (*PANK1*) gene, and its intronic miRNA miR-170[148]. Recruitment of p53 to the *PANK1* promoter is triggered by DNA damage, causing transcription of miR-170; which has been shown to target both *CDK6* and *RBL2*[148]. This demonstrates the complex role of RBL2 in controlling cell cycle progression, as its protein levels are indirectly decreased by p53 in response to DNA damage.

Several members of the miR-290 miRNA cluster bear identical seed sequences to the previously mentioned miRNAs present in C19MC that were demonstrated to target RBL2[2]. The miR-290 cluster is restricted to placental mammals and encompasses miR-290, miR-291-3p, miR-291-5p, miR-292-3p, miR-292-5p, miR-293, miR-294, and miR-295[149,150]. This cluster has been shown to target *Rbl2* post-transcriptionally. Transfection of the miR-290 cluster or single transfection of miR-291-3p, miR-292-3p, miR-294 or miR-295 causes a decrease in *Rbl2* expression to undetectable levels in mouse C2C12 cells[150]. This caused an increase in global DNA methylation and DNA methylation at subtelomeric regions, inferring that RBL2 controls the expression levels of the *de novo* DNA methyltransferases DNMT3A and DNMT3B[150].

1.5.3 - Role of RBL2 in DNMT3B Suppression

All three active members of the mammalian DNA methyltransferase (DNMT) family (DNMT1, DNMT3A and DNMT3B) have shown to be regulated in some part by members of the Rb protein family in the past. The *DNMT1* gene is a target of E2F transcription factors, and is irregularly overexpressed in the absence of RB1, leading to hypermethylation of DNA in prostate epithelial cells[151]. As mentioned prior, reduction of RBL2 protein levels as a result of introducing the KSHV miR-K12-4-5p in 293 cells results in an increase in DNMT1, DNMT3A and DNMT3B mRNA levels relative to control cells (and resulted in an increase in DNA methylation at both viral and cellular CpG loci)[147].

The ability of RBL2 to negatively regulate DNMTB was first addressed in a study by Blasco et al., in 2008. They demonstrated that ablation of the miRNA processing protein Dicer has been associated with a significant repression of DNMT1, DNMT3A and DNMT3B in mouse embryonic stem cells, resulting in a decrease in DNA methylation[150]. The mammalian specific

miRNA, miR-290 (previously described) is notably downregulated in murine *Dicer1*-null cells, resulting in increased levels of RBL2[150]. Re-introduction of the miR-290 cluster (which targets *Rbl2*) causes an increase in expression of DNMT3A and DNMT3B in mouse C2C12 cells, demonstrating its ability to inhibit the expression of these two DNA methyltransferases and indirectly control DNA methylation[150]. Similarly, another study demonstrated that knockdown of *Rbl2* in Dicer null cells had a positive effect on *Dnmt3a2* and *Dnmt3b* expression, causing them to retain an undifferentiated morphology[152].

Lastly, decreased expression of RBL2 in ETMR patient samples (as a result of overexpression of miR-519a, and miR-512-3p from the C19MC miRNA cluster) is believed to cause upregulated expression of both DNMT3B and a unique variant of DNMT3B transcribed from an upstream promoter (exon 1B-containing DNMT3B) that would normally become hypermethylated in the absence of these genetic changes (specifically the fusion between the promoter of *TTYH1* and C19MC)[2].

1.6 - DNMT3B

1.6.1 - DNA Methylation

Epigenetics refers to heritable alterations in the structure of DNA and chromatin that are not due to changes in the primary sequence of DNA nucleotides. The only known epigenetic modification of DNA itself in differentiated mammalian cells is the covalent attachment of a methyl group to the C5 position of cytosine nucleotides (5mC), usually in CpG dinucleotide sequences, hereby referred to as DNA methylation[153]. DNA methylation was first addressed in 1975 when methylation of cytosine residues in CpG dinucleotides were identified as a plausible epigenetic mark in vertebrates[154]. Since then, it has been discovered that more than half of the genes in the vertebrate genome contain CpG-rich regions that average around 1kb in length; these are known as CpG islands (CGIs)[155]. The lack of CGIs in the other half of the mammalian genome is believed to be caused by conversion of 5mC to thymine by spontaneous and enzymatic deamination over time in previously methylated regions[155]. Methylation of CGIs significantly correlates with gene silencing, as it causes DNA condensation and prevention of transcription factor recruitment at promoters[155]. In addition, silencing of gene transcription is also

achieved by DNA methylation at enhancers, through reducing activation of target genes by the enhancer that is methylated (demonstrated in the past through reporter gene assays)[155].

The presence of CGIs at the transcriptional start sites of genes allows for epigenetic transcriptional regulation of these genes through DNA methylation. The majority of CGIs are unmethylated in somatic cells, with transcriptional activation usually being regulated by nucleosome concentration (euchromatin has a low density of nucleosomes and favors transcription), histone PTMs (H3K4me3 is highly responsible for activation of transcription) and transcription factors[155]. DNA methylation can not only regulate gene expression spatially, but temporally, as *de novo* DNA methylation can cause inactivation of genes whose CGIs were previously hypomethylated. Methylation of CGIs in promoters can cause long-term transcriptional repression through prevention of transcription initiation (by preventing recruitment of transcription factors and formation of heterochromatin); this is seen in imprinted genes and genes silenced during X chromosome inactivation[155].

DNA methylation is not limited to CGIs, most gene bodies are extensively methylated, even though they are CpG-poor[155]. Methylation in exons is the leading cause of C to T transition mutations, which can lead to disease and cancer when occurring in germ and somatic cells, respectively[156]. CGIs can also be present within gene bodies, although their function in this context is unknown[155]. Contrary to methylation at transcription start sites, gene body methylation is associated with active transcription, as the 5mC mark blocks initiation but not elongation[155]. It has been hypothesized that DNA methylation is stimulated by the process of transcription, as the H3K36me3 PTM could recruit DNMTs to gene bodies[157]. Methylation in gene bodies has also been implicated in gene splicing, and controlling alternative promoter usage[158,159]. Methylation is also known to occur at enhancers (regions of DNA outside of a gene that regulate transcription) and insulators (sequences of DNA that block the effect of enhancers on genes). Enhancers are the most differentially methylated regions (DMR) in the context of oncogenesis; demonstrating that changes in their epigenetic patterns can greatly influence the transcriptome and behavioral properties of cells[160].

Demethylation of 5mC residues is not as well studied as the process of methylation, however a few studies have demonstrated this to be a natural occurring process. It was demonstrated that in adult brain, hydroxylation of 5mC by the TET1 enzyme (converting it to 5-hydroxymethylcytosine) followed by conversion to 5-hydroxymethyluracil (a recently identified

epigenetic modification of DNA, caused by oxidation of 5mC, that allows DNA enhanced flexibility and hydrophilicity) with the aid of AID/APOBEC deaminases ultimately leads to base excision repair replacing 5mC with a regular cytosine[161,162]. It was further discovered that thymine DNA glycosylase (TDG) is required for base excision repair of both 5mC and 5hmC[163]. In addition, absence of the ten-eleven translocation 3 (TET) methylcytosine dioxygenase causes a failure of CpG demethylation in genes that regulate embryonic development (such as *Oct4* and *Nanog*) on the paternal chromosomes, delaying embryogenesis[164]. Further inquiry into the mechanism of DNA demethylation is an ongoing area of research.

There are currently several techniques available to analyze patterns of DNA methylation. Prior to analysis of genomic DNA for the presence of 5mC marks, DNA must undergo multiple steps of processing, the major one being bisulfite conversion. The goal of bisulfite conversion is to allow methylated cytosine residues (5mC) to be differentiated from unmethylated cytosine residues during the final step of DNA methylation analysis (microarray analysis)[165]. The whole procedure of bisulphite conversion consists of numerous consecutive structural modifications; DNA must first be denatured to single stranded DNA (ssDNA) using high temperatures, it is then incubated with $NaHSO_3$ at pH 5.0 (causing deamination and the addition of a bisulphite group at C6 on cytosine residues), followed by incubation of the DNA at a high pH (triggering desulphonation)[165]. This process results in the conversion of unmethylated cytosine residues into uracil nucleotides, while methylated cytosine residues are not susceptible to bisulfite conversion and remain intact[165]. Uracils are amplified in subsequent PCRs as thymines while 5mC residues are amplified as cytosines, and comparison of sequenced bisulfite-treated DNA with reference genomic data allows DNA methylation patterns to be deduced. This can be done with the Infinium HumanMethylation450 BeadChip microarray (450k array) which contains 485,512 probes covering 99% of genes in The NCBI Reference Sequence (RefSeq) collection and 96% of CGIs present in the UCSC database[166,167]. The bisulphate converted DNA can also be sequenced with newly developed high-throughput sequencing (whole-genome bisulphate sequencing). This method is more costly but has the advantage to detect genome-wide 5-methylcytosine.

1.6.2 - DNA Methyltransferases

There are currently three separate families of DNA methyltransferase genes known to be present in mammalian cells: DNMT1, DNMT2, and DNMT3[168]. DNMT1 is known as the

maintenance methyltransferase, in that it catalyzes DNA methylation in cells undergoing mitosis, maintaining current patterns of DNA methylation[169]. DNMT2 is generally considered to be catalytically inactive when it comes to methylating DNA, but recent studies have demonstrated that DNMT2 is a tRNA methyltransferase that is required for proper differentiation of hematopoietic cells[170]. The DNMT3 family consists of two *de novo* methyltransferases (which are capable of establishing new methylation marks in DNA), DNMT3A and DNMT3B, and a third catalytically inactive homologue which increases the catalytic ability of these *de novo* methyltransferases, DNMT3L[171].

DNMT1 is localized to the replication fork during DNA replication, and methylates newly synthesized DNA so that it retains the 5mC marks present on the parent strand[171]. It is also capable of differentiating between hemimethylated and unmethylated DNA, allowing it to methylate the unmethylated complementary strand of the prior[171]. The protein itself is ~1620 amino acids in length; it contains a C-terminal methyltransferase (MTase) domain (~500 amino acids) and a N-terminal regulatory region (~621 amino acids) that consists of the proliferating cell nuclear antigen (PCNA) binding domain, a replication *foci* targeting sequence (RFTS) domain, a CXXC zinc finger domain and a two bromo adjacent homology (BAH) domains[171,172]. The MTase domain of DNMT1 shares a common structure with those found in DNMT3A and DNMT3B, known as "AdoMet-dependent methyltransferase" domains[171]. It is required for cofactor binding (through motifs I and X) and substrate catalysis (using motifs IV, VI, and VIII)[171]. The PCNA domain binds facilitates DNMT1-PCNA interaction; PCNA is a protein required in DNA replication, it helps maintain the separation of DNA strands at the replication fork[171]. The RFTS domain inhibits DNMT1 from binding substrates through interacting with the MTase domain, while the zinc finger and BAH domains aid in its ability to bind to DNA[172]. A linker sequence between the CXXC and BAH domain also plays a role in in DNMT1 autoinhibition, through both the CXXC and RFTS domains[172]. The activity of DNMT1 is also known to be regulated by both the CDK inhibitor p21 and the family of Rb proteins[201].

DNMT2, now referred to as tRNA aspartic acid methyltransferase 1 (TRDMT1) is a short (391 amino acid) protein that catalyzes the methylation of cytosine-38 in the anticodon loop of aspartic acid tRNA[173]. It has also been thought to be involved in recognizing damaged DNA, mutation repair and DNA recombination[174].

DNMT3A and DNMT3B are 912 amino acids and 853 amino acids, respectively[175]. They are responsible for establishing DNA methylation patterns early on in mammalian development and in germ cells through *de novo* methylation; they do not show a preference between hemimethylated and unmethylated DNA[175]. They also help maintain the heterochromatin in pericentric regions and telomeres through DNA methylation[175]. These two structurally related proteins contain three main domains; the C-terminal MTase domain (which is smaller than the one present in DNMT1, but contains the same MTase motifs), and the N-terminal proline-tryptophan-tryptophan-proline (PWWP) and ATRX-DNMT3-DNMT3L (ADD) domains[175,176]. The PWWP domain helps facilitate chromatin bonding, it interacts non-specifically with DNA in DNMT3B but has shown lower DNA-binding activity in DNMT3A[177]. The ADD domain, also found in ATRX (a histone H3.3 histone chaperone and chromatin remodelling protein), is a cysteine-rich domain that contains a C2-C2 zinc finger and an atypical PHD finger; it allows these proteins to interact with transcription factors (such as Myc), chromatin modifying enzymes (SUV39H1, SETDB1 and EZH2) and histone H3 when it bears the H3K4me0 mark[175,176]. DNMT3A and DNMT3B are crucial enzymes for mammalian development, as DNMTA null mice die within a month of being born and DNMTB null mice show embryonic lethality at E14.5-E18.5[171].

DNA methyltransferase 3-like (DNMT3L) is 421 amino acids, and encodes a catalytically inactive homologue of DNMT3A and DNMT3B[178,179]. It has a similar amino acid sequence to DNMTA and DNMT3B but is missing the catalytic C-terminal region and retains only the ADD domain[176]. It is known to simulate the DNA methylation activity of DNMT3A an DNMT3B, and plays an important role in germ cell development[180]. In cells that express *DNMT3L, de novo* methylation by DNMT3A2 (a splice variant of DNMT3A) is catalyzed by a tetrameric complex consisting of two molecules of DNMT3A and DNMT3L, and requires the presence of a nucleosome (as DNMT3L targets this tetramer to H3K4me0)[178]. DNMT3B also interacts directly with DNMT3L, forming a dimeric complex that greatly influences the effects of *de novo* DNA methylation[181].

1.6.3 - Effects of DNA Methylation on Developmental Processes

The initial pattern of DNA methylation is established at around the time of embryonic implantation and maintained throughout development[175]. *De novo* methylation, carried out by

DNMT3A and DNMT3B, is able to be established following fertilization as the zygote undergoes global DNA demethylation in order to create totipotency in the embryo[175]. A second wave of *de novo* methylation happens during germ line development, when primordial germ cells segregate into different paths[175]. Furthermore, DNA methylation is important in X chromosome inactivation (the silencing of one copy of the X chromosome in females to prevent twice as many gene products from forming) and imprinting (silencing of genes specific to one parent in gametes)[1]. DNMT3A and DNMT3B are both highly expressed in early embryogenesis, in both embryonic tissues and undifferentiated embryonic stem cells; expression of these proteins is inversely correlated with differentiation[175].

Methylation patterns of human DNA are created, in part, due to the flaking sequence preferences of DNMT3A and DNMT3B[182]. In addition, the DNMT3L-$(DNMT3A)_2$-DNMT3L complex has been shown to favor methylation of CpG sites separated by 8 to 10 base pairs, furthering the case that sequence specificity is partly responsible for how *de novo* methyltransferases target DNA[181]. Chromatin conformation and histone PTMs have also demonstrated influence on where DNMT3A and DNMT3B target DNA. These proteins read histone modifications through their N-terminal domains, and target nucleosomes with H3K4me0 and H3K36me3 marks[175]. H3K36me3 accumulates in the gene bodies of active genes present in euchromatin, while H3K4me3 is required for active transcription. A strong correlation between DNA methylation, absence of H3K4me3 and the presences of H3K36me3 has been demonstrated in the past, showing an epigenetic influence on the methylation pattern of cells[183]. *De novo* DNA methylation is incapable of occurring in nucleosomes bearing the H3K4me3 mark, and the higher the level of gene expression from a transcription start site, the less likely it is for a CGI to become methylated[155,184]. Furthermore, genes with CGI promoters that are already silenced through histone PTMs (such as H3K9me3 and H3K27me3) mediated by Polycomb complexes are much more likely to become methylated than other genes[185]. DNMTs can also be targeted to specific methylation sites through interacting with other proteins known to bind chromatin and DNA[175]. This includes the transcription factors Myc, p53 and RAR, and the chromatin modifying proteins EZH2 (the H3K27 trimethyltransferase present in the PRC2 complex), HP1 and SUV39H1[175,186,187].

As stated previously, both DNMT3A and DNMT3B are required for viable offspring in mice, demonstrating how important they are in embryogenesis and development[171]. DNMT3B

has been shown to have a key role in hypermethylation of CGIs proximal to promoters[188]. The high expression of these *de novo* methyltransferases in early embryonic cells is essential for determining the epigenetic landscape of DNA, and any changes in their expression can lead to drastic alterations of the transcriptome. The maintenance of these 5mC marks is carried out by DNMT1 in cellular division, due to its ability to target hemimethylated DNA and localize to the replication fork during cell division[205]. It should be noted that both DNMT3A and DNMT3B participate in maintenance methylation, and the presence of DNMT1 alone is not sufficient to maintain methylation in newly synthesized DNA[189]. However, DNMT1 null mouse embryos are also embryonic lethal, demonstrating that these enzymes are not capable of solely carrying out maintenance methylation[190]. The required presence of the three enzymatically active DNMTs during embryogenesis emphasizes their importance in shaping the genetic makeup of mammalian cells.

1.6.4 - DNA Methylation in Cancer

Genomic stability and basal gene expression levels are largely maintained by a fixed and predetermined pattern of DNA methylation[191]. Irregular patterns of DNA methylation can lead to changes in transcription of important regulatory genes, causing irregular progression of the cell cycle, resistance of cell death, genomic instability, angiogenesis and grant cells the ability to metastasize. DNMT1, DNMT3A and DNMT3B have all been shown to be overexpressed in tumors, suggesting that DNA methyltransferases are proto-oncogenes whose overactivity causes genome wide reprogramming that leads to a pro-oncogenic transcriptome[171]. Overexpression of the DNMT1 mouse variant in NIH3T3 cells results in cellular transformation, and human fibroblasts with constant overexpression of DNMT1 leads to hypermethylation of CGIs in a time-dependent manner[171]. Reduction of DNMT1 levels through gene deletion in HCT116 (a colon cancer cell line) has shown to result in slower growth and lower levels of genomic methylation[192]. Aberrant expression of these proteins has been detected in colon, pancreatic, breast, kidney and other cancers[171].

DNMT1 and DNMT3B are capable of changing different histone modifications present in colon cancer cells, and overexpression of DNMT3B1 in $APC^{Min/+}$ mice (a mouse model carrying a point mutation in the *APC* gene, which results in multiple intestinal neoplasia) enhances the formation of these neoplasms[193]. Overexpression of DNMT3B1 also causes tumor

suppressor gene methylation[193]. DNMT1 also functions in the maintenance of colon cancer stem cells (CSCs) and colon cancer initiating cells (CICs) through causing an increased expression of CD44 and CD24[171].

Abnormal overexpression of DNMT1 is associated with the development of pancreatic cancer (from normal tissue, to precancerous lesions, to cancer), and higher DNMT1 expression correlates with lower rates of overall survival[194]. Increased levels of DNMT3A and DNMT3B expression has also been found to correlate with overall lower survival rates in pancreatic cancer patients[195]. These results are attributed to an overall increase in DNA hypermethylation, which has been observed in early- and late-stage human pancreatic cancer lesions[171].

DNMT1, DNMT3A and DNMT3B are all overexpressed in breast cancer as well, with DNMT3B overexpression being present in 30% of patients[171,196]. Expression of *DNMT3A* and *DNMT3B* has also been found to correlate with expression of *JUNB* and *EREG* (both responsible for activation of cellular growth pathways), respectively[225]. Interestingly, the expression of the most common DNMT overexpressed in breast cancer (DNMT3B) was also the most variable in terms of range of expression (being expressed nearly 82 times higher in one patient) demonstrating that the level of overexpression of DNMTs influences their catalytic activity in cancer cells[225].

DNMT1 and DNMT3B have also showed 3-fold expression in kidney cancer cells, compared to control tissue, and was assocaited with DNA hypermethylation[197]. A significant overexpression of these two DNMTs has also been observed in gliomas, and was found to correlate with decreased expession of the PTEN and p21 tumor suppressors (through transcriptional silencing of these genes by CGI methylation)[198]. Hepatocellular carcinoma shows a progressive increase in methylation during cancer development; this was the result of increased expression of DNMT1, DNMT3A and DNMT3B, and was predictive of a low overall survival rate[168]. Upregulation of these three proteins has also been demonstrated in AML cells and CML cells in the acute phase of cancer progression[199]. Aside from the active DNMTs, DNMT3L overexpression has been detected in testicular germ cell tumors, and was found to be required for the growth of human embryonal carcinomas[200].

Inhibition of DNMT activity is associated with reduced tumorigenicity and a decrease in methylation of tumor suppressor genes[171]. Several catalytic inhibitors of DNMTs have been developed in order to treat cancer, with the main ones being 5-azaC, 5-azaCdR and zebularine[171].

These are nucleoside analogues of cytidine, which when incorporated into DNA prevent the enzymatic activity of DNMTs, stopping DNA methylation from occurring, and have been approved for use in myelodysplastic syndromes[171,201]. Further development of DNMT inhibitors in clinical settings will likely result in improved treatment capabilities by physicians in the future.

1.6.5 - *Oncogenic Effects of DNMT3B*

As mentioned prior, all of the catalytically active DNMTs are implicated in various forms of cancer. Due to the *de novo* nature of DNMT3B, its overexpression is particularly harmful when it comes to epigenomic reorganization. The altered expression of DNMT3B in cancer is ironically sometimes due to differential methylation at its promoter; this has been observed in glioma cell lines and tumor tissue samples, when compared to normal brain tissues[198]. DNMT3B depletion (and not DNMTA) has been shown to induce apoptosis in human cancer cells specifically, which outlines its importance in cancer cell survival[202]. Its oncogenic capability is largely attributed to its ability to methylate and inactivate tumor suppressors; this has been demonstrated to be the case in several cancers, as Nanaomycin A (an inhibitor of DNMT3B that binds to its MTase domain) selective inhibition of DNMT3B reactivates previously silenced tumor suppressor genes in colon, lung and bone marrow cancer cell lines[203].

Several polymorphisms of DNMT3B have been implicated in different types of cancer. A C → T transition in a novel promoter of DNMT3B, 149bp upstream from the transcription start site increases the transcription of DNMT3B, and is associated with an increased risk of lung cancer[204]. A T → C transition, located 283bp upstream from the start site, upregulates expression of DNMT3B and correlates with susceptibility to lung cancer[205]. The C46359T polymorphism of DNMT3B has also been shown to be statistically significant in leaving women more prone to developing cervical cancer; these are just some examples of polymorphisms in DNMT3B that leave people more susceptible to developing cancer[206].

Upregulation of DNMT3B expression has been detected in endometrial cancers, specifically those with a poorly differentiated cellular phenotype[207]. *DNMT3B* is the most frequently overexpressed DNMT gene in breast cancer as well, having a 30% prevalence and showing over 80 times basal expression levels in some cases[225]. It is also significantly overexpressed in human melanoma, and has been shown to trigger uncontrolled proliferation

through stimulation of the mTOR pathway through inhibition of mTORC2 signaling[208]. Although there are a plethora of examples in which overactive DNMT3B is pro-oncogenic, it has shown to possess tumor suppressive qualities in preventing the development of lymphoma in mice[209].

Finally, DNMT3B has the ability to regulate other epigenetic marks through protein interaction. For example, DNMT3B has exhibited the ability to have an effect on the creation of histone PTMs mediated by PRC1, influencing the susceptibility of genes that are hypermethylated in colon cancer[210]. Although it is unknown exactly why high levels of DNMT3B cause it to target specific sites (specifically the promoters of tumor suppressor genes) that are not methylated at a basal level, it is clear that overexpression of this protein leads to epigenetic changes that leave cells prone to undergoing oncogenic transformation.

1.6.5.1 - ICF Syndrome

Immunodeficiency, centromere instability and facial anomalies (ICF) syndrome is the only human disease that follows an autosomal recessive inheritance pattern of immunological dysfunction and aberrant DNA methylation[211]. It is diagnosed by lack of gamma globulin in blood plasma (causing immune deficiency) and DNA arrangements at pericentric heterochromatin of chromosomes in lymphoid cells undergoing mitogenic stimulation[211]. Observation of facial anomalies is present in most but not all patients, and consists of a broad flat nasal bridge, increased distance between the eyes, epicanthic folds (a skin fold of the upper eyelid, covering the the medial canthus of the eye), a high forehead, low set ears and tongue protrusion[211]. This is the only human disease known to be associated with mutations in a DNA methyltransferase gene[212].

ICF syndrome is usually diagnosed in infants and children, with the majority of patients succumbing to the disease prior to adolescence; it is an extremely rare disorder, with less than 50 cases having been presented since its discovery in the 1970s[212]. ICF syndrome is primarily caused by biallelic mutations in *DNMT3B*, usually in the C-terminal portion of the protein, which contains it catalytic domain[212]. This results in a hypomorphic function of the DNMT3B enzyme in lymphocytes of patients harboring these mutations, causing hypomethylation in DNA sequences of juxtacentromeric regions in chromosomes 1, 9 and 16[213]. The lack of methylation in these juxtacentromeric regions leaves them prone to chromatin decondensation, resulting in chromosomal rearrangements, whole-arm deletions and translocations[212].

As lymphocytes of ICF syndrome patients undergo mitogenic stimulation, chromosomal instability occurs due to the lack of methylation in pericentric heterochromatin, leading to multibranched configurations with different numbers and combinations of arms from chromosomes 1, 9 and 16[211,214]. This theory of chromosomal instability being a result of lack of DNA methylation was cemented by an experiment that showed treating a normal pro-B lymphoblastoid cell line with inhibitors of DNA methylation (5-azacytidine or azadeoxycytidine) caused similar chromosomal rearrangements to those seen in ICF syndrome patients[215].

Microarray analysis of patient and control lymphoblastoid cell lines (LCLs) revealed dysregulation of genes with ICF-specific dysregulation that encode proteins involved in lymphocyte migration, activation and survival[211]. Irregular transcript levels of these genes causes these B lymphocytes to mature incorrectly, producing irregular amounts of immunoglobulins[211]. The resulting lack of humoral immunity present in these patients leaves them susceptible to lethal infections. ICF syndrome stresses the importance of DNA methyltransferase genes (specifically DNMT3B) in controlling the epigenome of cells to prevent widespread transcriptional dysregulation and chromosomal instability; which is also demonstrated by the fact that complete inactivation of *DNMT3B* is embryonic lethal in mice[1].

1.6.5.2 - Exon 1B-containing DNMT3B

DNMT3B is required for several events in embryonic differentiation, through its ability to methylate genomic DNA and change the transcriptome of cells. Exon 1B-containing DNMT3B expression is restricted to a very narrow developmental window in fetal neuronal development, at around 8 weeks following conception[2]. The promoter of this DNMT3B variant is normally surrounded by heterochromatin in adult human brain, but is found to be demethylated in fetal brain and ETMRs (allowing transcription of exon 1B-containing *DNMT3B* to occur)[2]. Transcription of this specific variant of DNMT3B was only recently discovered in an analysis of a cohort of 12 ETMR samples[2]. Methylation of promoter 1B was found to correlate with gestational age and a decrease in *DNMT3B* exon-1B production[2]. Given that *DNMT3B* promoter 1B is only hypomethylated in fetal brain and ETMR samples (causing this unique isoform of DNMT3B to be exclusively expressed when this occurs) and the unique methylation pattern present in ETMR patient samples, exon 1B-containing DNMT3B is believed to significantly contribute to the formation of ETMRs[2].

1.7 - Project Rationale, Hypothesis and Objectives

Embryonal tumors with multilayered rosettes (ETMR) are a group of deadly embryonal central nervous system tumors that primarily affect infants. These tumors have been characterized by the amplification of a microRNA (miRNA) cluster located at chr19q13.41 (C19MC), which is a genetic hallmark of the disease, being present in up to 95% of samples tested[2,69]. Our lab recently identified that overexpression of this miRNA cluster is driven by the amplification of a gene fusion between the promoter of *TTYH1* (a gene whose expression is restricted to the brain) and C19MC, and that ETMRs have a highly specific global DNA methylation pattern[2]. Overexpression of the C19MC miRNA cluster is proposed to cause a targeted downregulation of retinoblastoma-like protein 2 (RBL2), a tumor suppressor that regulates the expression of the *de novo* DNA methyltransferases DNMT3A and DNMT3B[2]. This repression of RBL2, in turn, is believed to cause an altered (increased) expression of DNMT3B. It was further identified that a previously uncharacterized, embryonic brain development specific, isoform of DNMT3B (exon 1B-containing DNMT3B) was significantly overexpressed in ETMRs[2]. This unique isoform of DNMT3B is transcribed through the usage of an alternate promoter of this gene (promoter 1B) that is normally silenced by DNA methylation in both fetal and adult brain tissues[2]. A schematic summarizing the genomic organization of DNMT3B in various tissues, as well as the overexpression of exon 1B-containing DNMT3B due to the *TTYH1*-C19MC fusion in ETMRs can be seen in Figure 1.

Overexpression of the unique DNMT3B variant, exon 1B-containing DNMT3B (DNMT3B V6), in embryonal tumors with multilayered rosettes (ETMRs) is believed to be due to the absence of its inhibition by RBL2 as a result of its targeted downregulation by miRNAs that were overexpressed due to the amplification of a gene fusion between the promoter of *TTYH1* and C19MC[2]. Furthermore, methylation of the promoter that drives expression of DNMT3B V6 (exon 1B-containing DNMT3B) usually occurs after 8-10 weeks of gestation, and remains present throughout development as cells differentiate thereby silencing expression of this *de novo* methyltransferase[2]. Given the undifferentiated cellular phenotype of ETMRs, their unique DNA methylation profile and consistent overexpression of exon 1B-containing DNMT3B (DNMT3B V6), we believe that this unique DMMT3B variant is responsible for the epigenomic reorganization that occurs in ETMRs, resulting in oncogenesis (see Figure 1).

We therefore hypothesize that overexpression of exon 1B-containing DNMT3B plays an important role in the development of ETMRs, via a global reprogramming of DNA methylation within the pediatric brain, leading to a reestablishment of an early embryonal epigenetic landscape, and the formation of ETMRs.

In order to prove this, we set out to overexpress DNMT3B V6, along with the canonical DNMT3B variant (DNMT3B V1) as a control, to demonstrate that it is responsible for the formation of ETMRs. I therefore will generate 3rd generation VSV-G pseudotyped lentiviral particles capable of stably transducing cells with the coding sequences of DNMT3B (DNMT3B variant 1) and exon 1B-containing DNMT3B (DNMT3B variant 6). Transfection (nucleofection) of standard vectors containing the coding sequences of DNMT3B V1 and V6 will be performed in tandem as well. This will allow me to transduce and transfect cell lines likely to model ETMRs, specifically the primitive neuroectodermal tumor (PNET) cell line PFSK-1, with the coding sequences for these DNMT3B isoforms, in order to assay for phenotypic changes and global changes in DNA methylation. Confirmation of overexpression will be attempted through western blot, quantitative PCR (PCR), and immunofluorescence (IF) labelling analysis. Following overexpression of DNMT3B V6 (and plausibly DNMT3B V1) we expect to see a change in global patterns of DNA methylation within transformed cells that cluster with those found in ETMR samples, and a change in cellular morphology and oncogenic properties that mimic those found in ETMR cells.

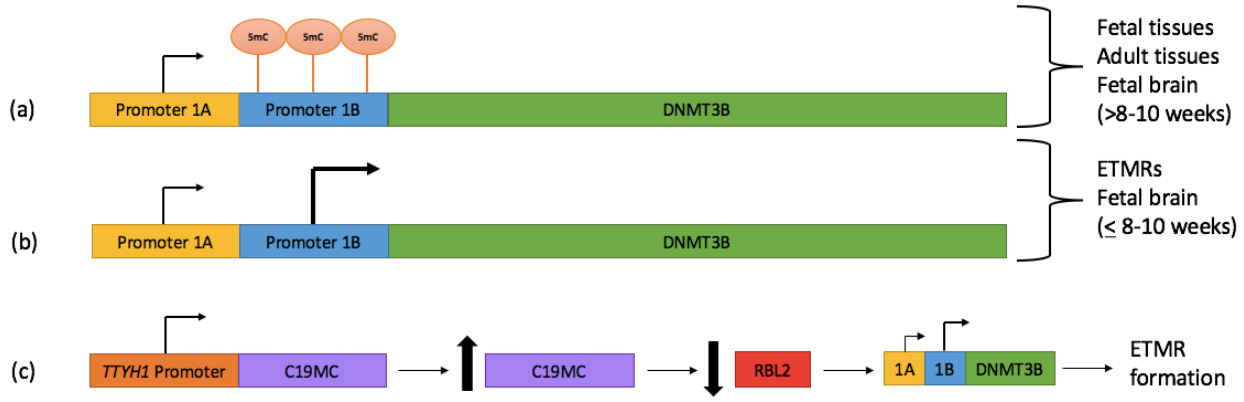

Figure 1 (a, b and c) - DNMT3B genomic organization and proposed model of ETMR tumorigenesis

(a) Promoter 1B of DNMT3B is methylated in fetal tissues, adult tissues and fetal brain 8-10 weeks following conception. This results in a closed conformation of chromatin surrounding this promoter, preventing transcription of exon 1B-containing DNMT3B from occurring. Only the canonical DNMT3B variant is transcribed from promoter 1A, which has exon 1B spliced out prior to translation. (b) Promoter 1B of DNMT3B remains unmethylated up to 8-10 weeks following conception in fetal brain. It is only otherwise found unmethylated in embryonal tumors with multilayered rosettes (ETMRs). When unmethylated, it allows for transcription of the unique DNMT3B variant exon 1B-containing DNMT3B (in addition to the canonical DNMT3B variant). (c) In our proposed model of ETMR tumorigenesis, the fusion between the promoter of TTYH1 and C19MC causes an increase in transcription of miRNAs from C19MC. Several of these miRNAs target the tumor suppressor RBL2 (known to inhibit DNMT3B expression), decreasing its expression. This results in the absence of methylation at promoter 1B of DNMT3B and an increase in expression of both DNMT3B and exon 1B-containing DNMT3B. The overexpression of exon 1B-containing DNMT3B specifically is believed to cause a significantly altered DNA methylation profile and retention of an early embryonic epigenetic state in pediatric brain, leading to ETMR formation.

Chapter 2 - Determining if Exon 1B-Containing DNMT3B is Responsible for the Formation of ETMRs

2.1 - Introduction

Embryonal tumors with multilayered rosettes (ETMRs) are an extremely rare form of childhood embryonal brain cancer originally thought to be a type of central nervous system primitive neuroectodermal tumor (CNS PNET)[47]. Since their initial discovery, there have been as few as 75 cases published in the literature, with survival times ranging between 24 and 36 months[47,54]. These uncommon and deadly cancers are characterized by the amplification of the 19q13.42 locus and the LIN28A protein expression[47]. More recently, it was discovered that nearly all of these cancers carry a gene fusion between the promoter of *TTYH1* (a brain specific chloride-ion channel) and the chromosome 19 microRNA cluster (C19MC), the largest microRNA cluster present in the human genome[2].

Amplification of this chromosomal fusion results in aberrant overexpression of microRNAs present in C19MC, specifically miR-519a, and miR-512-3p, that have been shown to cause a targeted downregulation of retinoblastoma-like 2 (RBL2)[2]. RBL2 is a member of the retinoblastoma tumor suppressor family, and has exhibited the ability to downregulate expression of the *de novo* DNA methyltransferase DNMT3B in the past[150]. Due to the decreased expression of RBL2 in ETMR cells (as a result of C19MC miRNA amplification), levels of DNMT3B are abnormally high[2]. Furthermore, absence of methylation was detected on an upstream promoter of DNMT3B (promoter 1B), and correlated with increased expression of a unique variant of DNMT3B (exon 1B-containing DNMT3B). This unique variant will hereby be referred to as DNMT3B variant 6 (V6), and the canonical splice variant will be referred to as DNMT3B variant 1 (V1). In addition, ETMR samples carry a unique DNA methylation profile that clusters separately from CNS PNET, fetal brain, pediatric brain and adult brain control samples[2].

In summary, amplification of the gene fusion between *TTYH1* and C19MC is believed to cause downregulation of RBL2, which in turn causes dysregulated overexpression of DNMT3B V6 through use of the unmethylated DNMT3B promoter 1B. This overexpression of DNMT3B V6 is hypothesized to be the cause of ETMR oncogenesis, the unique DNA methylation profile seen in ETMR samples, and responsible for the distinct transcriptional profile seen in ETMR patients (such as high expression levels of LIN28A)[2]. Therefore, we believe that overexpression

of exon 1B-containing DNMT3B in cells with characteristics similar to those of ETMR cells would result in a global change in methylation and cellular transformation towards an ETMR-like phenotype.

2.2 - Materials and Methods

Cell line characteristics and tissue culture conditions. All cell lines used throughout the experimental process were purchased from ATCC® (with the exception of BT183) and stored in a liquid nitrogen cryogenic tank (-196°C) until use. Cell lines were removed from the cryogenic tank and kept on dry ice (-79°C) prior to undergoing quick thawing in a 37°C water bath (~1min). The contents of the cryogenic tubes were then transferred to a 10mL centrifuge tube (using a 1000µL micropipette) with 10mL of cell line specific media (detailed below). Cells were pelleted in a tabletop centrifuge for 5 minutes at 1000rpm and media was aspirated. The cells were then resuspended in appropriate media and transferred to a vented T75 flask with 15mL (total) of cell line specific media. All cell lines were cultured at 37°C, 5% CO_2. 293FT virus-producing cells were cultured in complete DMEM containing 10% FBS supplemented with 0.1mM MEM non-essential amino acids, 1mM sodium pyruvate and 2mM L-glutamine. 293FT cells were maintained in 500µg/mL Geneticin® 24h after being placed in to culture until transfection. The PFSK-1 primitive neuroectodermal tumor cell line was cultured in RPMI-1640 medium supplemented with 10% FBS. The SF188 adult glioma cell line was grown in EMEM supplemented with 10% FBS. This cell line was previously transformed to stably express the *H3F3A* (H3.3) p.Gly34Arg mutant histone with a Myc-tag at the N-terminus. This transformed variant of SF188 (SF188 H3-G34R-Myc) was also grown in EMEM with 10% FBS, but was maintained in 500µg/mL Geneticin® to retain expression of the mutant histone. The HT1080 fibrosarcoma cell line was cultured in D-MEM containing 10% FBS supplemented with 4.5g/L glucose, 110mg/L sodium pyruvate and 2mM L-glutamine. The BT183 ETMR cell line was seeded onto BD Primaria™ 6-well plates coated with poly-L-ornithine and laminin, and grown in NeuroCult™ NS-A Basal Human Medium (STEMCELL Technologies) supplemented with NS-A proliferation supplement, 10ug/mL rhEGF, 10ug/mL bFGF and 0.2% heparin.

Vector amplification. The pCMV6-Entry Vectors were purchased from OriGene© (PS100001) with DNMT3B (DNMT3B Variant 1) and exon 1B-containing DNMT3B (DNMT3B Variant 6) coding sequences spliced into the multiple cloning site. In addition, the lentiviral destination vector (pLenti-C-Myc-DDK-IRES-Puro) in which the DNMT3B V1 and V6 sequences were cloned into (prior to lentiviral production) was also purchased from OriGene© (PS100069). Vectors were resuspended in 10µL of TE Buffer (pH = 8) to a final concentration of 1µg/µL. In order to amplify plasmid DNA, 2µL of each resuspended plasmid (diluted to 20ng/µL using ddH_2O) was used to transform 50µL of OneShot® Stbl3™ Chemically Competent *E. coli*. Stbl3™ cells were thawed on ice (4°C) for 10min prior to the addition of 2uL of plasmid DNA. The bacteria-plasmid mixture was left on ice for 20min after gentle mixing, and then heat shocked at 42°C for 30 seconds in a water bath before being placed back on ice. 950µL of S.O.C. Medium (Invitrogen™) was added to the transformed bacteria. Tubes were then incubated at 37°C, shaking at 220rpm for 1 hour. 50µL of each mixture was plated onto pre-warmed (37°C) LB-kanamycin (50µg/mL) agar plates (pCMV6-Entry Vectors) or LB-chloramphenicol (34µg/mL) agar plates (pLenti-C-Myc-DDK-IRES-Puro) and incubated overnight at 37°C. Several colonies were picked the following morning and cultured for 7 hours at 220rpm in 5mL of LB-kanamycin (50µg/mL), for the pCMV6-Entry vectors, or 5mL of LB-chloramphenicol (34µg/mL), for the lentiviral destination vectors. Following this, 100µL of each bacterial culture was used to inoculate 35mL of fresh LB-kanamycin (50µg/mL) or LB-chloramphenicol (34µg/mL) in a 250mL Erlenmeyer flask. Plasmid DNA was extracted from bacterial cultures using the QIAGEN® Plasmid *Plus* Midi kit, according to the manufactures instructions (eluting in 200µL of EB). Plasmid DNA was quantified using the NanoDrop 1000 spectrophotometer (Thermo Scientific™) according to the manufacturers instructions. Amplified pCMV6-Entry Vector clones were then sent for sequencing to confirm the presence of the DNMT3B genes. The presence of these sequences was confirmed through Sanger sequencing at the Genome Quebec Innovation Center©; the exact sequences present in each vector and the list of sequencing primers used can be found in Supplementary Data (Table S1 and S3).

Molecular cloning. In order to clone DNMT3B V1 and V6 sequences from the pCMV6-Entry Vectors into the lentiviral destination vector (pLenti-C-Myc-DDK-IRES-Puro), three separate restriction digestion reactions were set up. 5µg of plasmid DNA (pCMV6-Entry Vector

DNMT3B V1, pCMV6-Entry Vector DNMT3B V6 or pLenti-C-Myc-DDK-IRES-Puro) was digested using 70U of MluI, and 70U of SgfI in NEB© Buffer 3.1 for 90 minutes at 37°C (in a total reaction volume of 100µL). 30µL of each pCMV6-Entry Vector digestion was purified using the QiaQuick MiniElute PCR Purification Kit, according to the manufacturers instructions (eluting in 30µL of Buffer EB). 10µL of the lentiviral destination vector digestion was dephosphorylated using 1U of rAPid Alkaline Phosphatase (Roche Rapid DNA Ligation Kit) in a 20µL reaction volume for 10 minutes at 37°C. The alkaline phosphatase was inactivated for 2 minutes at 75°C following dephosphorylation. Two ligation reactions (using the Roche Rapid DNA Ligation Kit) were then set up with the following components:

Component	**Volume (Ligation 1)**	**Volume (Ligation 2)**
Dephosphorylated Vector DNA (pLenti-C-Myc-DDK-IRES-Puro)	2µL	2µL
DNMT3B V1 Insert DNA (purified)	3µL	-
DNMT3B V6 Insert DNA (purified)	-	3µL
DNA Dilution Buffer	2µL	2µL
ddH_2O	3µL	3µL
MIXED		
T4 DNA Ligase Buffer (2X)	10µL	10µL
T4 DNA Ligase	1µL	1µL

The ligation reactions were left at room temperature for 20 minutes before 2µL was used to transform Stbl3™ competent cells. Transformation and vector amplification was performed as previously described (see **vector amplification**) using the QIAGEN® Plasmid mini kit. Each selected colony was spotted on a backup LB-chloramphenicol (34µg/mL) agar plate; 6 clones each were picked for DNMT3B V1 and DNMT3B V6. In order to determine the presence of DNMT3B genes in the lentiviral destination vector clones prior to sequencing, 5uL of plasmid DNA from each clone (mini-prep) was digested with 2U of MluI and 2U of SgfI in NEB© Buffer 3.1 for 90 minutes at 37°C (in a total reaction volume of 20µL). Each digestion was run on a 1% w/v agarose gel for 70 minutes at 100V. The results of this digestion can be seen in supplementary data (see Figure S2). Clones #2, #3, #9 and #10 were sent for Sanger sequencing

at the Genome Quebec Innovation Center© to confirm the presence of DNMT3B V1 (clones #2 and #3) and DNMT3B V6 (clones #9 and #10) in the destination vector (pLenti-C-Myc-DDK-IRES-Puro). All clones sent for sequencing contained the original sequences found in the pCMV6-Entry Vector (sequences can be found in Table S1).

Nucleofection. Nucleofection was carried out using the Lonza Nucleofector™ 2b device (Entry Level Model) and the Lonza Human Stem Cell Nucleofector™ Starter Kit, according to the manufacturers instructions, under the following conditions: PFSK-1 cells were grown in culture for at least 72 hours and two passages prior to nucleofection. PFSK-1 cells were expanded to 150mm x 20mm tissue culture plates and grown to ~80% confluency before being detached with 5mL of Cellstripper™ per plate, and centrifuged for 5 minutes at 1000 rpm. Supernatant was aspirated and cells were resuspended in 1mL of appropriate tissue culture media (without antibiotics). 20µL of cells were stained with trypan blue and the number of cells was determined using a hemocytometer. Three separate aliquots of 1 million cells were then pipetted into 3 x 15mL centrifuge tubes and the supernatant was aspirated following another centrifugation for 5 minutes at 1000 rpm. 2µg of plasmid DNA (pCMV6 Empty Vector, pCMV6 DNMT3B V1 and pCMV6 DNMT3B V6) was then added to each aliquot before resuspending the cells in 100µL of nucleofector solution. Cells were nucleofected using program # A-033 and then immediately transferred to a 6-well plate in a final volume of 1.5mL of appropriate media (without antibiotics). The following day (24h later) media was aspirated and replaced with 2mL of fresh media. Cells were then passaged (48h following nucleofection) to 100mm x 20mm tissue culture plates in 10mL of media containing 300µg/mL Geneticin® (optimum concentration was determined prior to nucleofection by performing a titration kill curve on PFSK-1 with varying concentrations of Geneticin®). Cells were maintained in 300µg/mL Geneticin® until visible clones appeared on the 10cm plates (~2 weeks). The resulting clones were picked using Scienceware® cloning discs and transferred to wells in a 24-well plate; the remaining colonies were passaged together to a fresh 100 mm x 20 mm tissue culture plate to form a single "batch" culture. Cells were allowed to proliferate in 300µg/mL Geneticin® until enough 10 mm x 20 mm plates of each transformed PFSK-1 cell line (pCMV6 Empty Vector, pCMV6 DNMT3B V1 and pCMV6 DNMT3B V6) were available for mRNA and protein extraction.

Lentiviral production. 293FT cells cultured in three 150mm x 20mm tissue culture plates were left to proliferate until ~80% confluency (media was aspirated and replaced the prior evening with media containing no antibiotics). Media was also replaced 2 hours prior to cotransfection. Lentiviral plasmids (pLenti-C-Myc-DDK-IRES Puro Empty Vector, pLenti-C-Myc-DDK-IRES Puro DNMT3B V1 Clone #2 and pLenti-C-Myc-DDK-IRES Puro DNMT3B V6 Clone #9) were diluted to 0.5μg/μL, and 4.5μg (9μL) of plasmid DNA was mixed with 45μL of pPACKH1™ Packaging Plasmid mix (SBI) in 750μL of OptiMEM®. Transfection was carried out using a 3:1 ratio of Lipofectamine® 2000 to plasmid DNA, according to the manufacturers instructions. Cells were incubated overnight and media was replaced the following day with 23mL of fresh media. 48h following transfection, media (viral supernatant) was transferred to a 50mL centrifuge tube, underwent centrifugation for 5 minutes at 500g (4°C) to remove debris and was filtered (0.22μm syringe filter) into a fresh 50mL centrifuge tube; 23mL of fresh media was added to the cells and they were incubated overnight. Viral supernatant was collected the following day using the same procedure and 5.75mL of 5X PEG-*it*™ was added to each 50mL centrifuge tube. The resulting PEG-*it*™ viral supernatant mixture was refrigerated (4°C) for 48h to allow lentiviral particles to precipitate. All six 50mL centrifuge tubes (containing precipitated lentivirus) underwent centrifugation (at 4°C) for 30 minutes at 1500g to pellet the lentivirus. Supernatant was aspirated and the pairs of lentiviral pellets were resuspended and pooled in 400μL of cold (4°C) PBS. Aliquots of 50μL were then transferred to cryogenic tubes and the lentivirus was stored at -80°C until use.

Lentiviral infection. PFSK-1 cells were counted as previously described (see **nucleofection**) and 50,000 cells each were plated in 4 wells of a 24-well plate (in 500μL of appropriate media) the evening prior to lentiviral transduction. The day of transduction, media was aspirated and replaced with the following mixture (containing one of the three types of lentivirus): 50μL of lentivirus and 500μL of media with 8μg/mL of polybrene (vortexed for 8 seconds in a 1.5mL microfuge tube prior to being added to the wells). The fourth well was used as a non-transduced control (only media with 8μg/mL polybrene was added). Cells were incubated overnight. Media was replaced the following day with another fresh aliquot of lentivirus (50μL of lentivirus in 500μL of media with 8μg/mL polybrene). 48 hours following the first infection, cells were passaged to 100mm x 20mm tissue culture plates and incubated overnight. 72 hours following

the first infection, media was replaced with fresh media containing 2μg/mL of puromycin (a concentration determined previously by performing a titration kill curve on PFSK-1 with varying concentrations of puromycin). Transduced PFSK-1 cells (pLenti-C-Myc-DDK-IRES Puro Empty Vector, pLenti-C-Myc-DDK-IRES Puro DNMT3B V1, pLenti-C-Myc-DDK-IRES Puro DMNT3B V6) were maintained in 2μg/mL puromycin for ~3 weeks until all non-transduced cells had died. Cells were passaged in batch culture due to the low rate of proliferation observed, and maintained in 100 mm x 20 mm tissue culture plates (in 2μg/mL puromycin) until enough cells were available for protein extraction and immunofluorescence analysis.

RNA extraction. RNA was isolated from two ~80% confluent 100mm x 20mm tissue culture plates for each nucleofected PFSK-1 cell line, in addition to the non-nucleofected PFSK-1 control, using the Aurum™ Total RNA Mini Kit (according to the manufacturers instructions). Cells were washed with 10mL of room temperature (22°C) PBS prior to extraction and lysed directly on the 100mm x 20mm plates. RNA samples were stored at -80°C prior to qPCR analysis.

qPCR analysis. RNA was quantified using the NanoDrop 1000 spectrophotometer (Thermo Scientific™) according to the manufacturers instructions. All RNA samples were diluted to 50ng/μL prior to the cDNA synthesis reaction. cDNA synthesis was carried out using the iScript™ cDNA synthesis kit according to the manufacturers instructions, with both ddH_2O and no reverse transcription controls. Quantitative determination of *DNMT3B* expression levels (both Variant 1 and Variant 6) was performed using the Ssofast Evagreen kit (BioRad) with standard conditions indicated by the manufacturer, at an annealing temperature of 60°C, on the Roche LightCycler 480 using cDNA from nucleofected PFSK-1 clones and batch culture. Cycle threshold (C_t) values were normalized to β-actin (*ACTB*) using the $2^{-\Delta\Delta Ct}$ method. The following primer pairs were used:

5'-TAACAACGGCAAAGACCGAGGG-3', DNMT3B forward
5'-TCCTGCCACAAGACAAACAGCC-3', DNMT3B reverse
5'-GGCACCCAGCACAATGAAGATCAA-3', β-actin forward
5'-TAGAAGCATTTGCGGTGGACGATGGA-3', β-actin reverse

Protein purification and quantification. Protein was isolated from two ~80% confluent 100mm x 20mm tissue culture plates for each nucleofected and transduced PFSK-1 cell line, in addition to the wild type PFSK-1 controls. Cells were washed with 10mL of cold (4°C) PBS prior to extraction and lysed directly on the 100mm x 20mm plates with 200µL of RIPA buffer supplemented with cOmplete™ Lysis-M tablets. The lysed cells were then pooled into a 1.5mL microfuge tube and left on ice (4°C) for 20 minutes. Tubes were centrifuged (4°C) at maximum speed for 10 minutes and the protein supernatant was transferred to a fresh 1.5mL centrifuge tube for protein quantification. Protein was quantified by preparing a standard curve of BSA using Bio-Rad Protein Assay Dye Reagent, according to the manufacturers instructions. Protein samples were stored at -80°C prior to western blot analysis.

Western blot analysis. Protein samples stored at -80°C were thawed on ice and 20µg of each protein sample was transferred to a fresh 1.5mL microfuge tube. 4X SDS buffer was then added to each sample; protein was denatured for 5 minutes at 95°C before being placed back on ice (4°C). Protein was electrophoresed on a 1.0mm thick 10% acrylamide-gel (4% acrylamide stacking gel) in 1X running buffer for 75 minutes at 100V. Samples were then transferred to a PVDF membrane (activated in methanol for 1 minute prior to the transfer) in cold (4°C) 1X transfer buffer for 75 minutes at 100V. The PVDF membrane was then cut and blocked for 1 hour at room temperature in TBST (5% skim milk). The membrane was then washed in TBST three times (10 minutes each) before the addition of the primary antibody solution (in TBST). Membranes were probed with the following primary antibodies overnight (>16 hours) at the specified concentrations:

DNMT3B (H-230) Rabbit pAb sc-20704 (Santa Cruz Biotechnology) diluted 1 in 200 in TBST (5% skim milk)
Myc-Tag (9B11) Mouse mAb #2276 (Cell Signaling Technology) diluted 1 in 1000 in TBST (5% skim milk)
Myc-Tag (71D10) Rabbit mAb #2278 (Cell Signaling Technology) diluted 1 in 1000 in TBST (5% skim milk)
β-tubulin (ab6046) Rabbit pAb (abcam®) diluted 1 in 1000 in TBST (5% skim milk)

Following primary antibody incubation, membranes were washed in TBST three times (10 minute each) before the addition of the secondary antibody. Membranes were then probed with the following secondary antibodies for 1 hour at room temperature (in addition to Strep-Tactin® at a concentration of 1 in 10,000):

Donkey anti-rabbit IgG-HRP sc-2313 (Santa Cruz Biotechnology) diluted 1 in 5000 in TBST (5% skim milk)
Goat anti-mouse IgG-HRP sc-2005 (Santa Cruz Biotechnology) diluted 1 in 2000 in TBST (5% skim milk)

Following secondary antibody incubation, membranes were washed in TBST three times (10 minutes each) and proteins were then detected through autoradiography with enhanced chemiluminesence using the Pierce™ ECL Plus Western Blotting Substrate, according to the manufacturers instructions.

Immunofluorescence analysis. Lentiviral transduced PFSK-1 cells (along with the non-transduced control) and SF188 G34R-Myc cells were counted as previously described (see **nucleofection**) and plated, in duplicate, in 4-well chamber slides at 10,000 cells per chamber (in 500µL of cell line specific media). Cells were incubated overnight at 37°C, 5% CO_2 in order to facilitate the adherence of PFSK-1. Following overnight incubation, a pre-fixation was performed by adding 500µL of 4% PFA (in PBS) to each chamber. Cells were incubated for 2 minutes at room temperature (22°C) and the media/pre-fixation solution was then replaced with 1mL of ice-cold methanol (-20°C). Cells were then incubated in the freezer (-20°C) for 20 minutes and washed 3 times (5 minutes each) with PBS. 500µL of permeabilization solution was added to each chamber and cells were incubated for 15 minutes at room temperature (22°C). Cells were then washed 3 times (5 minutes each) with PBS and incubated in 500µL of blocking solution for 1 hour at room temperature (22°C). The blocking solution was aspirated and replaced with either a fresh 500µL aliquot of blocking solution (the no primary antibody control well) or 500µL of primary antibody solution. The primary antibody used was Myc-Tag (9B11) Mouse mAb #2276 (Cell Signaling Technology) diluted 1 in 200 (in blocking solution). Cells were then incubated at 4°C overnight (>16 hours). Following overnight incubation, cells were

washed 3 times (5 minutes each) in washing solution. Cells were then incubated in 500µL of secondary antibody solution. The secondary antibody used was Goat anti-Mouse IgG (H+L) Secondary Antibody, Alexa Fluor® 568 conjugate (Thermo Fischer Scientific) diluted 1 in 500 (blocking solution). Cells were incubated at room temperature for 1 hour and then washed 3 times (5 minutes each) in washing solution. Cells were washed one final time in PBS (5 minutes) and then mounted with one drop of ProLong® Gold Antifade Mountant with DAPI (Thermo Fischer Scientific) per chamber on a coverslip. Cells were left to dry in a flat dark (free of light) place overnight and visualized the following day on a fluorescence microscope (using an absorption wavelength of 578nm). The composition of the permeabilization, blocking and washing solutions can be seen below:

Solution	**Components**
Permeabilization solution (5mL)	50µL of 5% Igepal 2.5mL of 2% BSA (in PBS) 250µL of 10X PBS 2.25mL of ddH_2O
Blocking solution (15mL)	7.5mL of 2% BSA (in PBS) 750µL of 10X PBS 6.75mL of ddH_2O
Washing solution (50.25mL)	50mL of 1X PBS 250µL of Tween-20

URLs

IDT OligoAnalyzer 3.1, http://www.idtdna.com/calc/analyzer;

ExPASy Translate Tool, http://web.expasy.org/translate/;

NCBI BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi;

Ensembl, http://useast.ensembl.org/index.html;

UniProt, http://www.uniprot.org/;

EMBL-EBI ClustalW2, http://www.ebi.ac.uk/Tools/msa/clustalw2/.

2.3 - Results

2.3.1 - Experimental Overview

In an attempt to prove our hypothesis, we have stably transfected (nucleofected) and transduced the commercial PFSK-1 cell line (originating from a supratentorial PNET in a 22-month-old) with the coding sequences of DNMT3B V1 and V6. This cell line is likely to mimic the cellular properties of ETMRs when transformed with DNMT3B V6, as PNETs are similar to ETMRs in that they occur in the brains of children, are highly undifferentiated, share similar histology and have a similar prognosis. pCMV6-Entry Vectors from OriGene© containing the coding sequences of DNMT3B V1 and V6 were used for nucleofection; vectors and coding sequences can be seen in Supplementary Data (Table S1 and Figure S1). Single clones were selected for each variant and the remaining clones were passaged in a batch culture until enough cells were available to collect mRNA and protein. Western blot analysis was used to determine DNMT3B expression levels, and quantitative real-time PCR (qPCR) was used to compare levels of DNMT3B mRNA with control PFSK-1 cells. In addition, these coding sequences were cloned into pLenti-C-Myc-DDK-IRES-Puro Vectors (see Figure S1) also purchased from OriGene©, and used to produce concentrated 3rd generation lentiviral particles capable of stably transducing the PFSK-1 cell line. Following transduction, cells were selected in puromycin until distinct cellular colonies formed. Cells were passaged in batch culture as the cells that survived selection proliferated at an unusually slow rate. A small portion of cells were set aside for fluorescence microscopy and protein was harvested once cells had proliferated to the desired confluence. Protein levels were analyzed though immunofluorescence and western blot analysis. It appears as though both methods of overexpressing DNMT3B did not result in viable cells with high levels of either DNMT3B variant, suggesting that there was an error in the methodology or that unnaturally high levels of DNMT3B are cytotoxic in the PFSK-1 cell line.

2.3.2 - Transfection of PFSK-1 Cells Using Nucleofection

In order to determine whether or not the overexpression of DNMT3B V6 is responsible for the unique methylation pattern and oncogenic properties of ETMRs, we performed nucleofection on PFSK-1 cells using vectors which contained the coding sequences of DNMT3B V1, V6 and an empty vector control. Cells nucleofected with the empty vector (pCMV6-Entry

Vector) died during colony selection, and nucleofection was repeated, but analysis of these cells was not pursued when overexpression of either protein was not detected (when compared to non-transfected PFSK-1) in cells that had been selected over the course of weeks in G418. Initial transfection efficiency was not measured due to the absence of GFP in the entry vector, and because antibiotic selection should result in cell death of all non-transfected cells. Colony selection proved to be difficult, as the majority of colonies picked did not survive transfer from the 100mm x 20mm selection dish to the 24-well plate. Cells were propagated and passaged at least twice following nucleofection, prior to mRNA and protein extraction. mRNA was quantitated using a NanoDrop 1000 (ThermoScientific), all mRNA purified demonstrated an A_{260}/A_{280} ratio in the range of 2.08-2.14. Protein levels were quantified using a standard Bradford protein assay. DNMT3B mRNA levels were determined using real-time PCR (qPCR) and DNMT3B protein expression was measured using western blot analysis. The results of the qPCR and western blot analysis can be seen in Figure 2 and Figure 3, respectively.

PFSK-1 cells were transfected (using nucleofection) with the following constructs: pCMV6-Entry Vector-Empty Vector, pCMV6-Entry Vector-DNMT3B V1 and pCMV6-Entry Vector-DNMT3B V6. Following transfection, cells were cultured in an appropriate concentration of G418 (previously determined through the use of a kill curve) until colonies of cells appeared in the tissue culture plates the cells were grown in. pCMV6-Entry Vector-Empty Vector cells did not produce colonies following nucleofection and selection. Twelve clones (colonies of cells) for both DNMT3B V1 and DNMT3B V6 were then transferred to 24-well plates using cloning discs dipped in trypsin (this was repeated as no clones were viable following the first transfer attempt). The remaining clones were passaged in a batch culture (DNMT3B V1 Batch and DNMT3B V6 Batch) to a fresh plate. All cells were maintained in G418 following clonal selection. The majority of clones did not survive transfer to the 24-well plate, only one DNMT3B V1 clone (DNMT3B V1 Clone 20) and three DNMT3B V6 clones (DNMT3B V6 Clone 22, DNMT3B V6 Clone 23 and DNMT3B V6 Clone 25) proliferated following clonal transfer.

The four viable clones and two sets of batch cultures were expanded in culture until total mRNA and protein could be extracted. DNMT3B mRNA levels were quantified in triplicate (n = 3) using qPCR (with no reverese transcription, dH_2O reverse transcription and no template controls). The average levels of expression, normalized to total mRNA isolated from non-nucleofected PFSK-1 can be seen in Figure 2. The no reverese transcription, dH_2O reverse

transcription and no template controls all showed little to no signal following quantitative PCR, demonstrating absence of mRNA contamination. The qPCR analysis did not demonstrate any clear differences in DNMT3B mRNA expression levels between non-nucleofected PFSK-1 and the six sets of stably nucleofected PFSK-1 (DNMT3B V1 Batch, DNMT3B V1 Clone 20, DNMT3B V6 Batch, DNMT3B V6 Clone 22, DNMT3B V6 Clone 23 and DNMT3B V6 Clone 25). As seen in Figure 2, standard deviation of DNMT3B mRNA expression for all six samples encompassed the point of no change in expression (1.000E+0) or fell below it.

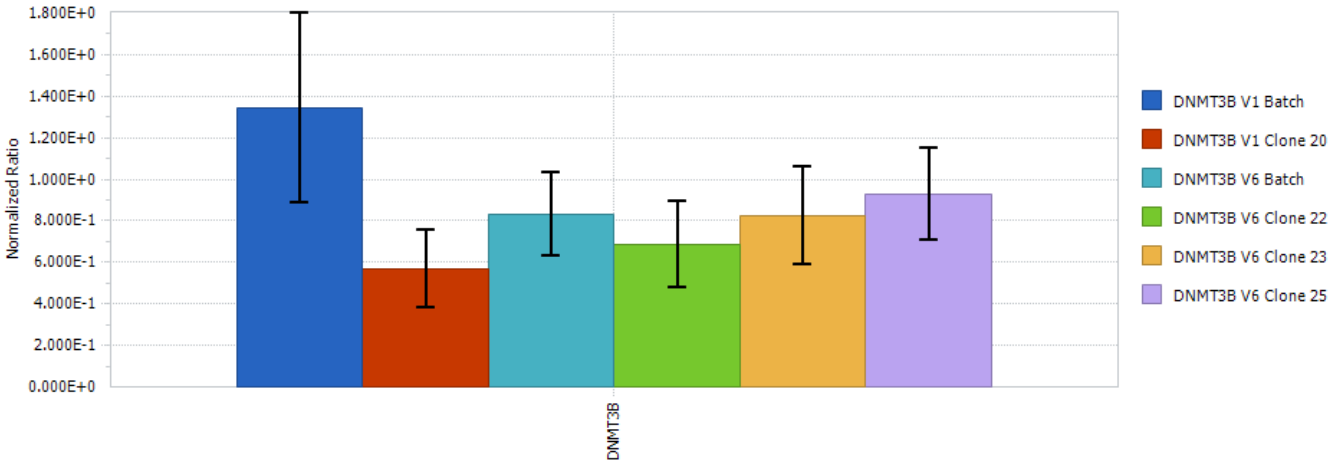

Figure 2 - Normalized Expression of DNMT3B by qPCR in Nucleofected PFSK-1

mRNA expression levels of viable clones and batch cultures of stably nucleofected PFSK-1 (PNET) were analyzed using qPCR analysis. DNMT3B V1 Clone 20, DNMT3B V6 Clone 22, DNMT3B V6 Clone 23 and DNMT3B V6 Clone 25 are cell lines propagated from a single colony of cells following nucleofection (using either the DNMT3B V1 or DNMT3B V6 vector constructs) and antibiotic selection. DNMT3B V1 Batch and DNMT3B V6 Batch are the remaining colonies of cells that were not isolated, and passaged together in cell culture following selection. Expression of mRNA was normalized to non-nucleofected PFSK-1. Although the nucleofected clones and batch culture survived in antibiotic resistance at a concentration previously determined to be lethal (within 2 days), there is no overexpression of either DNMT3B variant in the cells examined.

DNMT3B protein expression levels of the four viable nucleofected PFSK-1 clones (DNMT3B V1 Clone 20, DNMT3B V6 Clone 22, DNMT3B V6 Clone 23 and DNMT3B V6

Clone 25) and the two remaining nucleofected PFSK-1 batch cultures (DNMT3B V1 Batch and DNMT3B V6 Batch) were measured in tandem using western blot analysis. In addition, two total protein controls were loaded onto the same western blot gel. Protein isolated from non-nucleofected PFSK-1 (PFSK-1 WT), which should harbor basal levels of DNMT3B expression, and the ETMR cell line BT183, which is known to have significant overexpression of DNMT3B. β-tubulin protein levels were also measured as a loading control. The results of this western blot can be seen in Figure 3. There was no clear detection of DNMT3B overexpression in any of the protein samples, while the β-tubulin signal was strong in all samples analyzed. Futher western blot analysis was not pursued due to the lack of DNMT3B mRNA expression demonstrated in qPCR analysis (see Figure 2).

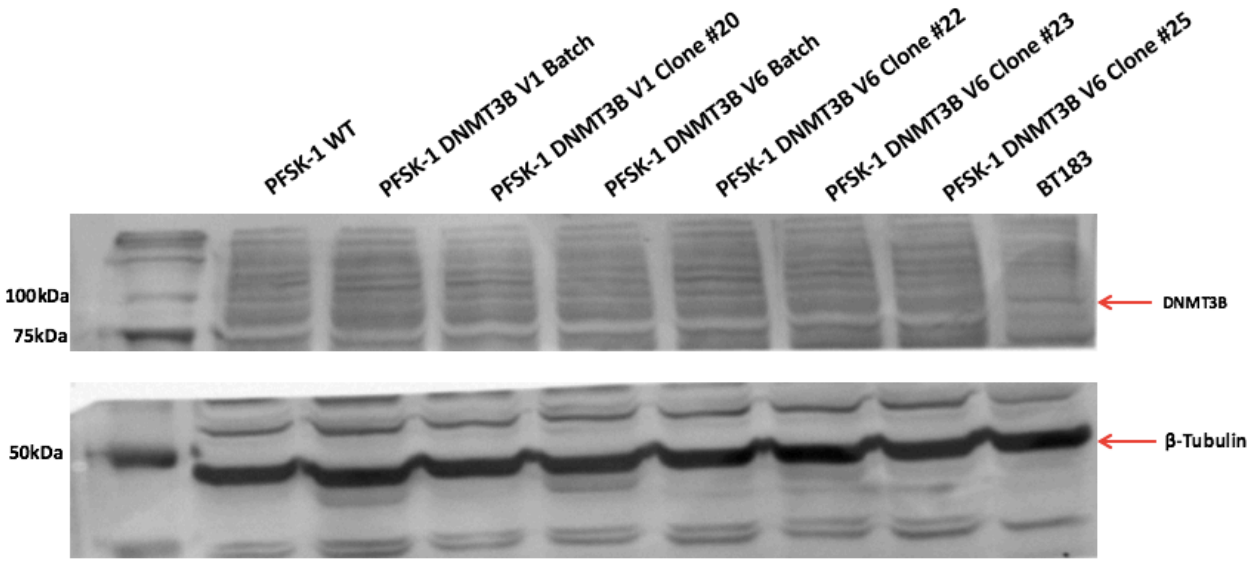

Figure 3 – Nucleofected PFSK-1 DNMT3B Western Blot

Protein expression levels of viable clones and batch cultures of stably nucleofected PFSK-1 (PNET) were examined through western blot analysis. DNMT3B V1 Clone 20, DNMT3B V6 Clone 22, DNMT3B V6 Clone 23 and DNMT3B V6 Clone 25 are cell lines propagated from a single colony of cells following nucleofection (using either the DNMT3B V1 or DNMT3B V6 vector constructs) and antibiotic selection. DNMT3B V1 Batch and DNMT3B V6 Batch are the remaining colonies of cells that were not isolated, and passaged together following selection. PFSK-1 WT refers to non-nucleofected PFSK-1 and BT183 is a control ETMR cell line known to have a significant overexpression of DNMT3B. β-tubulin expression was measured in each sample to provide a loading control. No clear overexpression of either protein was detected.

2.3.3 - Lentiviral Transdution of PFSK-1 Cells

The DNMT3B V1 and V6 coding sequences were then cloned into pLenti-C-Myc-DDK-IRES-Puro Vectors, and 3rd generation lentiviral particles were produced and concentrated using PEG-it Virus Precipitation Solution (SBI). A third batch of Lentiviral particles were produced using the empty vector. The three types of concentrated lentivirus (produced from pLenti-C-Myc-DDK-Empty Vector (EV), pLenti-C-Myc-DDK-DNMT3B V1 and pLenti-C-Myc-DDK-DNMT3B V6 lentiviral destination vectors) were used to transduce PFSK-1 cells, and transduced cells were selected over the course of weeks in puromycin. It took several weeks for clear cellular colonies to form in tissue culture during antibiotic selectoin, the size of these viable

clones were noticeably smaller than the ones produced from nucleofection. Therefore, the transduced clones were passaged together to speed up the collection of protein from successfully transformed cells, resulting in three sets of stably transduced PFSK-1 batch cultures: EV (pLenti), DNMT3B V1 Batch and DNMT3B V6 Batch. Following at least two passages, cells were set aside for immunofluorescence analysis and the remainder of cells were used to extract protein. Protein levels were quantified using a standard Bradford protein assay.

Given that high levels of DNMT3B were not detected in the BT183 cell line through western blot analysis, and β-tubulin was found to be at normal levels in the protein isolated (in nucleofected PFSK-1), it is highly probable that the primary DNMT3B antibody provided was not properly binding to DNMT3B V1 or DNMT3B V6 during incubation. In order to see if any expression of our DNMT3B variants was present in our transduced PFSK-1 cell lines, we decided to perform alternative western blot analysis targeting the C-terminal Myc-tag (using two separate primary antibodies) believed to be attached to our proteins of interest (see Figure S1). Western blot analysis of these three transduced PFSK-1 cultures included total protein isolated from non-transduced PFSK-1 (PFSK-1 WT), non-transduced PFSK-1 treated with polybrene (PFSK-1 Polybrene), and protein isolated from a pediatric glioblastomas cell line transformed to express the H3-G34R histone protein tagged with Myc (SF188 H3-G34R-Myc), as a positive control. β-tubulin was also measured as a loading control. The histone protein sample (SF188 H3-G34R-Myc) should not be expected to have the presence of β-tubulin, as all other proteins are degraded during the process of histone isolation. The results of this western blot analysis can be seen in Figure 4a-b.

The first primary antibody used (Cell Signaling, Rabbit mAb, 1:1000) did not demonstrate the presence of Myc-tagged DNMT3B in any of the samples analyzed, nor did it detect the presence of the Myc tag in our SF188 H3-G34R-Myc positive control (see Figure 4a). β-tubulin signal was strong in all protein samples analyzed (see Figure 4a). The second primary antibody used (Cell Signaling, Mouse mAb, 1:1000) demonstrated that the first antibody was ineffective, as the SF188 H3-G34R-Myc positive control showed a strong signal (see Figure 4b). β-tubulin signal was again strong in all protein samples analyzed. The presence of Myc-tagged DNMT3B in the two transduced PFSK-1 samples analyzed (DNMT3B V1 Batch and DNMT3B V6 Batch) was questionable, as the band present where Myc-tagged DNMT3B should have

shown up for PFSK-1 DNMT3B V1 Batch and DNMT3B V6 Batch (~97kDa) was also present at a similar intensity for PFSK-1 WT and PFSK-1 Polybrene (see Figure 4b).

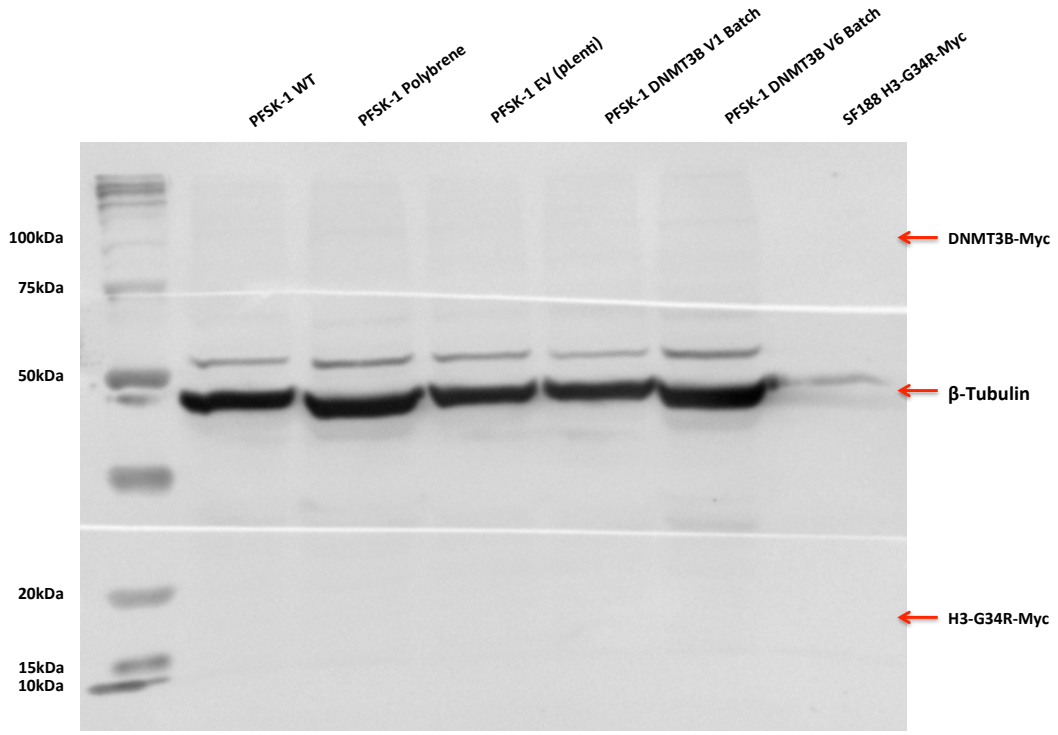

Figure 4(a) - Transduced PFSK-1 Myc Tag Western (Cell Signaling, Rabbit mAb, 1:1000)

Protein expression levels of DNMT3B tagged with Myc at the C-terminal end of the protein were analyzed through western blot analysis. PFSK-1 WT refers to non-transduced PFSK-1, protein was also extracted and loaded from PFSK-1 cells treated with polybrene only (PFSK-1 Polybrene). SF188 H3-G34R-Myc is a previously transformed cell line known to express a G34R mutant histone H3 with a Myc tag, and served as a positive control. β-tubulin expression was measured in each sample to provide a loading control. Given the lack of detection of Myc tag in the positive control, it is likely the primary antibody did not properly bind to the Myc tag during incubation. Lack of β-tubulin in the positive control is due to the fact that histone extraction results in degradation of all other proteins.

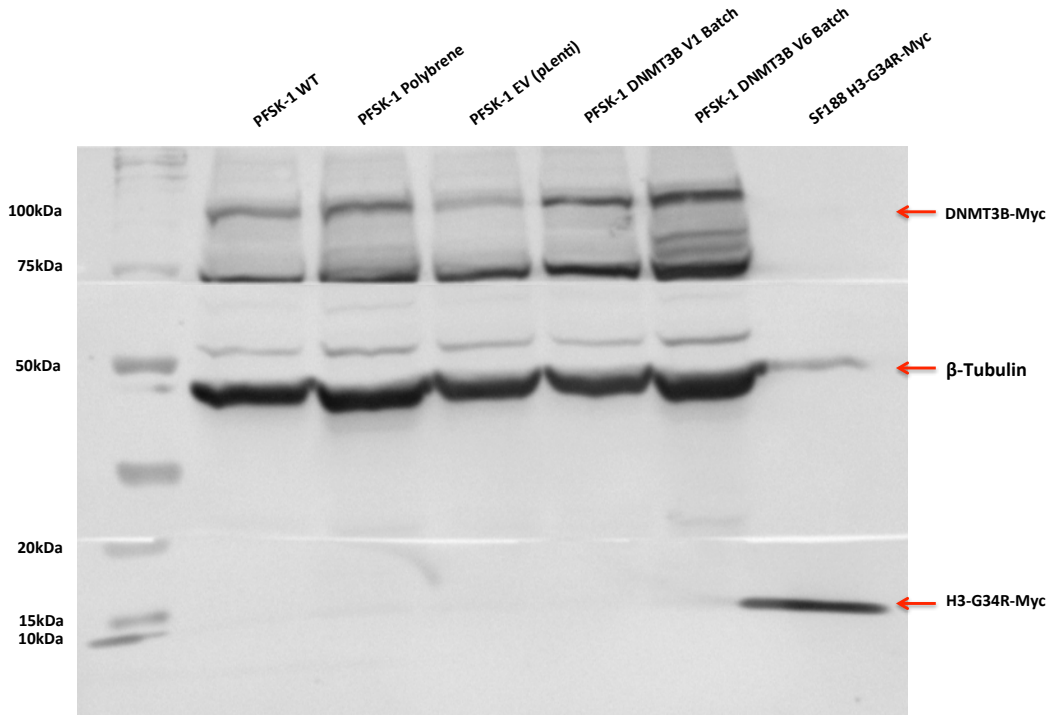

Figure 4(b) - Transduced PFSK-1 Myc Tag Western (Cell Signaling, Mouse mAb, 1:1000) *Protein expression levels of DNMT3B tagged with Myc at the C-terminal end of the protein were analyzed through western blot analysis. PFSK-1 WT refers to non-transduced PFSK-1, protein was also extracted and loaded from PFSK-1 cells treated with polybrene only (PFSK-1 Polybrene). SF188 H3-G34R-Myc is a previously transformed cell line known to express a G34R mutant histone H3 with a Myc tag, and served as a positive control. β-tubulin expression was measured in each sample to provide a loading control. Lack of β-tubulin in the positive control is due to the fact that histone extraction results in degradation of all other proteins. Slight overexpression appears to be present in the DNMT3B V1 and DNMT3B V6 transduced cells when compared to wild type, polybrene treated and empty vector transduced PFSK-1.*

In order to determine whether or not the presence of Myc-tagged DNMT3B in the two stably transduced PFSK-1 samples (DNMT3B V1 Batch and DNMT3B V6 Batch) seen through western blot analysis (see Figure 4b) was just background noise, immunofluorescence (IF) targeting the Myc-tag was performed on the three sets of stably transduced PFSK-1 batch cultures (EV, DNMT3B V1 and DNMT3B V6). In addition, immunofluorescence targeting the

Myc-tag was also performed on both negative (non-transduced PFSK-1 or PFSK-1 WT) and positive (SF188 H3-G34R-Myc) controls. Cells were plated in duplicate in 4-well chamber slides prior to immunofluorescence analysis. The second well served as a control for the presence of background fluorescence had the secondary antibody not been properly washed away, and lacked the addition of the primary antibody during immunofluorescence analysis (No 1° Ab). The nuclei of cells in this IF experiment were stained using the DAPI fluorescent stain. Visualization of the Myc-tag was performed using an Alexa Fluor 568 conjugated secondary antibody.

The IF protocol proved to be successfully carried out, as SF188 H3-G34R-Myc demonstrated localized Myc-tag expression in the nuclei of all cells plated and PFSK-1 WT had no presence of Myc-tag in any of the cells plated (see Figure 5a-b). The No 1° Ab control wells did not show fluorescence in any of the samples analyzed, demonstrating that any fluorescent signals seen were not due to background noise (see Figure 5a-e). All stably transduced PFSK-1 batch cultures (EV, DNMT3B V1 and DNMT3B V6) had a low percentage of cells expressing the Myc-tag (see Figure 5c-e). This further confirms the insignificant levels of overexpression of DNMT3B V1 and DNMT3B V6 in the stably transduced PFSK-1 batch cultures seen in western blot analysis (Figure 4b). The presence of Myc-tagged DNMT3B in some cells but not others following transduction and antibitioc selection (Figure 5c-e) raises several questions about how cells survived antibotioc selection but either did not integrate or express the appropriate lentiviral sequences.

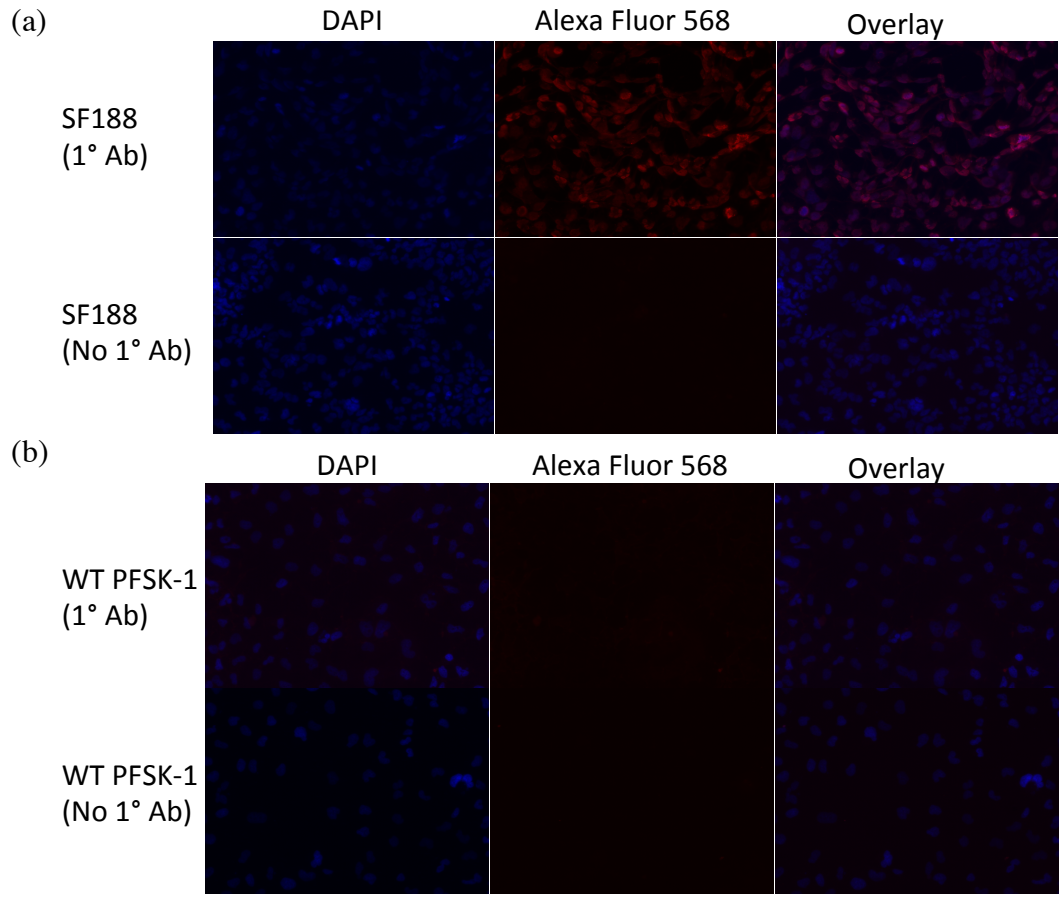

Figure 5 (a and b) – Transduced PFSK-1 Myc-Tag IF (Positive and Negative Controls)
Immunofluorescence analysis of the SF188 H3-G34R-Myc (positive control) and non-transduced PFSK-1 (negative control) cell lines using the Alexa Fluor 568 fluorescent dye conjugated secondary antibody. The DAPI fluorescent stain was used to determining the location of nuclei in these cells. The non-transduced PFSK-1 negative control (b) shows a complete lack of Myc tag expression in the nucleus, while the SF188 H3-G34R-Myc positive control (a) shows a dense concentration of Myc tag expression in the nucleus (where histone H3 is localized). This confirms the immunofluorescence protocol was functioning properly.

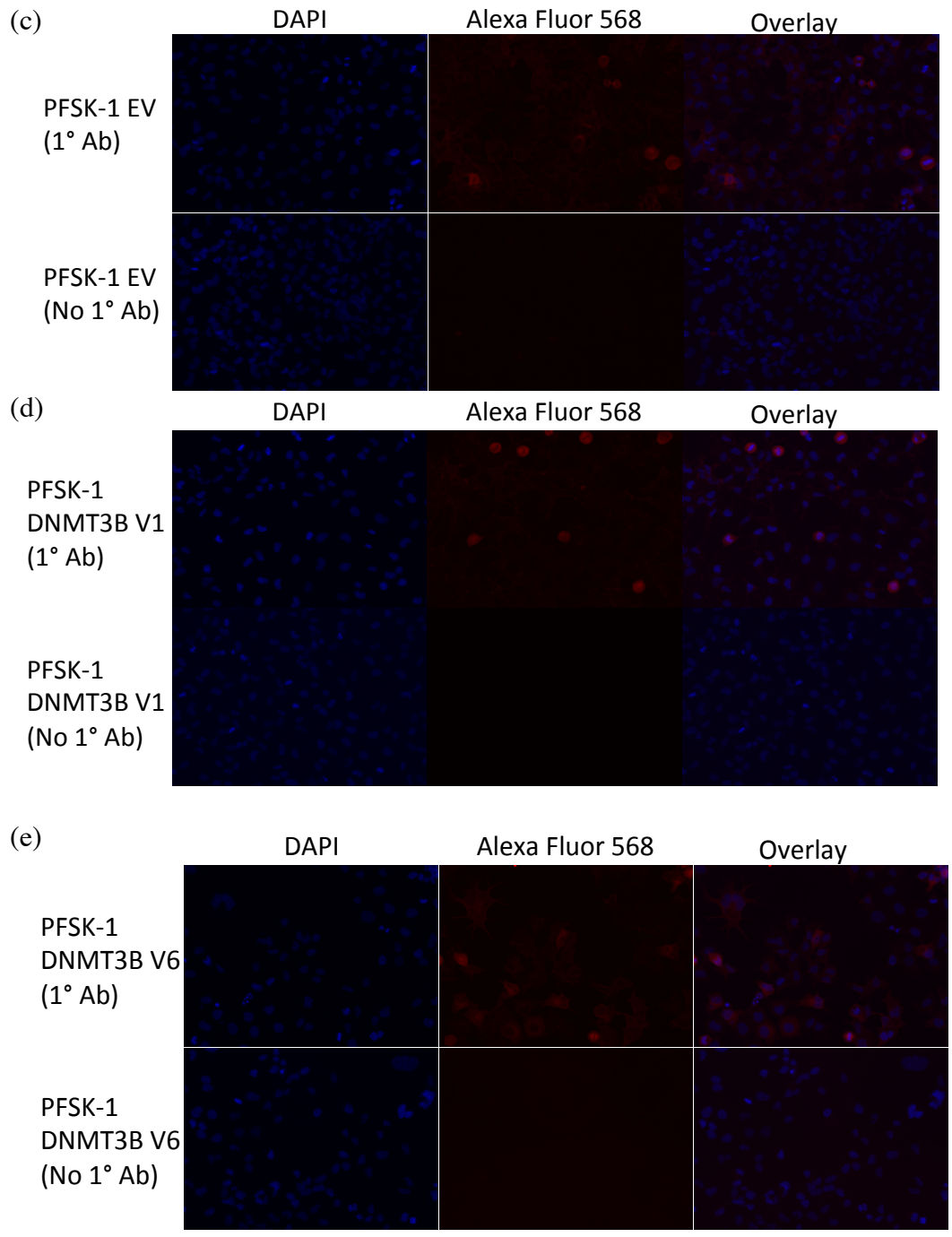

Figure 5 (c, d and e) - Transduced PFSK-1 Myc-Tag IF

Immunofluorescence analysis of the PFSK-1 cell lines transduced with lentiviral particles containing Myc-tagged DNMT3B V1 and DNMT3B V6 coding sequences, and lentiviral particles produced with the empty vector (EV) construct, using the Alexa Fluor 568 fluorescent dye conjugated secondary antibody. The DAPI fluorescent stain was used to determining the location

of nuclei in these cells. Expression of the Myc tag was present in a subset of cells in all three transduced cell types (c), (d) and (e). However, the expression levels of this Myc tag are abnormally low for lentiviral transduction, and it is abnormal for all cells to not be expressing the Myc tag given that cells were selected for and maintained in puromycin following transduction.

Overexpression of DNMT3B V1 and DNMT3B V6 was attempted in the PFSK-1 (PNET) cell line using both stable nucleofection and stable lentiviral transduction. mRNA and protein expression levels of stably nucleofected PFSK-1 cell lines (DNMT3B V1 Clone 20, DNMT3B V1 Batch, DNMT3B V6 Clone 22, DNMT3B V6 Clone 23, DNMT3B V6 Clone 25 and DNMT3B V6 Batch) were analyzed using qPCR and western blot analysis, respectively. Neither the qPCR or western blot analysis demonstrated an increased expression of either DNMT3B variant in nucleofected PFSK-1 cells that underwent antibiotic selection (see Figures 1 and 2). It is plausible that the entry vector used was unable to transcribe our proteins of interest in this particular cell line once stably transfected.

Lentiviral transduction yielded similarly questionable results. PFSK-1 was transduced using lentiviral particles produced with three lentiviral destination vectors containing a C-terminal Myc-tag at the multiple cloning site (an empty vector, a vector containing the coding sequence of DNMT3B V1 and a vector containing the coding sequence of DNMT3B V6). Following transduction and antibiotic selection, total protein was isolated from these cell lines and expression levels of Myc-tagged DNMT3B were analyzed using western blot analysis. Overexpression of Myc-tagged DNMT3B V1 and V6 was not clear from this experiment, as the signal detected in PFSK-1 DNMT3B V1 Batch and PFSK-1 DNMT3B V6 Batch was similar to those seen in negative controls (see Figure 4b). The transduced PFSK-1 cell lines also underwent immunofluorescence analaysis and showed a low percentage of cells expressing the Myc-tag in all three successfully transduced PFSK-1 cell lines (see Figure 5c-e). The ability of these cells to survive antibiotic selection but not express our construct (Myc-tagged DNMT3B V1 or V6) is highly puzzling. It is plausible that transcription of lentiviral sequences were silenced in viable cells following integration due to the cytotoxic effects of high levels of DNMT3B. Further inquiry into the effects of overexpression of either variant of DNMT3B *in vitro* is required.

2.4 - Issues Encountered Throughout Experimental Process

Numerous problems arose throughout our attempts to overexpress DNMT3B V1 and V6 in PFSK-1. Standard transfection of the pCMV6-Entry Vectors containing the coding sequences of these DNMT3B variants was first performed using Lipofectamine 2000 (ThermoFisher Scientific), however cells did not survive during the antibiotic selection that followed. Transfection of PFSK-1 with a GFP control plasmid demonstrated a high efficiency of plasmid integration, and it is unknown why transfection with the plasmids containing our proteins of interest resulted in cell death. Prior studies have been successful in transfecting PFSK-1 with plasmid DNA, so the nature of this experimental failure could have been dependent on the plasmids themselves[2]. It is also possible that PFSK-1 is highly sensitive to the cytotoxicity of the lipofectamine transfection reagent. As such, we switched to nucleofection (a method shown to be highly effective on cells with an undifferentiated phenotype in the past) in the hope that our plasmids would stably integrate into PFSK-1. Following antibiotic selection, these cells were extremely sensitive to any type of passaging. Almost all the clones picked were unable to survive transfer from the $10cm^2$ selection dish to a 24-well plate. In fact, the first attempt at transferring clones following nucleofection and selection resulted in the death of all clones selected, causing the experiment to have to be repeated. The ability of these cells to survive antibiotic selection following nucleofection, yet not express the proteins encoded in the plasmids now present in them is a highly confusing result, and raises questions on the integrity of the plasmids themselves.

While attempting to clone the coding sequences of these DNMT3B variants from the pCMV6-Entry Vector into the pLenti-C-Myc-IRES-DDK-Puro Vector, the polymerase chain reaction (PCR) to isolate and amplify these sequences repeatedly failed. In order to overcome this, several primers had to be designed and many PCRs with an annealing temperature gradient were performed to identify ideal conditions. Once we overcame this hurdle, our attempts to ligate this sequence into our lentiviral destination vector were also repeatedly unsuccessful. During the troubleshooting process, we ran many ligation reactions with different incubation temperatures, incubation times, amounts of vector DNA and amounts of insert DNA until the restriction digest yielded positive results. The pLenti-C-Myc-IRES-DDK-Puro Vector was the second lentiviral destination vector that we cloned the DNMT3B sequences into. The first vector we cloned into was pLenti-C-Myc-DDK-IRES-GFP. The lentivirus produced using this vector failed to express

GFP in PFSK-1 following transduction. As a result, we transfected this vector using Lipofectamine 2000 (ThermoFisher Scientific) into 293FT cells (as well as a GFP control plasmid). The cells transfected with pLenti-C-Myc-DDK-IRES-GFP failed to express GFP, while the cells transfected with the control plasmid expressed GFP at very high levels. We then sequenced the vector and found inconsistencies in the sequence provided on the OriGene© website and our sequencing results (specifically in the IRES sequence and vector backbone). This caused us to have to repeat the cloning process, lentiviral production and transduction with a new vector (pLenti-C-Myc-IRES-DDK-Puro). PFSK-1 cells were also highly resistant to attaching to the chamber slides during immunofluorescence. The cells consistently washed away following fixation, until a pre-fixation using 4% PFA was found to fix the problem; this caused the immunofluorescence to have to be repeated several times. Similar to the pCMV6-Entry Vector used in nucleofection, it is difficult to explain the lack of protein expression in PFSK-1 cells transduced and selected for in puromycin for weeks. Even more puzzling is the fact that some transduced cells demonstrated expression of Myc-tagged DNMT3B while others did not. It is plausible that the functionality of these plasmids were not ideal, as the coding sequences of DNMT3B V1 and V6 were confirmed to be present in them at the time of lentiviral production (see Table S1).

2.5 - Discussion

Our attempts to overexpress DNMT3B V1 and V6 in the PFSK-1 cell line proved to be much more difficult than we anticipated. Although the cells survived antibiotic selection following nucleofection and lentiviral transduction, it appears that no significant change in DNMT3B expression was present in either type of transformed cell. Nucleofection yielded particularly negative results of DNMT3B variant overexpression. The western blot did not demonstrate any increased presence of DNMT3B in nucleofected PFSK-1 compared to non-nucleofected cells. However, the lack of DNMT3B overexpression in the BT183 ETMR cell line (which is characterized by DNMT3B V1 and V6 overexpression) predicts that the DNMT3B primary antibody used may not have been ideal. Protein extraction was not an issue as the β-tubulin loading control showed a strong signal for all samples examined. qPCR clearly demonstrated that there was no increased expression of DNMT3B V1 or V6 mRNA in nucleofected cells. When expression levels of DNMT3B mRNA were normalized to non-

nucleofected PFSK-1, the only sample which showed slight increase was DNMT3B V1 processed in batch culture, however this increase was negligible and the standard deviation fell below the expression level of DNMT3B mRNA in non-transfected PFSK-1. Intitial transfection efficiency was not measured as antibiotic selection should have killed off any non-stably transfected cells, and no fluorescent protein was encoded within the vector used.

Examination of the expression of Myc-tagged DNMT3B V1 and V6 in transduced PFSK-1 cells yielded questionable results. Although there may have been a presence of Myc-tagged DNMT3B in the western blot, the bands present in non-transduced PFSK-1 and polybrene treated PFSK-1 samples indicate that the slightly stronger signal in the DNMT3B V1 and V6 transduced samples may just be due to background noise or human error when loading. The immunofluorescence on the other hand did show that there was Myc-tagged DNMT3B present in some cells that survived antibiotic selection. However, the low percentage of cells expressing the DNMT3B variants and lack of protein overexpression in the western blot (confirmed to be accurate through the β-tubulin loading control and SF188-H3-G34R-Myc positive control) indicate that expression levels were not high enough to continue with cellular analysis. Similar to nucleofection, efficiency of lentiviral transduction was not calculated because of the antibiotic selection process and lack of GFP in the lentiviral destination vector.

Had demonstrated overexpression of DNMT3B V1 and V6 occurred in the PFSK-1 cell line (through either stable nucleofection or lentiviral transduction), the next step would have been to extract DNA from the proliferated clones and batch cultures, and analyze the methylation profile in the cells exhibiting overexpression of these proteins (as well as the control cell lines). This would be carried out using the 450K methylation array. Analysis of the methylation profile (and seeing whether or not this methylation pattern clustered with the distinct ETMR methylation phenotype observed in prior studies) would give great insight into whether or not overexpression of DNMT3B V1 or V6 is responsible for the unique methylation profile present in ETMR samples[2]. It also would have been ideal to perform analysis of cellular proliferation (using the MTT assay), cellular migration and invasion (using Boyden chamber assays) in both the transformed and control cells to determine if the oncogenic properties in PFSK-1 changed following transformation and transduction. In addition, the appearance of cell morphology in cells before and after transformation would need to be examined, given the distinct difference in appearance between PNETs and ETMRs in tissue culture.

Given that PFSK-1 is far more stable in cell culture than human neural stem cells (hNSC), and less sensitive to intrusive attempts of protein overexpression, it was the initial target for DNMT3B V1 and V6 overexpression. Had these experiments succeeded, we would have gone on to stably overexpress these proteins in hNSC, of different ages, in stem cell culture and analyze them in a similar way as described above. This is due to the fact that oncogenesis of ETMRs occurs during early infancy, when expression of DNMT3B V6 is unnaturally present in neuronal stem cells. hNSCs provide the cellular context that is most similar to the ETMR cancer initiating cells in nature, and overexpression of DNMT3B V6 in these cells is the ideal functional study for this subset of embryonal brain cancer[62]. hNSC of the earliest age would provide the best cellular environment for oncogenesis of ETMRs would be the ideal cell type to target for overexpression of DNMT3B V6 in future studies.

2.6 - Experimental Contributions

Andrea Bajic maintained the ETMR BT-183 cells in stem cell culture, and provided the protein extracts (extracted and quantified) from this cell line used in western blot analysis. Damien Faury transformed the pediatric glioma SF188 cell line to express G34R mutant histone H3 with the Myc tag (resulting in the SF188 H3-G34R-Myc cell line). Histone proteins were also extracted and quantified by Damien Faury.

Chapter 3 - General Discussion

3.1 - Summary of Experiments Performed

Recent evidence has shown that the amplification of the gene fustion between *TTYH1* and C19MC causes a targeted downregulation of *RBL2* expression, resulting in an increased expression of exon 1B-containing DNMT3B in ETMRs[2]. The lack of methylation at the promoter driving exon 1B-containing DNMT3B expression in ETMRs (promoter 1B), the unique methylation profile of ETMRs and their highly undifferentiated cellular nature all point to exon 1B-containing DNMT3B being a key oncogenic driver of this rare cancer. As ETMRs were once classified under the umbrella term of central nervous system primitive neuroectodermal tumors (CNS PNETs), due to their undifferentiated phenotypes, histological similarity, similar anatomical location and comparably poor prognosis, we believed that overexpression of DNMT3B V6 in a PNET cell line would recapitulate the early embryonic cellular state seen in ETMRs (specifically the DNA methylation profile and oncogenic characteristics). We set out to overexpress DNMT3B V6 in the PNET cell line PFSK-1 (as well as the DNMT3B V1 control) in order to demonstrate its unique transformative properties.

Nucleofection and lentiviral transduction were performed on the PFSK-1 cell line, and antibiotic selection following transfection and transduction occurred for weeks following transformation (using G418, and puromycin, respectively). Following this, expression of DNMT3B mRNA and protein was examined in PFSK-1 nucleofected with vectors containing the coding sequences of DNMT3B V1, DNMT3B V6 and an empty vector control. Clones were isolated and propagated in cell culture (one clone for DNMT3B V1 and three clones for DNMT3B V6) with the remaining colonies expanded as a batch culture. Increased expression of either variant was not detected, following antibiotic selection, by both qPCR and western blot analysis.

Lentiviral transduction of PFSK-1 with lentiviral particles thought to be capable of transducing cells to express Myc-tagged DNMT3B V1 and DNMT3V6 protein occurred in tandem (as well as lentivirus produced with an empty vector control). Following weeks of proliferation and passaging in a concentration of puromycin previously determined to be lethal to the PFSK-1 cell line, expression of Myc-tagged DNMT3B was measured in transduced PFSK-1. Western blot analysis did not reveal any significant overexpression of either variant, and

immunofluorescence demonstrated that a low percentage of cells expressed Myc-tagged DNMT3B following selection. As neither method of overexpression was successful, analysis of cellular characteristics and the DNA methylation profile of these PFSK-1 cells did not occur.

3.2 - Hypothesized Effect of Exon 1B-Containing DNMT3B

Given the recent evidence that exon-1B containing DNMT3B expression peaks at around 8 weeks post conception, and is exclusively expressed in fetal brain tissue, it is likely responsible for the establishment of a methylation pattern that facilitates proliferation and maintaining stem-cell like properties during early neurogenesis[2]. As it is found to be exclusively overexpressed in ETMRs (which have stem-cell like properties) and silenced in pediatric and adult brain through methylation of promoter 1B, its aberrant expression likely causes an unnatural reestablishment of epigenomic organization that is only present during fetal development[2]. Due to the fact that the N-terminal region of DNMT3B is responsible for DNA targeting and binding to chromatin, the extra 11 amino acids present at the start of the N-terminus in DNMT3B V6 (see Table S2) likely dictate where the catalytic activity of this novel *de novo* methyltransferase takes place. DNMT3B has proved to be oncogenic in numerous types of cancer, so it is highly likely that overexpression of this unique variant is the cause of ETMR formation[171]. However, overexpression of DNMT3B (V1 and V6) in PFSK-1, a type of PNET that already contains a genome unstable enough to form a malignant cancer, is plausibly too much for the cells to handle (and could induce cell death). Indeed, although there are many studies in which DNMT3B is downregulated through siRNA treatment or inhibited pharmacologically; a thorough examination of current literature shows no studies in which catalytically active DNMT3B is overexpressed in cells *in vitro*[216,217]. Perhaps the overexpression of DNMT3B in ETMRs and other cancers requires a highly specific cellular context and fine tuned level of expression in order for the cytotoxic effects of overexpression to not take priority of over the oncogenic ones. Further analysis of the effects of DNMT3B overexpression would need to take place in order to confirm this theory.

3.3 - Future Directions

As the oncogenic effects and DNA methylation activity of DNMT3B V6 are still unknown, future studies examining the ability of DNMT3B V6 to transform cells and change their methylation profile should be an ongoing area of research. Although overexpression of this

protein failed in PFSK-1, stable lentiviral integration of the coding sequence of DNMT3B V6 into cells more likely to model ETMR formation, such as hNSC would be a better place to start. Several experiments could take place to demonstrate the theory that the amplification of fusion between the promoter of *TTYH1* and C19MC is responsible for upregulation of DNMT3B V6, and therefore ETMR oncogenesis.

As stated prior, lentiviral transduction of hNSC to stably overexpress the coding sequences of DNMT3B V1 and DNMT3B V6 should be the next experimental approach. In addition, other attempts to recapitulate the effects of the *TTYH1*-C19MC gene fusion amplification in hNSC could take place in tandem, such as RBL2 knockdown mediated by siRNA, and co-transfection of the two siRNAs from C19MC demonstrated to downregulate RBL2 (miR-519a and miR-512-3p). Following these three functional experiments in hNSC, expression levels of DNMT3B V1 and V6 should be measured through western blot and immunofluorescence analysis. RBL2 expression should also be examined in cells that underwent RBL2 knockdown through siRNA and miR co-transfection. Once stable changes in protein levels are demonstrated, genomic DNA from the hNSC should undergo 450K methylation array analysis, and the results should be clustered with the current ETMR methylation profile data. This would give insight into whether or not DNMT3B V6 is the enzyme responsible for the unique methylation profile seen in this rare cancer. Oncogenic properties of these transformed cells (rates of proliferation and metastatic capability) could also be examined through MTT and Boyden chamber assays.

In addition to experiments taking place in hNSC, functional studies on ETMR cell lines such as the BT183 cell line should occur as well. Overexpression of RBL2 through lentiviral transduction could be one approach, while inhibition of DNMT3B V1 and V6 through siRNA transfection could be another. Similarly, changes in protein levels of RBL2 and DNMT3B should be confirmed through western blot and immunofluorescence analysis. Subsequent changes in DNA methylation and oncogenic characteristics (cellular proliferation and metastatic ability) should be then be analyzed through 450K methylation arrays, MTT assays, and Boyden chamber assays. Should either of these *in vitro* studies demonstrate significant results, it would be ideal to attempt develop a mouse model of ETMRs through intracranial injection of lentiviral particles capable of transducing cells to overexpress DNMT3B V6. Should brain tumors develop in these mice, histology should be examined to see if they contain the multilayered rosette

histology unique to ETMRs and the DNA methylation profile of these tumors should be analyzed through 450K methylation. In summary, there are several approaches to demonstrate the oncogenic effects of DNMT3B V6, and its role in ETMR formation, that should occur in the future. Establishing that this protein is largely responsible for the formation of these deadly brain tumors would be a great step forward in developing new treatment approaches (such as DNMT3B inhibitors) for affected children.

3.4 - General Conclusion

Extensive evidence has shown that exon 1B-containing DNMT3B (DNMT3B V6) is responsible for the unique methylation profile seen in ETMR samples[2]. It is aberrantly overexpressed in ETMRs due to hypomethylation of DNMT3B promoter 1B, and is only otherwise found expressed in fetal brain during weeks 8-10 of gestation[2]. Determining whether or not this DNMT3B variant is responsible for the formation of ETMRs (as an indirect result of the amplification of the gene fusion between the promoter of *TTYH1* and C19MC), and examining its oncogenic properties should be an active area of embryonal brain cancer research. Demonstrating that this is the key oncoprotein causing ETMR formation would give physicians and scientists the opportunity to develop new treatment methods for this rare and deadly form of pediatric brain cancer. Should overexpression of DNMT3B V6 in hNSC result in oncogenic transformation, and DNA methylation profiles that cluster with current ETMR methylation profile data, the development of novel DNMT3B inhibitors would greatly influence the ability to treat children affected by this deadly disease.

Although there were no substantial results produced throughout this experimental process, one can speculate on the effect of overexpressing variants of DNMT3B in PNET cell lines, such as PFSK-1, by extrapolating on the selection that occurred following transduction. It is plausible that the PFSK-1 cells that integrated either DNMT3B variant (exon 1B-containing and variant 1) at a high copy number were unable to maintain expression of DNMT3B at a significant level following genomic integration, or were consequently lost in passaging due to cell death. Transcription of these genes could have also been silenced through methylation in promoter regions of the virally integrated sequence, preventing expression of the enzymes and downstream tags. Overexpression of *de novo* methyltransferases can severely affect the genomic integrity of cells, causing widespread chromosomal instability, and it is feasible that

overexpressing DNMT3B V1 and V6 in PFSK-1 could have resulted in cell death. It is also possible that cells expressing these DNMT3B variants at high enough levels to alter the methylation profile of the genome were lost throughout the G418 and puromycin selection processes due to the cytotoxic effects of high levels of DNMT enzymes (triggering apoptosis) when in the presence of other drugs[218]. This would explain the low percentage of cells expressing the Myc-protein tag in the immunofluorescence assay and lack of Myc tag detection in western blots performed. However, due to the inconsistency in vector backbone sequencing results and frequently failed transfection experiments with vectors provided by OriGene in the past, it is tempting to hypothesize that the issues encountered throughout this process originated with the starting material provided.

Due to the cost of 450K methylation arrays, it is unknown whether or not the methylation profile of successfully transduced cells (that demonstrated expression of Myc-tagged DNMT3B V1 and V6) was altered. 450K analysis of these cells was not pursued as the efficiency of transduction using lentivirus should not have been nearly as low as it was. The low percentage of cells expressing DNMT3B V1, and DNMT3B V6, would have skewed the 450K methylation results towards an intermediate profile that would not properly cluster with ETMRs had overexpression of DNMT3B V6 produced methylation patterns similar to those seen in ETMR samples. Again, the low percentage of successfully transduced PFSK-1 cells may have been due to cellular cytotoxicity caused by overexpression of *de novo* methyltransferases because of their capability of inducing genome wide epigenomic reorganization. Further experimentation on hNSC, the BT183 ETMR cell line and the use of mouse models will hopefully elucidate whether or not overexpression of DNMT3B V6 is responsible for the oncogenesis of ETMRs, and the distinct methylation pattern seen in this rare childhood cancer. These future studies will help facilitate the creation of novel therapies, such as inhibitors of the *de novo* methyltransferase DNMT3B, for this deadly malignant pediatric brain cancer.

3.5 - References

1. Okano, M., Bell, D. W., Haber, D. A. & Li, E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99,** 247–257 (1999).

2. Kleinman, C. L. *et al*. Fusion of TTYH1 with the C19MC microRNA cluster drives expression of a brain-specific DNMT3B isoform in the embryonal brain tumor ETMR. *Nat Genet* **46,** 39–44 (2014).

3. Miller, R. W., Young Jr., J. L. & Novakovic, B. Childhood cancer. *Cancer* **75,** 395–405 (1995).

4. Pfister, S. M. *et al*. Molecular diagnostics of CNS embryonal tumors. *Acta Neuropathol*. **120,** 553–566 (2010).

5. McGovern, S. L., Grosshans, D. & Mahajan, A. Embryonal brain tumors. *Cancer J* **20,** 397–402 (2014).

6. Al-Halabi, H. *et al*. Preponderance of sonic hedgehog pathway activation characterizes adult medulloblastoma. *Acta Neuropathol* **121,** 229–239 (2011).

7. Maleci, A., Cervoni, L. & Delfini, R. Medulloblastoma in children and in adults: a comparative study. *Acta Neurochir* **119,** 62–67 (1992).

8. Packer, R. J., Macdonald, T., Vezina, G., Keating, R. & Santi, M. *Medulloblastoma and primitive neuroectodermal tumors*. *Handbook of clinical neurology* **105,** (2012).

9. Louis, D. N. *et al*. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* **114,** 97–109 (2007).

10. Northcott, P. A. *et al*. Medulloblastoma comprises four distinct molecular variants. *J Clin Oncol* **29,** 1408–1414 (2011).

11. Wang, X. *et al*. Medulloblastoma subgroups remain stable across primary and metastatic compartments. *Acta Neuropathol* **129,** 449–457 (2015).

12. Clifford, S. C. *et al*. Wnt/Wingless Pathway Activation and Chromosome 6 Loss Characterise a Distinct Molecular Sub-Group of Medulloblastomas Associated with a Favourable Prognosis. *Cell Cycle* **5,** 2666–2670 (2006).

13. Taylor, M. D. *et al*. Molecular subgroups of medulloblastoma: The current consensus. *Acta Neuropathol*. **123,** 465–472 (2012).

14. Satir, P., Pedersen, L. B. & Christensen, S. T. The primary cilium at a glance. *J. Cell Sci*. **123,** 499–503 (2010).

15. Mobassarah, N. J. *et al*. Sonic Hedgehog Signalling Pathway: A Complex Network. *Ann. Neurosci*. **21,** 19–22 (2014).

16. Denef, N., Neubüser, D., Perez, L. & Cohen, S. M. Hedgehog induces opposite changes in turnover and subcellular localization of patched and smoothened. *Cell* **102,** 521–531 (2000).

17. Low, J. a. & De Sauvage, F. J. Clinical experience with hedgehog pathway inhibitors. *J. Clin. Oncol*. **28,** 5321–5326 (2010).

18. Spassky, N. *et al*. Primary cilia are required for cerebellar development and Shh-dependent expansion of progenitor pool. *Dev. Biol*. **317,** 246–59 (2008).

19. Raffel, C. *et al*. Sporadic medulloblastomas contain PTCH mutations. *Cancer Res* **57,** 842–845 (1997).

20. Zurawel, R. H., Allen, C., Wechsler-Reya, R., Scott, M. P. & Raffel, C. Evidence that haploinsufficiency of Ptch leads to medulloblastoma in mice. *Genes. Chromosomes Cancer* **28,** 77–81 (2000).

21. ICH. International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. in *Clinical Investiagion of Medicinal Products in the Pediatric Population E11* (2000).

22. Kool, M. *et al*. Genome sequencing of SHH medulloblastoma predicts genotype-related response to smoothened inhibition. *Cancer Cell* **25,** 393–405 (2014).

23. Zhukova, N. *et al*. Subgroup-specific prognostic implications of TP53 mutation in medulloblastoma. *J. Clin. Oncol*. **31,** 2927–2935 (2013).

24. Ramaswamy, V. *et al*. Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. *Acta Neuropathol*. (2016). doi:10.1007/s00401-016-1569-6

25. Kool, M. *et al*. Integrated genomics identifies five medulloblastoma subtypes with distinct genetic profiles, pathway signatures and clinicopathological features. *PLoS One* **3,** e3088 (2008).

26. Niehrs, C. The complex world of WNT receptor signalling. *Nat. Rev. Mol. Cell Biol*. **13,** 767–79 (2012).

27. Logan, C. Y. & Nusse, R. The Wnt signaling pathway in development and disease. *Annu. Rev. Cell Dev. Biol*. **20,** 781–810 (2004).

28. González-Sancho, J. M., Brennan, K. R., Castelo-Soccio, L. a & Brown, A. M. C. Wnt proteins induce dishevelled phosphorylation via an LRP5/6- independent mechanism,

irrespective of their ability to stabilize beta-catenin. *Mol. Cell. Biol.* **24,** 4757–4768 (2004).

29. Holland, J. D., Klaus, A., Garratt, A. N. & Birchmeier, W. Wnt signaling in stem and cancer stem cells. *Curr. Opin. Cell Biol.* **25,** 254–64 (2013).

30. Moon, R. T., Kohn, A. D., De Ferrari, G. V & Kaykas, A. WNT and beta-catenin signalling: diseases and therapies. *Nat. Rev. Genet.* **5,** 691–701 (2004).

31. Inestrosa, N. C. & Arenas, E. Emerging roles of Wnts in the adult nervous system. *Nat. Rev. Neurosci.* **11,** 77–86 (2010).

32. Ellison, D. W. *et al.* Definition of disease-risk stratification groups in childhood medulloblastoma using combined clinical, pathologic, and molecular variables. *J. Clin. Oncol.* **29,** 1400–7 (2011).

33. Kool, M. *et al.* Molecular subgroups of medulloblastoma: An international meta-analysis of transcriptome, genetic aberrations, and clinical data of WNT, SHH, Group 3, and Group 4 medulloblastomas. *Acta Neuropathol.* **123,** 473–484 (2012).

34. Pugh, T. J. *et al.* Medulloblastoma exome sequencing uncovers subtype-specific somatic mutations. *Nature* **488,** 106–10 (2012).

35. DeSouza, R.-M., Jones, B. R. T., Lowis, S. P. & Kurian, K. M. Pediatric medulloblastoma - update on molecular classification driving targeted therapies. *Front. Oncol.* **4,** 176 (2014).

36. Kunder, R. *et al.* Real-time PCR assay based on the differential expression of microRNAs and protein-coding genes for molecular classification of formalin-fixed paraffin embedded medulloblastomas. *Neuro. Oncol.* **15,** 1644–51 (2013).

37. Carvalho, C. M. B. & Lupski, J. R. Copy number variation at the breakpoint region of isochromosome 17q. *Genome Res.* **18,** 1724–1732 (2008).

38. Milde, T. *et al.* Histone Deacetylases 5 and 9 in Medulloblastoma: Novel Markers for Risk Stratification and Role in Tumor Cell Growth. *Clin.Cancer Res* (2010).

39. Verma, V. *et al.* Atypical teratoid rhabdoid tumor: long-term survival after chemoradiotherapy. *Child's Nerv. Syst.* 1–7 (2015). doi:10.1007/s00381-015-2723-5

40. Versteege, I. *et al.* Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394,** 203–206 (1998).

41. Rorke, L. B., Packer, R. & Biegel, J. Central nervous system atypical teratoid/rhabdoid tumors of infancy and childhood. *J Neurosurg* **85,** 56–65 (1996).

42. Biegel, J. Molecular genetics of atypical teratoid/rhabdoid tumors. *Neurosurg. Focus* **20,** 1–7 (2006).

43. Krishnan, C., Vogel, H. & Perry, A. Atypical teratoid/rhabdoid tumor with ganglioglioma-like differentiation: case report and review of the literature. *Hum. Pathol.* **45,** 185–8 (2014).

44. Sévenet, N. *et al*. Spectrum of hSNF5/INI1 somatic mutations in human cancer and genotype-phenotype correlations. *Hum. Mol. Genet.* **8,** 2359–68 (1999).

45. Sansam, C. G. & Roberts, C. W. M. Epigenetics and cancer: Altered chromatin remodeling via Snf5 loss leads to aberrant cell cycle regulation. *Cell Cycle* **5,** 621–624 (2006).

46. Vries, R. G. J. *et al*. Cancer-associated mutations in chromatin remodeler hSNF5 promote chromosomal instability by compromising the mitotic checkpoint. *Genes Dev.* **19,** 665–70 (2005).

47. Korshunov, A. *et al*. Embryonal tumor with abundant neuropil and true rosettes (ETANTR), ependymoblastoma, and medulloepithelioma share molecular similarity and comprise a single clinicopathological entity. *Acta Neuropathol* **128,** 279–289 (2014).

48. Packer, R. J. Brain tumors in children. *Arch Neurol* **56,** 421–425 (1999).

49. Peifer, M. *et al*. Telomerase activation by genomic rearrangements in high-risk neuroblastoma. *Nature* **526,** 700–4 (2015).

50. Picard, D. *et al*. Markers of survival and metastatic potential in childhood CNS primitive neuro-ectodermal brain tumours: An integrative genomic analysis. *Lancet Oncol.* **13,** 838–848 (2012).

51. Reifenberger, J. *et al*. Missense mutations in SMOH in sporadic basal cell carcinomas of the skin and primitive neuroectodermal tumors of the central nervous system. *Cancer Res.* **58,** 1798–803 (1998).

52. Molenaar, J. J. *et al*. Cyclin D1 and CDK4 activity contribute to the undifferentiated phenotype in neuroblastoma. *Cancer Res.* **68,** 2599–2609 (2008).

53. Gessi, M. *et al*. H3.3 G34R mutations in pediatric primitive neuroectodermal tumors of central nervous system (CNS-PNET) and pediatric glioblastomas: possible diagnostic and therapeutic implications? *J. Neurooncol.* **112,** 67–72 (2013).

54. Eberhart, C. G., Brat, D. J., Cohen, K. J. & Burger, P. C. Pediatric neuroblastic brain tumors containing abundant neuropil and true rosettes. *Pediatr Dev Pathol* **3,** 346–352 (2000).

55. Louis, D. N. *et al*. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol* **114,** 97–109 (2007).

56. Manjila, S. *et al*. Embryonal tumors with abundant neuropil and true rosettes: 2 illustrative cases and a review of the literature. *Neurosurg. Focus* **30,** E2 (2011).

57. Adamek, D., Sofowora, K. D., Cwiklinska, M., Herman-Sucharska, I. & Kwiatkowski, S. Embryonal tumor with abundant neuropil and true rosettes: an autopsy case-based update and review of the literature. *Childs. Nerv. Syst.* **29,** 849–54 (2013).

58. Gessi, M. *et al*. Embryonal tumors with abundant neuropil and true rosettes: a distinctive CNS primitive neuroectodermal tumor. *Am. J. Surg. Pathol.* **33,** 211–7 (2009).

59. Pfister, S. *et al*. Novel genomic amplification targeting the microRNA cluster at 19q13.42 in a pediatric embryonal tumor with abundant neuropil and true rosettes. *Acta Neuropathol.* **117,** 457–464 (2009).

60. Korshunov, A. *et al*. Focal genomic amplification at 19q13.42 comprises a powerful diagnostic marker for embryonal tumors with ependymoblastic rosettes. *Acta Neuropathol.* **120,** 253–60 (2010).

61. Korshunov, A. *et al*. LIN28A immunoreactivity is a potent diagnostic marker of embryonal tumor with multilayered rosettes (ETMR). *Acta Neuropathol.* **124,** 875–881 (2012).

62. Ceccom, J. *et al*. Embryonal tumor with multilayered rosettes: diagnostic tools update and review of the literature. *Clin. Neuropathol.* **33,** 15–22 (2014).

63. Shyh-Chang, N. & Daley, G. Q. Lin28: Primal regulator of growth and metabolism in stem cells. *Cell Stem Cell* **12,** 395–406 (2013).

64. Ambros, V. & Horvitz, H. R. Heterochronic mutants of the nematode Caenorhabditis elegans. *Science* **226,** 409–416 (1984).

65. Moss, E. G. & Tang, L. Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Dev. Biol.* **258,** 432–442 (2003).

66. Heo, I. *et al*. Lin28 Mediates the Terminal Uridylation of let-7 Precursor MicroRNA. *Mol. Cell* **32,** 276–284 (2008).

67. Piskounova, E. *et al*. Lin28A and Lin28B inhibit let-7 MicroRNA biogenesis by distinct mechanisms. *Cell* **147,** 1066–1079 (2011).

68. Yong, S. L. & Dutta, A. The tumor suppressor microRNA let-7 represses the HMGA2 oncogene. *Genes Dev.* **21,** 1025–1030 (2007).

69. Spence, T. *et al*. A novel C19MC amplified cell line links Lin28/let-7 to mTOR signaling in embryonal tumor with multilayered rosettes. *Neuro Oncol* **16,** 62–71 (2014).

70. Hovestadt, V. *et al*. Decoding the regulatory landscape of medulloblastoma using DNA methylation sequencing. *Nature* **510,** 537–41 (2014).

71. Nobusawa, S. *et al*. Analysis of chromosome 19q13.42 amplification in embryonal brain tumors with ependymoblastic multilayered rosettes. *Brain Pathol*. **22,** 689–97 (2012).

72. Li, M. *et al*. Frequent Amplification of a chr19q13.41 MicroRNA Polycistron in Aggressive Primitive Neuroectodermal Brain Tumors. *Cancer Cell* **16,** 533–546 (2009).

73. Bentwich, I. *et al*. Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet*. **37,** 766–770 (2005).

74. Morales-Prieto, D. M., Ospina-Prieto, S., Chaiwangyen, W., Schoenleben, M. & Markert, U. R. Pregnancy-associated miRNA-clusters. *J. Reprod. Immunol*. **97,** 51–61 (2013).

75. Ren, J., Jin, P., Wang, E., Marincola, F. M. & Stroncek, D. F. MicroRNA and gene expression patterns in the differentiation of human embryonic stem cells. *J. Transl. Med*. **7,** 20 (2009).

76. Bar, M. *et al*. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. *Stem Cells* **26,** 2496–2505 (2008).

77. Ambros, V. MicroRNA pathways in flies and worms: Growth, death, fat, stress, and timing. *Cell* **113,** 673–676 (2003).

78. Tutar, L., Tutar, E. & Tutar, Y. MicroRNAs and cancer; an overview. *Curr. Pharm. Biotechnol*. **15,** 430–7 (2014).

79. Seitz, H., Ghildiyal, M. & Zamore, P. D. Argonaute loading improves the 5' precision of both MicroRNAs and their miRNA strands in flies. *Curr Biol* **18,** 147–151 (2008).

80. Dykxhoorn, D., Novina, C. & Sharp, P. Killing the messenger: short RNAs that silence gene expression. *Nat. Rev. Mol. cell Biol*. **4,** 457–67 (2003).

81. Chan, J. A., Krichevsky, A. M. & Kosik, K. S. MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res*. **65,** 6029–6033 (2005).

82. Buscaglia, L. E. B. & Li, Y. Apoptosis and the target genes of microRNA-21. *Chin. J. Cancer* **30,** 371–80 (2011).

83. Wang, W. & Luo, Y. MicroRNAs in breast cancer: oncogene and tumor suppressors with clinical potential. *J. Zhejiang Univ. Sci. B* **16,** 18–31 (2015).

84. Kwak, P. B., Iwasaki, S. & Tomari, Y. The microRNA pathway and cancer. *Cancer Science* **101,** 2309–2315 (2010).

85. Ma, L., Teruya-Feldstein, J. & Weinberg, R. a. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature* **449,** 682–8 (2007).

86. Uziel, T. *et al*. The miR-17 92 cluster collaborates with the Sonic Hedgehog pathway in medulloblastoma. *Proc. Natl. Acad. Sci.* **106,** 2812–2817 (2009).

87. Mendell, J. T. miRiad Roles for the miR-17-92 Cluster in Development and Disease. *Cell* **133,** 217–222 (2008).

88. Faraoni, I., Antonetti, F. R., Cardone, J. & Bonmassar, E. miR-155 gene: A typical multifunctional microRNA. *Biochimica et Biophysica Acta - Molecular Basis of Disease* **1792,** 497–505 (2009).

89. Voorhoeve, P. M. *et al*. A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. in *Advances in Experimental Medicine and Biology* **604,** 17–46 (2007).

90. Medina, P. P. & Slack, F. J. microRNAs and cancer: an overview. *Cell Cycle* **7,** 2485–92 (2008).

91. Calin, G. A. *et al*. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 15524–9 (2002).

92. Raveche, E. S. *et al*. Abnormal microRNA-16 locus with synteny to human 13q14 linked to CLL in NZB mice. *Blood* **109,** 5079–5086 (2007).

93. Pekarsky, Y. & Croce, C. M. Role of miR-15/16 in CLL. *Cell Death Differ*. 1–6 (2014). doi:10.1038/cdd.2014.87

94. Johnson, S. M. *et al*. RAS is regulated by the let-7 microRNA family. *Cell* **120,** 635–47 (2005).

95. Chen, P.-S., Su, J.-L. & Hung, M.-C. Dysregulation of microRNAs in cancer. *J. Biomed. Sci*. **19,** 90 (2012).

96. Kumar, M. S. *et al*. Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 3903–8 (2008).

97. Garzon, R. *et al*. MicroRNA 29b functions in acute myeloid leukemia. *Blood* **114,** 5331–5341 (2009).

98. Fiserova, B., Kubiczkova, L., Sedlarikova, L., Hajek, R. & Sevcikova, S. The miR-29 family in hematological malignancies. *Biomed. Pap. Med. Fac. Univ. Palacky. Olomouc. Czech. Repub.* **158,** 1–8 (2014).

99. Welch, C., Chen, Y. & Stallings, R. L. MicroRNA-34a functions as a potential tumor suppressor by inducing apoptosis in neuroblastoma cells. *Oncogene* **26,** 5017–22 (2007).

100. Chang, T.-C. *et al*. Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. *Mol. Cell* **26,** 745–52 (2007).

101. Chang, C.-J. *et al*. p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. *Nat. Cell Biol.* **13,** 317–23 (2011).

102. Paroo, Z., Ye, X., Chen, S. & Liu, Q. Phosphorylation of the Human MicroRNA-Generating Complex Mediates MAPK/Erk Signaling. *Cell* **139,** 112–122 (2009).

103. Indovina, P., Marcelli, E., Casini, N., Rizzo, V. & Giordano, A. Emerging roles of RB family: New defense mechanisms against tumor progression. *Journal of Cellular Physiology* **228,** 525–535 (2013).

104. Knudson, A. & G. Mutation and Cancer: Statistical Study of Retinoblastoma. *Proc. Natl. Acad. Sci. U. S. A.* **68,** 820–823 (1971).

105. Horowitz, J. M. *et al*. Frequent inactivation of the retinoblastoma anti-oncogene is restricted to a subset of human tumor cells. *Proc. Natl. Acad. Sci. U. S. A.* **87,** 2775–9 (1990).

106. Lee, E. Y. *et al*. Inactivation of the retinoblastoma susceptibility gene in human breast cancers. *Science* **241,** 218–21 (1988).

107. Harbour, J. W. *et al*. Abnormalities in structure and expression of the human retinoblastoma gene in SCLC. *Science (80-.).* **241,** 353–357 (1988).

108. Sun, A., Bagella, L., Tutton, S., Romano, G. & Giordano, A. From G0 to S phase: A view of the roles played by the retinoblastoma (Rb) family members in the Rb-E2F pathway. *Journal of Cellular Biochemistry* **102,** 1400–1404 (2007).

109. Buchkovich, K., Duffy, L. A. & Harlow, E. The retinoblastoma protein is phosphorylated during specific phases of the cell cycle. *Cell* **58,** 1097–1105 (1989).

110. Narita, M. *et al*. Rb-mediated heterochromatin formation and silencing of E2F target genes during cellular senescence. *Cell* **113,** 703–716 (2003).

111. McConnell, B. B., Gregory, F. J., Stott, F. J., Hara, E. & Peters, G. Induced expression of p16(INK4a) inhibits both CDK4- and CDK2-associated kinase activity by reassortment of cyclin-CDK-inhibitor complexes. *Mol. Cell. Biol.* **19,** 1981–1989 (1999).

112. Bisteau, X. *et al*. CDK4 T172 Phosphorylation Is Central in a CDK7-Dependent Bidirectional CDK4/CDK2 Interplay Mediated by p21 Phosphorylation at the Restriction Point. *PLoS Genet*. **9,** (2013).

113. Lapenna, S. & Giordano, A. Cell cycle kinases as therapeutic targets for cancer. *Nat. Rev. Drug Discov*. **8,** 547–566 (2009).

114. Dick, F. a & Rubin, S. M. Molecular mechanisms underlying RB protein function. *Nat. Rev. Mol. Cell Biol*. **14,** 297–306 (2013).

115. Hassler, M. *et al*. Crystal structure of the retinoblastoma protein N domain provides insight into tumor suppression, ligand interaction, and holoprotein architecture. *Mol. Cell* **28,** 371–85 (2007).

116. Ajioka, I. Coordination of proliferation and neuronal differentiation by the retinoblastoma protein family. *Development Growth and Differentiation* **56,** 324–334 (2014).

117. De Luca, A. *et al*. A unique domain of pRb2/p130 acts as an inhibitor of Cdk2 kinase activity. *J. Biol. Chem*. **272,** 20971–20974 (1997).

118. Genovese, C., Trani, D., Caputi, M. & Claudio, P. P. Cell cycle control and beyond: emerging roles for the retinoblastoma gene family. *Oncogene* **25,** 5201–9 (2006).

119. Gabellini, C., del Bufalo, D. & Zupi, G. Involvement of RB gene family in tumor angiogenesis. *Oncogene* **25,** 5326–5332 (2006).

120. Burkhart, D. L. & Sage, J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat. Rev. Cancer* **8,** 671–682 (2008).

121. Mayhew, C. N. *et al*. Liver-specific pRB loss results in ectopic cell cycle entry and aberrant ploidy. *Cancer Res*. **65,** 4568–4577 (2005).

122. Van Harn, T. *et al*. Loss of Rb proteins causes genomic instability in the absence of mitogenic signaling. *Genes Dev*. **24,** 1377–1388 (2010).

123. Gonzalo, S. *et al*. Role of the RB1 family in stabilizing histone methylation at constitutive heterochromatin. *Nat. Cell Biol*. **7,** 420–8 (2005).

124. Campisi, J. Cellular senescence as a tumor-suppressor mechanism. *Trends in Cell Biology* **11,** (2001).

125. Xu, H. J. *et al*. Reexpression of the retinoblastoma protein in tumor cells induces senescence and telomerase inhibition. *Oncogene* **15,** 2589–96 (1997).

126. Serrano, M., Hannon, G. J. & Beach, D. A new regulatory motif in cell-cycle control

causing specific inhibition of cyclin D/CDK4. *Nature* **366,** 704–707 (1993).

127. Steiner, M. S., Wang, Y., Zhang, Y., Zhang, X. & Lu, Y. p16/MTS1/INK4A suppresses prostate cancer by both pRb dependent and independent pathways. *Oncogene* **19,** 1297–306 (2000).

128. Helmbold, H., Kömm, N., Deppert, W. & Bohn, W. Rb2/p130 is the dominating pocket protein in the p53-p21 DNA damage response pathway leading to senescence. *Oncogene* **28,** 3456–3467 (2009).

129. Kapić, a *et al*. Cooperation between p53 and p130(Rb2) in induction of cellular senescence. *Cell Death Differ*. **13,** 324–334 (2006).

130. Jackson, J. G. & Pereira-Smith, O. M. Primary and compensatory roles for RB family members at cell cycle gene promoters that are deacetylated and downregulated in doxorubicin-induced senescence of breast cancer cells. *Mol. Cell. Biol.* **26,** 2501–10 (2006).

131. Dimaras, H. *et al*. Loss of RB1 induces non-proliferative retinoma: Increasing genomic instability correlates with progression to retinoblastoma. *Hum. Mol. Genet*. **17,** 1363–1372 (2008).

132. Chau, B. N. & Wang, J. Y. J. Coordinated regulation of life and death by RB. *Nat. Rev. Cancer* **3,** 130–8 (2003).

133. Wang, R. H., Liu, C. W., Avramis, V. I. & Berndt, N. Protein phosphatase 1alpha-mediated stimulation of apoptosis is associated with dephosphorylation of the retinoblastoma protein. *Oncogene* **20,** 6111–22 (2001).

134. Bowen, C., Spiegel, S. & Gelmann, E. P. Radiation-induced apoptosis mediated by retinoblastoma protein. *Cancer Res*. **58,** 3275–3281 (1998).

135. Jiang, Y., Saavedra, H. I., Holloway, M. P., Leone, G. & Altura, R. A. Aberrant regulation of survivin by the RB/E2F family of proteins. *J. Biol. Chem*. **279,** 40511–40520 (2004).

136. Skapek, S. X., Pan, Y.-R. & Lee, E. Y.-H. P. Regulation of cell lineage specification by the retinoblastoma tumor suppressor. *Oncogene* **25,** 5268–76 (2006).

137. Bellan, C. *et al*. Missing expression of pRb2/p130 in human retinoblastomas is associated with reduced apoptosis and lesser differentiation. *Invest. Ophthalmol. Vis. Sci*. **43,** 3602–8 (2002).

138. Kiess, M., Gill, R. M. & Hamel, P. A. Expression and activity of the retinoblastoma protein (pRB)-family proteins, p107 and p130, during L6 myoblast differentiation. *Cell Growth Differ*. **6,** 1287–98 (1995).

139. Stanelle, J., Stiewe, T., Theseling, C. C., Peter, M. & Pützer, B. M. Gene expression changes in response to E2F1 activation. *Nucleic Acids Res*. **30,** 1859–1867 (2002).

140. Claudio, P. P. *et al*. RB2/p130 gene-enhanced expression down-regulates vascular endothelial growth factor expression and inhibits angiogenesis in vivo. *Cancer Res* **61,** 462–468 (2001).

141. Bellacchio, E. & Paggi, M. G. Understanding the targeting of the RB family proteins by viral oncoproteins to defeat their oncogenic machinery. *J. Cell. Physiol*. **228,** 285–291 (2013).

142. Schaal, C., Pillai, S. & Chellappan, S. P. The Rb-E2F transcriptional regulatory pathway in tumor angiogenesis and metastasis. *Adv. Cancer Res*. **121,** 147–174 (2014).

143. Hui, A. M., Li, X., Makuuchi, M., Takayama, T. & Kubota, K. Over-expression and lack of retinoblastoma protein are associated with tumor progression and metastasis in hepatocellular carcinoma. *Int J.Cancer* **84,** 604–608 (1999).

144. Schaffer, B. E. *et al*. Loss of p130 accelerates tumor development in a mouse model for human small-cell lung carcinoma. *Cancer Res*. **70,** 3877–83 (2010).

145. Liu, F. *et al*. MicroRNA-106b-5p boosts glioma tumorigensis by targeting multiple tumor suppressor genes. *Oncogene* 1–10 (2013). doi:10.1038/onc.2013.428

146. Wang, Q. *et al*. miR-17-92 cluster accelerates adipocyte differentiation by negatively regulating tumor-suppressor Rb2/p130. *Proc. Natl. Acad. Sci. U. S. A*. **105,** 2889–94 (2008).

147. Lu, F., Stedman, W., Yousef, M., Renne, R. & Lieberman, P. M. Epigenetic regulation of Kaposi's sarcoma-associated herpesvirus latency by virus-encoded microRNAs that target Rta and the cellular Rbl2-DNMT pathway. *J. Virol*. **84,** 2697–2706 (2010).

148. Böhlig, L., Friedrich, M. & Engeland, K. P53 activates the PANK1/miRNA-107 gene leading to downregulation of CDK6 and p130 cell cycle proteins. *Nucleic Acids Res*. **39,** 440–453 (2011).

149. Houbaviy, H. B., Murray, M. F. & Sharp, P. A. Embryonic stem cell-specific microRNAs. *Dev. Cell* **5,** 351–358 (2003).

150. Blasco, M. *et al*. A mammalian microRNA cluster controls DNA methylation and telomere recombination via Rbl2-dependent regulation of DNA methyltransferases. *Nat. Struct. Mol. Biol*. **15,** 268–79 (2008).

151. McCabe, M. T., Davis, J. N. & Day, M. L. Regulation of DNA methyltransferase 1 by the pRb/E2F1 pathway. *Cancer Res*. **65,** 3624–3632 (2005).

152. Sinkkonen, L. *et al*. MicroRNAs control de novo DNA methylation through regulation of transcriptional repressors in mouse embryonic stem cells. *Nat. Struct. Mol. Biol.* **15,** 259–267 (2008).

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

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

155. Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13,** 484–92 (2012).

156. Rideout, W. M., Coetzee, G. A., Olumi, A. F. & Jones, P. A. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science (80-.).* **249,** 1288–1290 (1990).

157. Hahn, M. A., Wu, X., Li, A. X., Hahn, T. & Pfeifer, G. P. Relationship between gene body DNA methylation and intragenic H3K9ME3 and H3K36ME3 chromatin marks. *PLoS One* **6,** (2011).

158. Laurent, L. *et al*. Dynamic changes in the human methylome during differentiation. *Genome Res.* **20,** 320–331 (2010).

159. Maunakea, A. K. *et al*. Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* **466,** 253–257 (2010).

160. Lay, F. D. *et al*. The role of DNA methylation in directing the functional organization of the cancer epigenome. *Genome Res.* **25,** 467–477 (2015).

161. Guo, J. U., Su, Y., Zhong, C., Ming, G. L. & Song, H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell* **145,** 423–434 (2011).

162. Carson, S., Wilson, J., Aksimentiev, A., Weigele, P. R. & Wanunu, M. Hydroxymethyluracil modifications enhance the flexibility and hydrophilicity of double-stranded DNA. *Nucleic Acids Res.* gkv1199– (2015). doi:10.1093/nar/gkv1199

163. Cortellino, S. *et al*. Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146,** 67–79 (2011).

164. Gu, T.-P. *et al*. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477,** 606–610 (2011).

165. Carless, M. A. Determination of DNA methylation levels using illumina HumanMethylation450 BeadChips. *Methods Mol. Biol.* **1288,** 143–192 (2015).

166. Morris, T. J. & Beck, S. Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data. *Methods* **72,** 3–8 (2015).

167. Bibikova, M. *et al*. High density DNA methylation array with single CpG site resolution. *Genomics* **98,** 288–295 (2011).

168. Park, H.-J., Yu, E. & Shim, Y.-H. DNA methyltransferase expression and DNA hypermethylation in human hepatocellular carcinoma. *Cancer Lett*. **233,** 271–8 (2006).

169. Miremadi, A., Oestergaard, M. Z., Pharoah, P. D. P. & Caldas, C. Cancer genetics of epigenetic genes. *Human Molecular Genetics* **16,** (2007).

170. Tuorto, F. *et al*. The tRNA methyltransferase Dnmt2 is required for accurate polypeptide synthesis during haematopoiesis. *EMBO J*. **34,** 2350–62 (2015).

171. Subramaniam, D., Thombre, R., Dhar, A. & Anant, S. DNA methyltransferases: a novel target for prevention and therapy. *Front. Oncol*. **4,** 80 (2014).

172. Zhang, Z.-M. *et al*. Crystal Structure of Human DNA Methyltransferase 1. *J. Mol. Biol*. **427,** 2520–2531 (2015).

173. Goll, M. G. *et al*. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. *Science (80-.).* **311,** 395–8 (2006).

174. Hermann, A., Schmitt, S. & Jeltsch, A. The human Dnmt2 has residual DNA-(Cytosine-C5) methyltransferase activity. *J. Biol. Chem*. **278,** 31717–31721 (2003).

175. Jurkowska, R. Z., Jurkowski, T. P. & Jeltsch, A. Structure and Function of Mammalian DNA Methyltransferases. *ChemBioChem* **12,** 206–222 (2011).

176. Liao, H. F. *et al*. Functions of DNA methyltransferase 3-like in germ cells and beyond. *Biology of the Cell* **104,** 571–587 (2012).

177. Chen, T., Tsujimoto, N. & Li, E. The PWWP domain of Dnmt3a and Dnmt3b is required for directing DNA methylation to the major satellite repeats at pericentric heterochromatin. *Mol. Cell. Biol*. **24,** 9048–58 (2004).

178. Ooi, S. K. T. *et al*. DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448,** 714–7 (2007).

179. Wienholz, B. L. *et al*. DNMT3L modulates significant and distinct flanking sequence preference for DNA methylation by DNMT3A and DNMT3B in vivo. *PLoS Genet*. **6,** (2010).

180. Matsuoka, T. *et al*. DNA methyltransferase-3 like protein expression in various

histological types of testicular germ cell tumor. *Jpn. J. Clin. Oncol.* (2016). doi:10.1093/jjco/hyw012

181. Van Emburgh, B. O. & Robertson, K. D. Modulation of Dnmt3b function in vitro by interactions with Dnmt3L, Dnmt3a and Dnmt3b splice variants. *Nucleic Acids Res.* **39,** 4984–5002 (2011).

182. Handa, V. & Jeltsch, A. Profound flanking sequence preference of Dnmt3a and Dnmt3b mammalian DNA methyltransferases shape the human epigenome. *J. Mol. Biol.* **348,** 1103–1112 (2005).

183. Meissner, A. *et al*. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454,** 766–70 (2008).

184. Hitchins, M. P. *et al*. Dominantly Inherited Constitutional Epigenetic Silencing of MLH1 in a Cancer-Affected Family Is Linked to a Single Nucleotide Variant within the 5'UTR. *Cancer Cell* **20,** 200–213 (2011).

185. Schlesinger, Y. *et al*. Polycomb-mediated methylation on Lys27 of histone H3 pre-marks genes for de novo methylation in cancer. *Nat. Genet.* **39,** 232–236 (2007).

186. Fuks, F., Hurd, P. J., Deplus, R. & Kouzarides, T. The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res.* **31,** 2305–2312 (2003).

187. Viré, E. *et al*. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439,** 871–874 (2005).

188. Li, H. *et al*. The histone methyltransferase SETDB1 and the DNA methyltransferase DNMT3A interact directly and localize to promoters silenced in cancer cells. *J. Biol. Chem.* **281,** 19489–19500 (2006).

189. Jeong, S. *et al*. Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol. Cell. Biol.* **29,** 5366–76 (2009).

190. Nakagaki, A., Osanai, H. & Kishino, T. Imprinting analysis of the mouse chromosome 7C region in DNMT1-null embryos. *Gene* **553,** 63–68 (2014).

191. Esteller, M. Relevance of DNA methylation in the management of cancer. *Lancet Oncology* **4,** 351–358 (2003).

192. Rhee, I. *et al*. CpG methylation is maintained in human cancer cells lacking DNMT1. *Nature* **404,** 1003–1007 (2000).

193. Linhart, H. G. *et al*. Dnmt3b promotes tumorigenesis in vivo by gene-specific de novo methylation and transcriptional silencing. *Genes Dev.* **21,** 3110–3122 (2007).

194. Wang, W., Gao, J., Man, X. H., Li, Z. S. & Gong, Y. F. Significance of DNA methyltransferase-1 and histone deacetylase-1 in pancreatic cancer. *Oncol. Rep.* **21,** 1439–1447 (2009).

195. Zhang, J. J. *et al.* Association of increased DNA methyltransferase expression with carcinogenesis and poor prognosis in pancreatic ductal adenocarcinoma. *Clin. Transl. Oncol.* **14,** 116–124 (2012).

196. Girault, I., Tozlu, S., Lidereau, R. & Bièche, I. Expression analysis of DNA methyltransferases 1, 3A, and 3B in sporadic breast carcinomas. *Clin. Cancer Res.* **9,** 4415–4422 (2003).

197. Arai, E. *et al.* Regional DNA hypermethylation and DNA methyltransferase (DNMT) 1 protein overexpression in both renal tumors and corresponding nontumorous renal tissues. *Int. J. Cancer* **119,** 288–296 (2006).

198. Rajendran, G. *et al.* Epigenetic regulation of DNA methyltransferases: DNMT1 and DNMT3B in gliomas. *J. Neurooncol.* **104,** 483–494 (2011).

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

200. Minami, K. *et al.* DNMT3L is a novel marker and is essential for the growth of human embryonal carcinoma. *Clin. Cancer Res.* **16,** 2751–2759 (2010).

201. Han, H., Yang, X., Pandiyan, K. & Liang, G. Synergistic re-activation of epigenetically silenced genes by combinatorial inhibition of DNMTs and LSD1 in cancer cells. *PLoS One* **8,** e75136 (2013).

202. Beaulieu, N. *et al.* An essential role for DNA methyltransferase DNMT3B in cancer cell survival. *J. Biol. Chem.* **277,** 28176–28181 (2002).

203. Kuck, D., Caulfield, T., Lyko, F. & Medina-Franco, J. L. Nanaomycin A selectively inhibits DNMT3B and reactivates silenced tumor suppressor genes in human cancer cells. *Mol. Cancer Ther.* **9,** 3015–23 (2010).

204. Shen, H. *et al.* A novel polymorphism in human cytosine DNA-methyltransferase-3B promoter is associated with an increased risk of lung cancer. *Cancer Res* **62,** 4992–4995 (2002).

205. Lee, S. J. *et al.* DNMT3B polymorphisms and risk of primary lung cancer. *Carcinogenesis* **26,** 403–409 (2005).

206. Hernández-Sotelo, D. *et al.* The 46359CT polymorphism of DNMT3B is associated with

the risk of cervical cancer. *Mol. Biol. Rep.* **40,** 4275–80 (2013).

207. Jin, F. *et al.* Up-regulation of DNA methyltransferase 3B expression in endometrial cancers. *Gynecol. Oncol.* **96,** 531–538 (2005).

208. Micevic, G. *et al.* DNMT3b Modulates Melanoma Growth by Controlling Levels of mTORC2 Component RICTOR. *Cell Rep.* **14,** 2180–92 (2016).

209. Hlady, R. A. *et al.* Loss of Dnmt3b function upregulates the tumor modifier Ment and accelerates mouse lymphomagenesis. *J. Clin. Invest.* **122,** 163–177 (2012).

210. Jin, B. *et al.* DNMT1 and DNMT3B modulate distinct polycomb-mediated histone modifications in colon cancer. *Cancer Res.* **69,** 7412–7421 (2009).

211. Ehrlich, M. The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clin. Immunol.* **109,** 17–28 (2003).

212. Xu, G. L. *et al.* Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402,** 187–191 (1999).

213. Fryns, J. P. *et al.* Centromeric instability of chromosomes 1, 9, and 16 associated with combined immunodeficiency. *Hum. Genet.* **57,** 108–110 (1981).

214. Tiepolo, L. *et al.* Multibranched chromosomes 1, 9, and 16 in a patient with combined IgA and IgE deficiency. *Hum. Genet.* **51,** 127–37 (1979).

215. Hernandez, R., Frady, A., Zhang, X. Y., Varela, M. & Ehrlich, M. Preferential induction of chromosome 1 multibranched figures and whole-arm deletions in a human pro-B cell line treated with 5-azacytidine or 5-azadeoxycytidine. *Cytogenet. Cell Genet.* **76,** 196–201 (1997).

216. Fujimori, H. *et al.* A comprehensive analysis of radiosensitization targets; functional inhibition of DNA methyltransferase 3B radiosensitizes by disrupting DNA damage regulation. *Sci. Rep.* **5,** 18231 (2015).

217. Sandhu, R., Rivenbark, A. G. & Coleman, W. B. Enhancement of chemotherapeutic efficacy in hypermethylator breast cancer cells through targeted and pharmacologic inhibition of DNMT3b. *Breast Cancer Res. Treat.* **131,** 385–99 (2012).

218. Oka, M. *et al.* De novo DNA methyltransferases Dnmt3a and Dnmt3b primarily mediate the cytotoxic effect of 5-aza-2'-deoxycytidine. *Oncogene* **24,** 3091–3099 (2005).

Appendix 1 - Supplementary Data

Table S1 - DNMT3B Variant 1 and Variant 6 DNA Sequences

Protein	DNA Sequence (confirmed by Sanger Sequencing)
DNMT3B Variant 1	ATGAAGGGAGACACCAGGCATCTCAATGGAGAGGAGGACGCCGGCGGGAGGGAAGACTCGATC CTCGTCAACGGGGCCTGCAGCGACCAGTCCTCCGACTCGCCCCCAATCCTGGAGGCTATCCGCAC CCCGGAGATCAGAGGCCGAAGATCAAGCTCGCGACTCTCCAAGAGGGAGGTGTCCAGTCTGCTA AGCTACACACAGGACTTGACAGGCGATGGCGACGGGGAAGATGGGGATGGCTCTGACACCCCAG TCATGCCAAAGCTCTTCCGGGAAACCAGGACTCGTTCAGAAAGCCCAGCTGTCCGAACTCGAAAT AACAACAGTGTCTCCAGCCGGGAGAGGCACAGGCCTTCCCCACGTTCCACCCGAGGCCGGCAGG GCCGCAACCATGTGGACGAGTCCCCCGTGGAGTTCCCGGCTACCAGGTCCCTGAGACGGCGGGC AACAGCATCGGCAGGAACGCCATGGCCGTCCCCTCCCAGCTCTTACCTTACCATCGACCTCACAG ACGACACAGAGGACACACATGGGACGCCCCAGAGCAGCAGTACCCCCTACGCCCGCCTAGCCCA GGACAGCCAGCAGGGGGGCATGGAGTCCCCGCAGGTGGAGGCAGACAGTGGAGATGGAGACAG TTCAGAGTATCAGGATGGGAAGGAGTTTGGAATAGGGGACCTCGTGTGGGGAAAGATCAAGGGC TTCTCCTGGTGGCCCGCCATGGTGGTGTCTTGGAAGGCCACCTCCAAGCGACAGGCTATGTCTGG CATGCGGTGGGTCCAGTGGTTTGGCGATGGCAAGTTCTCCGAGGTCTCTGCAGACAAACTGGTGG CACTGGGGCTGTTCAGCCAGCACTTTAATTTGGCCACCTTCAATAAGCTCGTCTCCTATCGAAAA GCCATGTACCATGCTCTGGAGAAAGCTAGGGTGCGAGCTGGCAAGACCTTCCCCAGCAGCCCTG GAGACTCATTGGAGGACCAGCTGAAGCCCATGTTGGAGTGGGCCCACGGGGGCTTCAAGCCCAC TGGGATCGAGGGCCTCAAACCCAACAACACGCAACCAGAGAACAAGACTCGAAGACGCACAGC TGACGACTCAGCCACCTCTGACTACTGCCCCGCACCCAAGCGCCTCAAGACAAATTGCTATAACA ACGGCAAAGACCGAGGGGATGAAGATCAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACA ACAAGAGCAGCCTGGAAGATGGCTGTTTGTCTTGTGGCAGGAAAAACCCCGTGTCCTTCCACCCT CTCTTTGAGGGGGGGCTCTGTCAGACATGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGAT GACGATGGCTATCAGTCTTACTGCACTGTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCAGCAA CACGAGCTGCTGCCGGTGTTTCTGTGTGGAGTGCCTGGAGGTGCTGGTGGGCACAGGCACAGCG GCCGAGGCCAAGCTTCAGGAGCCCTGGAGCTGCTACATGTGTCTCCCGCAGCGCTGTCATGGCGT CCTGCGGCGCCGGAAGGACTGGAACGTGCGCCTGCAGGCCTTCTTCACCAGTGACACGGGGCTT GAATACGAAGCCCCCAAGCTGTACCCTGCCATTCCCGCAGCCCGAAGGCGGCCCATTCGAGTCCT GTCATTGTTTGATGGCATCGCGACAGGCTACCTAGTCCTCAAAGAGTTGGGCATAAAGGTAGGAA AGTACGTCGCTTCTGAAGTGTGTGAGGAGTCCATTGCTGTTGGAACCGTGAAGCACGAGGGGAA TATCAAATACGTGAACGACGTGAGGAACATCACAAAGAAAAATATTGAAGAATGGGGCCCATTT GACTTGGTGATTGGCGGAAGCCCATGCAACGATCTCTCAAATGTGAATCCAGCCAGGAAAGGCC TGTATGAGGGTACAGGCCGGCTCTTCTTCGAATTTTACCACCTGCTGAATTACTCACGCCCCAAG GAGGGTGATGACCGGCCGTTCTTCTGGATGTTTGAGAATGTTGTAGCCATGAAGGTTGGCGACAA GAGGGACATCTCACGGTTCCTGGAGTGTAATCCAGTGATGATTGATGCCATCAAAGTTTCTGCTG CTCACAGGGCCCGATACTTCTGGGGCAACCTACCCGGGATGAACAGGATCTTTGGCTTTCCTGTG CACTACACAGACGTGTCCAACATGGGCCGTGGTGCCCGCCAGAAGCTGCTGGGAAGGTCCTGGA GCGTGCCTGTCATCCGACACCTCTTCGCCCCTCTGAAGGACTACTTTGCATGTGAA
DNMT3B Variant 6	ATGGAACCAAGTCCTGAGCCTCCAAGCTTGGAAAGCATGAAGGGAGACACCAGGCATCTCAATG GAGAGGAGGACGCCGGCGGGAGGGAAGACTCGATCCTCGTCAACGGGGCCTGCAGCGACCAGT

(exon 1B-containing DNMT3B)	CCTCCGACTCGCCCCCAATCCTGGAGGCTATCCGCACCCCGGAGATCAGAGGCCGAAGATCAAG CTCGCGACTCTCCAAGAGGGAGGTGTCCAGTCTGCTAAGCTACACACAGGACTTGACAGGCGAT GGCGACGGGGAAGATGGGGATGGCTCTGACACCCCAGTCATGCCAAAGCTCTTCCGGGAAACCA GGACTCGTTCAGAAAGCCCAGCTGTCCGAACTCGAAATAACAACAGTGTCTCCAGCCGGGAGAG GCACAGGCCTTCCCCACGTTCCACCCGAGGCCGGCAGGGCCGCAACCATGTGGACGAGTCCCCC GTGGAGTTCCCGGCTACCAGGTCCCTGAGACGGCGGGCAACAGCATCGGCAGGAACGCCATGGC CGTCCCCTCCCAGCTCTTACCTTACCATCGACCTCACAGACGACACAGAGGACACACATGGGACG CCCCAGAGCAGCAGTACCCCCTACGCCCGCCTAGCCCAGGACAGCCAGCAGGGGGGCATGGAGT CCCCGCAGGTGGAGGCAGACAGTGGAGATGGAGACAGTTCAGAGTATCAGGATGGGAAGGAGT TTGGAATAGGGGACCTCGTGTGGGGAAAGATCAAGGGCTTCTCCTGGTGGCCCGCCATGGTGGT GTCTTGGAAGGCCACCTCCAAGCGACAGGCTATGTCTGGCATGCGGTGGGTCCAGTGGTTTGGCG ATGGCAAGTTCTCCGAGGTCTCTGCAGACAAACTGGTGGCACTGGGGCTGTTCAGCCAGCACTTT AATTTGGCCACCTTCAATAAGCTCGTCTCCTATCGAAAAGCCATGTACCATGCTCTGGAGAAAGC TAGGGTGCGAGCTGGCAAGACCTTCCCCAGCAGCCCTGGAGACTCATTGGAGGACCAGCTGAAG CCCATGTTGGAGTGGGCCCACGGGGGCTTCAAGCCCACTGGGATCGAGGGCCTCAAACCCAACA ACACGCAACCAGAGAACAAGACTCGAAGACGCACAGCTGACGACTCAGCCACCTCTGACTACTG CCCCGCACCCAAGCGCCTCAAGACAAATTGCTATAACAACGGCAAAGACCGAGGGGATGAAGAT CAGAGCCGAGAACAAATGGCTTCAGATGTTGCCAACAACAAGAGCAGCCTGGAAGATGGCTGTT TGTCTTGTGGCAGGAAAAACCCCGTGTCCTTCCACCCTCTCTTTGAGGGGGGGCTCTGTCAGACA TGCCGGGATCGCTTCCTTGAGCTGTTTTACATGTATGATGACGATGGCTATCAGTCTTACTGCACT GTGTGCTGCGAGGGCCGAGAGCTGCTGCTTTGCAGCAACACGAGCTGCTGCCGGTGTTTCTGTGT GGAGTGCCTGGAGGTGCTGGTGGGCACAGGCACAGCGGCCGAGGCCAAGCTTCAGGAGCCCTGG AGCTGTTACATGTGTCTCCCGCAGCGCTGTCATGGCGTCCTGCGGCGCCGGAAGGACTGGAACGT GCGCCTGCAGGCCTTCTTCACCAGTGACACGGGGCTTGAATATGAAGCCCCCAAGCTGTACCCTG CCATTCCCGCAGCCCGAAGGCGGCCCATTCGAGTCCTGTCATTGTTTGATGGCATCGCGACAGGC TACCTAGTCCTCAAAGAGTTGGGCATAAAGGTAGGAAAGTACGTCGCTTCTGAAGTGTGTGAGG AGTCCATTGCTGTTGGAACCGTGAAGCACGAGGGGAATATCAAATACGTGAACGACGTGAGGAA CATCACAAAGAAAAATATTGAAGAATGGGGCCCATTTGACTTGGTGATTGGCGGAAGCCCATGC AACGATCTCTCAAATGTGAATCCAGCCAGGAAAGGCCTGTATGAGGGTACAGGCCGGCTCTTCTT CGAATTTTACCACCTGCTGAATTACTCACGCCCCAAGGAGGGTGATGACCGGCCGTTCTTCTGGA TGTTTGAGAATGTTGTAGCCATGAAGGTTGGCGACAAGAGGGACATCTCACGGTTCCTGGAGTGT AATCCAGTGATGATTGATGCCATCAAAGTTTCTGCTGCTCACAGGGCCCGATACTTCTGGGGCAA CCTACCCGGGATGAACAGGCCCGTGATAGCATCAAAGAATGATAAACTCGAGCTGCAGGACTGC TTGGAATACAATAGGATAGCCAAGTTAAAGAAAGTACAGACAATAACCACCAAGTCGAACTCGA TCAAACAGGGGAAAAACCAACTTTTCCCTGTTGTCATGAATGGCAAAGAAGATGTTTTGTGGTGC ACTGAGCTCGAAAGGATCTTTGGCTTTCCTGTGCACTACACAGACGTGTCCAACATGGGCCGTGG TGCCCGCCAGAAGCTGCTGGGAAGGTCCTGGAGCGTGCCTGTCATCCGACACCTCTTCGCCCCTC TGAAGGACTACTTTGCATGTGAA

Table S2 - DNMT3B Variant 1 and Variant 6 Protein Sequence

Protein	Protein Sequence
DNMT3B Variant 1	Met K G D T R H L N G E E D A G G R E D S I L V N G A C S D Q S S D S P P I L E A I R T P E I R G R R S S S R L S K R E V S S L L S Y T Q D L T G D G D G E D G D G S D T P V Met P K L F R E T R T R S E S P A V R T R N N N S V S S R E R H R P S P R S T R G R Q G R N H V D E S P V E F P A T R S L R R R A T A S A G T P W P S P P S S Y L T I D L T D D T E D T H G T P Q S S S T P Y A R L A Q D S Q Q G G Met E S P Q V E A D S G D G D S S E Y Q D G K E F G I G D L V W G K I K G F S W W P A Met V V S W K A T S K R Q A Met S G Met R W V Q W F G D G K F S E V S A D K L V A L G L F S Q H F N L A T F N K L V S Y R K A Met Y H A L E K A R V R A G K T F P S S P G D S L E D Q L K P Met L E W A H G G F K P T G I E G L K P N N T Q P E N K T R R R T A D D S A T S D Y C P A P K R L K T N C Y N N G K D R G D E D Q S R E Q Met A S D V A N N K S S L E D G C L S C G R K N P V S F H P L F E G G L C Q T C R D R F L E L F Y Met Y D D D G Y Q S Y C T V C C E G R E L L L C S N T S C C R C F C V E C L E V L V G T G T A A E A K L Q E P W S C Y Met C L P Q R C H G V L R R R K D W N V R L Q A F F T S D T G L E Y E A P K L Y P A I P A A R R R P I R V L S L F D G I A T G Y L V L K E L G I K V G K Y V A S E V C E E S I A V G T V K H E G N I K Y V N D V R N I T K K N I E E W G P F D L V I G G S P C N D L S N V N P A R K G L Y E G T G R L F F E F Y H L L N Y S R P K E G D D R P F F W Met F E N V V A Met K V G D K R D I S R F L E C N P V Met I D A I K V S A A H R A R Y F W G N L P G Met N R I F G F P V H Y T D V S N Met G R G A R Q K L L G R S W S V P V I R H L F A P L K D Y F A C E
DNMT3B Variant 6 (exon 1B-containing DNMT3B)	Met E P S P E P P S L E S Met K G D T R H L N G E E D A G G R E D S I L V N G A C S D Q S S D S P P I L E A I R T P E I R G R R S S S R L S K R E V S S L L S Y T Q D L T G D G D G E D G D G S D T

	P V **Met** P K L F R E T R T R S E S P A V R T R N N N S V S S R E R H R P S P R S T R G R Q G R N H V D E S P V E F P A T R S L R R R A T A S A G T P W P S P P S S Y L T I D L T D D T E D T H G T P Q S S S T P Y A R L A Q D S Q Q G G **Met** E S P Q V E A D S G D G D S S E Y Q D G K E F G I G D L V W G K I K G F S W W P A **Met** V V S W K A T S K R Q A **Met** S G **Met** R W V Q W F G D G K F S E V S A D K L V A L G L F S Q H F N L A T F N K L V S Y R K A **Met** Y H A L E K A R V R A G K T F P S S P G D S L E D Q L K P **Met** L E W A H G G F K P T G I E G L K P N N T Q P E N K T R R R T A D D S A T S D Y C P A P K R L K T N C Y N N G K D R G D E D Q S R E Q **Met** A S D V A N N K S S L E D G C L S C G R K N P V S F H P L F E G G L C Q T C R D R F L E L F Y **Met** Y D D D G Y Q S Y C T V C C E G R E L L L C S N T S C C R C F C V E C L E V L V G T G T A A E A K L Q E P W S C Y **Met** C L P Q R C H G V L R R R K D W N V R L Q A F F T S D T G L E Y E A P K L Y P A I P A A R R R P I R V L S L F D G I A T G Y L V L K E L G I K V G K Y V A S E V C E E S I A V G T V K H E G N I K Y V N D V R N I T K K N I E E W G P F D L V I G G S P C N D L S N V N P A R K G L Y E G T G R L F F E F Y H L L N Y S R P K E G D D R P F F W **Met** F E N V V A **Met** K V G D K R D I S R F L E C N P V **Met** I D A I K V S A A H R A R Y F W G N L P G **Met** N R P V I A S K N D K L E L Q D C L E Y N R I A K L K K V Q T I T T K S N S I K Q G K N Q L F P V V **Met** N G K E D V L W C T E L E R I F G F P V H Y T D V S N **Met** G R G A R Q K L L G R S W S V P V I R H L F A P L K D Y F A C E

Table S3 - DNMT3B and Vector Sanger Sequencing Primers

Primer	DNA Sequence (5'-3')
DNMT3B Rev10	CAGCACACAGTGCAGTAAGA
DNMT3B 1F	GAGGACACACATGGGACG
DNMT3B 2F	GCAACCAGAGAACAAGACTCGA
DNMT3B 3F	TGTACCCTGCCATTCCCG
VP1.5	GGACTTTCCAAAATGTCG
XL39	ATTAGGACAAGGCTGGTGGG
HA IRES Rev	TGTGGCCATATTATCATCGTGTT
IRES Rev	CAACAGACCTTGCATTTCTTT
Myc/DDK/IRES (MDIR) Rev	AGGTAGGGGCTGAAGGTCAG
LR50	CAGAGGTTGATTATCGATAAG
VP2.0	AGAGCTCGTTTAGTGAA

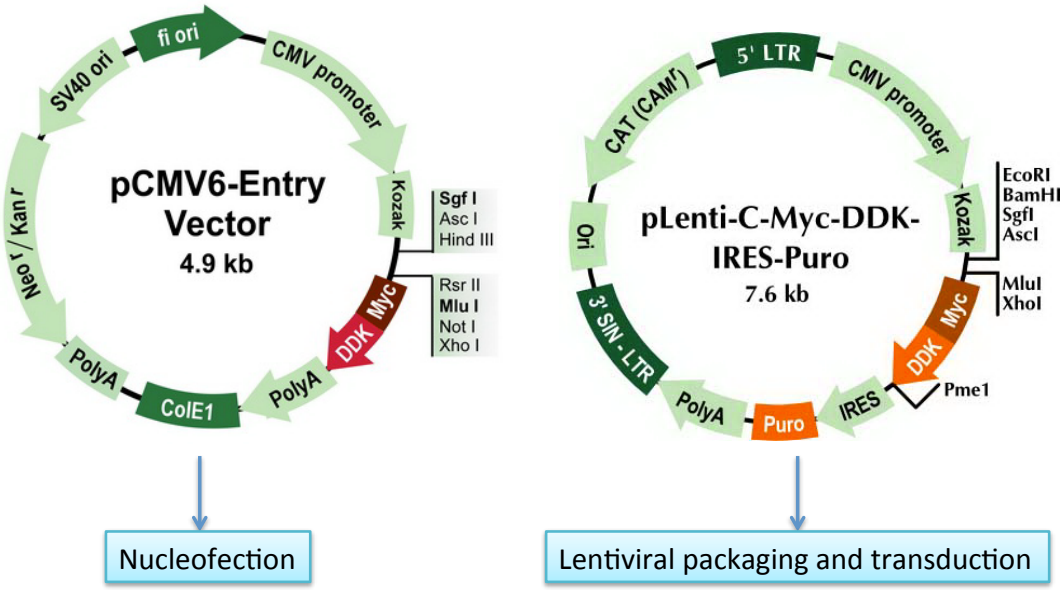

Figure S1 - DNMT3B Entry and Lentiviral Destination Vectors

Appendix 2 - Raw Data

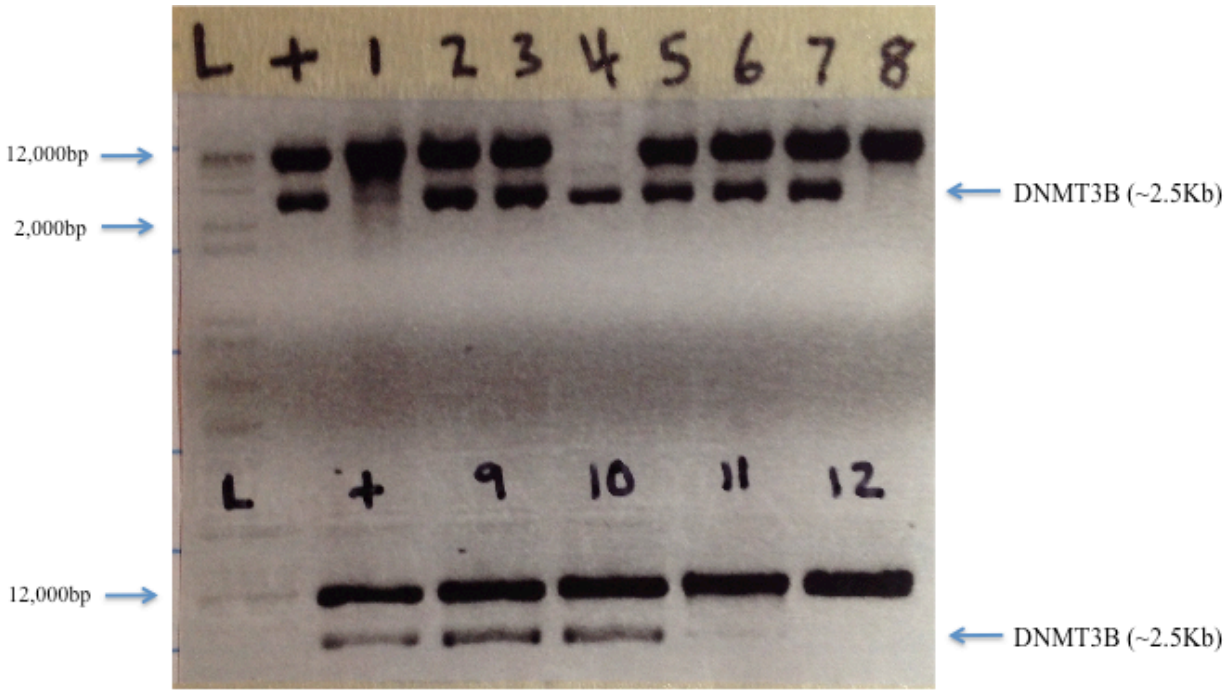

Figure S2 - MluI/SgfI Restriction Digest of pLenti-C-Myc-DDK-IRES Puro Ligated with DNMT3B V1 and V6

Plasmids were digested with MluI/SgfI for 90 minutes at 37°C prior to undergoing gel electrophoresis on an agarose (1% w/v) gel. ***L****: 1Kb Plus DNA Ladder,* ***+****: pCMV6 DNMT3B V1 clone #7 (positive control), #****1-6****: pLenti-C-Myc-DDK-IRES Puro DNMT3B V1, #****7-12****: pLenti-C-Myc-DDK-IRES Puro DNMT3B V6. The presence of DNMT3B and vector backbone sequences were detected in clones 2, 3, 5, 6, 7, 9 and 10.*