MUSASHI- 1 IN GROUP 3 MEDULLOBLASTOMA

CHARACTERIZATION OF MUSASHI-1 IN PEDIATRIC GROUP 3 MEDULLOBLASTOMA

By: Michelle Kameda-Smith Hon. BSc. Co-op, MSc., MBChB

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirement for the Degree of Doctor of Philosophy

Stem Cell and Cancer Research Institute & Department of Biochemistry
McMaster University

Descriptive Note

DOCTOR OF PHILOSOPHY (2019) McMaster University
(Biochemistry & Biomedical Sciences) Hamilton, Ontario

TITLE: Characterization of Musashi-1 in pediatric Group 3 medulloblastoma

AUTHOR: Dr. Michelle Kameda-Smith Hon. BSc. Co-op, MSc., MBChB

SUPERVISOR: Dr. Sheila K. Singh

NUMBER OF PAGES: 160 pages

Lay abstract

Brain tumours are the leading cause of childhood cancer death with medulloblastoma (MB) representing the most frequent malignant childhood brain tumour. Analysis of the data retrieved from multiple genetic studies of MB, we have determined that there are 4 genetic subgroups of MB: Wnt, Shh, Group 3 (G3) and Group 4 (G4). The subgroup with the worse prognosis is Group 3, and unique to this subgroup is the overproduction of the *MYC* gene products (i.e., *MYC* amplification). In fact, *MYC* amplification alone is associated with a poor prognosis in these children. As such many researchers and clinicians have been working together to find a way to target MYC. Although many pre-clinical experimental studies have cured MYC-amplified G3 MB using gene-targeting therapy, these results unfortunately have not translated into early clinical trials. Therefore, alternative targets that mediate the aggressiveness of MYC-amplified G3 MB is being sought. As cancer stem cells (CSC) have been implicated in tumour development and maintenance, a gene worthy of investigation in a neurodevelopmental tumour such as MB, is Musashi-1 (MSI1). MSI1 protein has been identified in high levels in many human cancers, been observed to play a crucial role in promoting normal stem cell features, and is also implicated in driving cancer. The protein that the *MSI1* gene produces binds to genes and modifies them to either stabilize or destabilize their path to becoming a protein. By manipulating *MSI1* in both NSC and MB CSC, I will observe how these cells either display greater or less cancer associated features. Further, with a new technology allowing researchers to identify MSI1 binding sites, we aim to determine how MSI1 modifies cancer causing and normal neural stem cell genes. Moreover, I will be studying both the gene-, pre-protein- and protein-level changes after experimentally manipulating *MSI1* gene levels to tease out its' main cancer associated

function. Altogether, we found a core list of genes that MSI1 modulates with functional significance giving us clues for a therapeutic targeting strategy for G3 MB.

Abstract

Pediatric medulloblastoma (MB) is the most common solid malignant brain neoplasm, with group 3 (G3) MB representing the most aggressive subgroup. Despite *MYC* amplification representing an independent poor prognostic factor in G3 MB, efforts to target the MYC pathway have met with limited therapeutic success. As such, alternative mediators of G3 MB continue to be sought. The RNA binding protein and neural stem cell determinant Musashi-1 (MSI1) has been implicated in a number of adult stem cells in various organs (e.g., brain, gut, ovaries/testes) with mounting evidence that MSI1 is an essential regulator of cancer stem cells (e.g., brain, gut, lung). Early studies in MB have shown MSI1 to be essential for tumour maintenance, however the direct interactions and specific mechanisms conferring tumours with high MSI1 expression (i.e., G3 MB) are yet to be determined. Here, I show MSI1 is an essential moderator of G3 MB in both a *MYC* amplified and p53 mutated (MP) mouse model of G3 MB and patient-derived xenograft (PDX) models. *MSI1* inhibition resulted in an abrogation of tumour initiation in both models, translating to a significantly prolonged survival. To determine how MSI1 regulates the post-transcriptional landscape of human G3 MB, an unbiased multiplatform approach was undertaken, using enhanced cross-linking and immunoprecipitation (eCLIP), and differential analyses post-*MSI1* inhibition at the transcriptome-, proteome-, and translatome-wide scale, revealing MSI1's key role in moderating G3 MB-associated cancer driving genes. In summary, employing innovational multi-platform integrative approach to stem cell cancer biology, I show the neural RNA binding protein MSI1, an essential master stem cell regulator, is hijacked from its normal neural developmental function to orchestrate the aberrant translational landscape of G3 MB.

Acknowledgements

For my family,

who have allowed me to dwell in endless possibility
and without whom none of this would have been conceivable.

Table of Contents

List of Figures and Tables

List of Abbreviations and Symbols

Adenomatous polyposis coli	APC
Alkylguanine DNA transferase	AGT
Atypical-teratoid rhabdoid tumour	ATRT
Brain tumour initiating cell	BTIC
Cancer stem cell	CSC
Carbapenem-resistant Enterobacteriaceae	CRE
Central nervous system	CNS
Chormatin immunoprecipitation	ChIP
Cyclin Dependent Kinase 6	CDK6
Cytometry by time of flight	CyTOF
Dimeythyl sulfoxide	DMSO
Deoxyribonucleic acid	DNA
Double stranded breaks	DSB
Double stranded RNA	dsRNA
enhanced cross-linking and immunoprecipitation	eCLIP
False discovery rate	FDR
Fluorescent activated cell sorting	FACS
Gamma-Aminobutyric acid	GABA
GLI family zinc finger 1	GLI1
Glioblastoma	GBM
Granule neuron precursor	GNP
Green fluorescent protein	GFP
Group 3	G3
Group 4	G4
Homeobox protein NANOG	NANOG
International Cancer Genome Consortium	ICGC
Knock down	LCA
Large cell anaplastic	KD
Magnetic resonance imaging	MRI
Mass Spectrometry	MS
Medulloblastoma	MB
MB Advanced Genomics Consortium	MAGIC
Medulloblastoma with extensive nodularity	MBEN
Messenger RNA	mRNA
Methylguanine methyltransferase	MGMT
microRNA	miRNA
microRNA induced silencing complex	miRISC
Musashi-1	MSI1
Musashi-2	MSI2
Neural stem cells	NSC
Next generation sequencing	NGS
Not significant	ns

O-6-methylguanine-DNA methyltransferase	MGMT
Orthodenticle homeobox 2	OTX2
Patched-1	Ptch1
Patient derived xenograft	PDX
Pediatric Cancer Genome Project	PCGP
Poly-A binding protein	PABP
Polycomb recessive complex	PRC
Primitive neuroectodermal tumour	PNET
RNA binding protein	RBP
RNA induced silencing complex	RISC
RNA immunoprecipitation	RIP
RNA-sequencing	RNA-seq
Sequential Window Acquisition of all Theoretical Mass Spectra	SWATH
Sex determining region Y(SRY)–box 2	SOX2
Short interfering RNA	siRNA
Short hairpin RNA	shRNA
Single stranded RNA	ssRNA
Smoothened	SMO
Sonic hedgehog	Shh
Short hairpin RNA	shRNA
Standard of care	SOC
St Jude's Medulloblastoma trial 12	SJMB12
Ribonucleic acid	RNA
Tandem Mass Tag and Mass Spectrometry	TMT-MS
Target of RNA Identification by Editing	TRIBE
The Cancer Genome Atlas	TGCA
Transforming growth factor-β	TGF-β
Tumour protein p53	TP53
Ultraviolet	UV
Unique molecular identifier	UMI
Untranslated region	UTR
Wingless	Wnt
World Health Organization	WHO

Declaration of Academic Achievement

I contributed to the design and execution of the research presented herein, as well as contributed to the data analysis and writing of all sections of the thesis. Drs. Sheila Singh, Robert Wechlser-Reya and Gene Yeo supervised the project. Dr. David Bakhshinyan and Ashley A Adile from the Singh lab "Team Medulloblastoma" helped to performed experiments and contributed to data acquisition and presentation. Collaborators EnChing Luo, Brian Yee and Frederick Tan (Gene Yeo lab), Helen Zhu (Jüri Reimand lab), and Kevin Brown performed the initial bioinformatics analyses to integrate the multiplatform work. Their specific contributions are also acknowledged in the manuscript.

CHAPTER 1

INTRODUCTION

1.0 Preamble

This thesis presents evidence and for the crucial role of an RNA-binding protein in the regulation of self-renewal in metastatic Group 3 medulloblastoma (G3 MB). The paradigm within which this work is performed is the cancer stem cell (CSC) theory with the establishment of CD133+ brain cancer stem cells identified as the rare fraction of cells maintaining neoplastic clones in high grade brain tumours (Singh et al., 2003; Singh et al., 2004). The inspiration for this work came from the initial work by *Vo et al.* and their discovery that a neural stem cell determinate, MSI1 plays an integral role in the self-renewal capacity of medulloblastoma. Specifically, that MSI1-positive primary MB tumours were associated with a poorer clinical outcome than MSI1 negative tumours (Vo et al., 2012). The aim of this introductory chapter is to take the reader through integral concepts related to the stem cell theory as it relates to pediatric medulloblastoma and further delves into the clinical implications of the molecular discoveries over that last decade.

1.1 Stem cell theory of cancer

The origin of cancer consists of multiple theoretical frameworks, which are far from reaching a consensus among researchers. One theory borne out of the classic theory cancer (Adlakha and Seth, 2016; Leedham et al., 2005; Rahman and Scott, 2007; Yeang et al., 2008) is the cancer stem cell (CSC) theory. The CSC theory

states that tumour growth is fueled by a small reservoir of tumour stem cells that reproduce themselves and give rise to neoplastic clones that sustains the cancer. This theory explains the clinical observations of tumour recurrence after seemingly successful eradication of the tumour after surgery and adjuvant therapy. The important implications of the CSC is that 1) cancer is primarily driven but a smaller population of stem cells, 2) can give rise to metastases and can act as a reservoir of cells that cause relapse after surgery and adjuvant therapy had radiologically eradicated any observable signs of disease, and 3) are closely related to normal stem cells sharing many of the behaviours and features of their normal counterparts. With growing body of knowledge that there exists a subpopulation of cells within the hematopoietic and solid tumours with stem cell properties possessing the capacity for propagating disease (Al-Hajj et al., 2003; Dick, 2003; Grichnik, 2006; Jordan, 2004; Singh et al., 2003; Wang et al., 2006), it has become one of the predominant theories to study cancer research. CSC theory has shifted the treatment paradigm from managing the tumour bulk to strategies aimed to eradicate the reservoir of CSC that repopulate the tumour bulk at recurrence. Despite its advantages to the alternative theories of cancer (i.e., epigenetic/genetic theories) (Paduch, 2015), the CSC theory is not without controversy as the theory is based largely as a result of xenotransplantation assays where human tumour cells were engrafted in immunocompromised mice. In fact, opponents of the theory argue that there is a possibility that CSCs may preferentially adapt and proliferate when propagated in a foreign species (i.e., mouse effect). In order to address these concerns, lineage tracing experiments in

the absence of transplantation assays were performed in multiple solid organ systems (Chen et al., 2012; Driessens et al., 2012; Schepers et al., 2012). These experiments collectively proved the existence of a fraction of cells with high clonogenic potential within different solid organ systems (i.e., intestinal, brain, skin) that has unique expression profile and possesses stem cell features. Furthermore, Chen *et al (*2012) and Driessens *et al* (2012) reported tumour cells similarly organize into a distinct hierarchy mirroring behaviour of normal tissue. All in all, CSC theory continues to provide an invaluable framework generating considerable insights to facilitate the understanding of the biology of human cancer.

1.2 Cancer and cells of origin

An important component of the CSC is how cancers arise. For a cell to deviate from its destined fate, it must undergo essential changes in the DNA sequence that regulates the cell. The early effort of James Till and Ernest McCulloch resolved some of the questions surrounding hematopoietic cell lineages and the existence of a multi-potent stem cell occupying the top of the hierarchy. Specifically, they discerned the number of transplanted cells required for radio-resistance in mice followed by the landmark discovery that macroscopic colonies of hematopoietic cells in the donor recipient's spleen (CFU-S) were multi-lineage, homogeneous, genetically distinct, and scaled in frequency with the number of transplanted bone marrow cells (Becker et al., 1963; McCulloch and Till, 1960; Till and McCulloch, 1961). These early observations formed the foundation for contemporary

approaches to quantify hematopoietic stem cell frequency by limiting dilution, a technique later extrapolated to solid tumour biology (Al-Hajj et al., 2003; Singh et al., 2004).

1.3 Discovery of human brain cancer stem cells

The CSC theory without a means to identify and isolate the CSC for further experimentation would not be feasible. In 2004, Singh *et al* reported *in vitro* and *in vivo* identification and purification of a cell, termed the "brain tumour initiating cell (BTIC)" from a primary human glioblastoma (GBM) and MB that was observed to have a significant capacity for self-renewal, proliferation, differentiation (Hemmati et al., 2003; Singh et al., 2003; Singh et al., 2004) and possess the ability to evade conventional adjuvant therapies (Bao et al., 2006). The BTIC represented a tumour cell population with marked expression of the cell surface marker CD133, identifiable through flow cytometry, and for the first time, experiments to characterize brain tumour CSC became feasible. The subsequent identification of CD15 as an additional marker of MB BTICs in mouse models further corroborated the existence of BTICS and facilitated the cancer stem cell modeling of MB in mice (Read et al., 2009; Ward et al., 2009). Clinical validation of the BTIC model was carried out by Panosyan *et al* using stem cell properties (i.e., self-renewal) of the BTICs, in pediatric glial and embryonal brain tumours, where correlates with outcome in patients with MB were observed (Panosyan et al., 2010). Specific to MB, a higher MB stem cell self-renewal index was associated with a more aggressive the clinical course of

MB prompting researchers to study cancer through the CSC model to determine how these CSC with BTIC properties promote tumourigenesis, maintenance, dissemination and treatment evasion (Figure 1.1). Additionally the observation that key brain tumour pathway overlap significantly with stem cell pathways may suggest the origin of these cells may be a normal neural stem cell gone awry. The inertia that of modern stem cell theorists have developed in collaboration with the neurooncological community further solidified the idea that stem cells or their close derivatives may underlie tumour formation.

1.4 Medulloblastoma cell of origin

With the appreciation that a variety of neural tumours harbor an essential subpopulation of cells that maintains the growth of neoplastic tissues and can initiate new tumours when transplanted in a receptive host, proposals of their cell of origin were postulated. For MB, attention was initially directed to the most abundant cells in the early post-natal period of the brain, the cerebellar granule neuron. These cells arise from granule neuron precursors (GNPs) in the external granular layer during the early post-natal period where their remarkable expansion is principally governed by the growth factor, Shh (Kenney et al., 2004). Building on this foundation of normal neurodevelopment of the cerebellum, the implication of an aberrant predominant signaling pathway was hypothesized after observing the sporadic development of MB in individuals with familial syndromes with aberrant Shh and canonical Wnt pathways (e.g. Gorlin's and Turcot's syndrome respectively). In 2008, two groups reported significant progress towards

understanding the cell of origin for both familial and sporadic cases of MB induced by the Shh pathway (Schuller et al., 2008; Yang et al., 2008). Yang *et al* showed activation of the Shh signaling in stems cells that are committed to the neuronal lineage causes tumours reminiscent of MB. The authors used CRE-recombinase-mediated deletion of patched (*Ptch1*), a repressor of Shh signaling, driven by a *Math1* enhancer element (i.e., an active in cerebellar GNPs), resulting in MB formation in both Ptch1 inhibition *in utero* or post-natally. Of note, the tumours were initiated in neural stem cells within the granule lineage developed more rapidly that those initiated in progenitors, explaining the clinical observation of a bimodal distribution of this subgroup of MB (Kool et al., 2012; Yang et al., 2008). Another group arrived at a similar conclusion by constitutive activating of the *Smo* allele (*SmoM2*), mimicking Shh signaling (Schuller et al., 2008). Schuller *et al* beautifully showed mice with targeted *SmoM2* expression in distinct cell populations within the developing brain (e.g., multi-potent $GFAP^+$ and $Olig2^+$ cell, cells committed to the cerebellar GNP lineage by *Math1*, *Tlx3* or *Gli1* expression and Purkinje cell neurons), uniformly formed tumours with histological features of MB. The initial work to identify the cell of origin of Shh MB by Yang *et al* (2008) and Shuller *et al* (2008) has paved the way for further identification of the cell of origin in the remaining three MB molecular subgroups. The second pathway to gain attention after the observation of MB in Turcot's syndrome (i.e., germline mutation of the gene, adenomatous polyposis coli (*APC*), causing over-activity of the Wnt pathway). During embryonic neurodevelopment, activation of the canonical Wnt signaling pathway has been observed to play an

important role in regulating neural stem cell proliferation. After Gibson *et al* (2010) observed genes identifying the Wnt subgroup more frequently expressed genes that were frequently expressed in the lower rhombic lip and the embryonic dorsal brainstem, a different cell of origin of Wnt MB was postulated (Gibson et al., 2010; Rogers et al., 2012) (Figure 1.2).

Employing an alternative approach to cell of origin determination, multiple studies have highlighted the important relationship between the molecular subgroups and the spatio-temporal origins of NSCs in which the initiating tumourigenic events occur. Specifically, expression of *mycn* in embryonal day 16 (E16) cerebellar mouse NSCs can induce Shh MB while expression of the same mutant at P0 produces MB that is reminiscent of a human G3 MB (Swartling et al., 2010) suggestive that a more stem like cell is induced by the same mutant to form G3 MB. This theory was further corroborated by Hooper *et al,* who made additional observations of MB molecular subgroups aligning along a neuronal differentiation continuum. The authors noted Wnt and Shh MB clustered close to NSCs and G3 and G4 MB closer to normal fetal germinal matrix and normal fetal brain (Hooper et al., 2014) in line with finding that cells cultured from higher grade brain tumours were characterized by an enrichment of stem cell features (Singh et al., 2003). Altogether, these findings suggest subtype specific mutational events, as well as spatiotemporal associations during embryonic development, in medulloblastomagenesis.

1.5 Pediatric medulloblastoma

Clinically, pediatric brain tumours are the leading cause of cancer mortality with medulloblastoma (MB) representing the most frequent malignant tumour (Ellison et al., 2011). MB is a highly malignant embryonal tumour that was first described as a distinct central nervous system (CNS) tumour entity in 1925 (Bailey, 1925). Current clinical risk stratification for pediatric MB delineates children into average risk and high-risk groups. High-risk disease is identified by the presence of an age less than 3 years, presence of metastasis at diagnosis and post-operative residual disease of greater than 1.5cm^2 (Gajjar et al., 2006; Ramaswamy et al., 2011). The multimodal approach to the treatment of pediatric MB has made significant strides over the last 30 years (Ellison, 2002). Initially, radiation therapy was introduced to reduce the rate of local recurrence and along the craniospinal axis (Lampe and Macintyre, 1954; Smith et al., 1973). The subsequent addition of maintenance chemotherapeutic agents (i.e., lomustine and vincristine) did not observe statistically significant survival benefits (Michiels et al., 2015) but a subset analysis did identify a particular group of MB (i.e., non-metastatic MB) in whom a survival benefit was apparent (Evans et al., 1990; Tait et al., 1990; Taylor et al., 2003).

For children under 3 years of age, the balance between the devastating neurocognitive side effects of radiation therapy and disease control has prompted the neurooncological community to learn from early experiences of post-operative irradiation of brain tumours (Packer et al., 1989). Traditional treatment for children >3 years of age encompasses maximal safe resection surgery, adjuvant whole cranio-spinal irradiation and chemotherapy with vincristine, cisplatin,

cyclophosphamide and lomustine (Packer, 2007). In children under 3 years of age however, a radiation sparing, high dose chemotherapy protocol is administered with radiotherapy reserved for salvage therapy. Epidemiologically, the incidence of these tumours has remained relatively constant in the US and Canada with approximately 5 per 1,000,000 children diagnosed annually (Johnston et al., 2014; Partap et al., 2009). Though survivorship has witnessed a significant improvement for these children, nearing 80% for standard-risk patients, the treatment toxicity and long-term sequelae of the current regime continues to considerably impact their quality of life (Bull et al., 2014; Ellison, 2010; Mulhern et al., 2004; Palmer et al., 2003; Palmer et al., 2001; Ris et al., 2001; Yoo et al., 2016).

In 2000, the World Health Organization (WHO) identified 4 histological variants: desmoplastic, large cell, medullomyoblastoma and melanocytic (Kleihues, 2000). Subsequently in 2007, the better known histopathological classification, incorporating prognostic and some cytogenetic data, further refined the previous classification to desmoplastic/nodular, MB with extensive nodularity (MBEN), large cell anaplastic (LCA) (Louis et al., 2007). Nodular desmoplastic tumours may have a better prognosis than classic variants. MBEN predominantly occur in infants and have an excellent prognosis (Rutkowski et al., 2005). The diagnosis of LCA MB on a basis of histology can often be difficult as the phenotype may be restricted to a small component of the tumour but for the most part have been associated with poor prognosis. Though improvements in outcome with chemoradiotherapy has witnessed a 2-fold reduction in mortality of patients with

medulloblastoma by 2002, it became clear that further improvements would require a better understanding of the biology of MB (Ellison, 2002).

1.6 Genomic stratification into subgroups with clinical significance

At the turn of the millennium, high-throughput methods for studying the transcriptome facilitated the early studies in histologically similar leukemias, identifying their further subdivision into clinically and molecularly distinct groups (Golub et al., 1999). A subsequent study by Pomeroy *et al* utilized transcriptomics to study histologically similar tumours (i.e., primitive neuroectodermal tumours (PNETs) of the CNS and MB) observing these entities were molecularly distinct from atypical teratoid-rabdoid tumours (ATRT) and unexpectedly from each other (Pomeroy et al., 2002). Furthermore, they showed MBs with classic histopathology were biologically distinct from nodular desmoplastic histology providing the initial evidence for what is now regarded as the Shh MB subgroup. Subsequent high throughput characterization of primary MB tumour samples identified definitive clusters that appeared to be distinct from one another (Cho et al., 2011; Kool et al., 2008; Northcott et al., 2011; Thompson et al., 2006). In 2010, an international expert panel gathered at a consensus conference in Boston, and determined that the evidence supported the existence of four main medulloblastoma subgroups (Wnt, Sonic hedgehog (Shh), Group 3 and Group 4) based on multiple genomic platforms with distinct demographics and clinical features (Taylor et al., 2012). The Wnt and Shh pathways were characterized by upregulation of genes in the canonical Wnt or Shh pathways respectively. These 2

subgroups were separated from each other and other subgroups on principal components analysis and were associated with improved clinical outcomes when compared to the Group 3 and 4 subgroups which were less well characterized but clinically correlated with a greater likelihood of metastatic disease and poor clinical outcomes (Cho et al., 2011; Clifford et al., 2006; Ellison et al., 2005; Kool et al., 2012; Kool et al., 2008; Northcott et al., 2011; Taylor et al., 2012). Following this international consensus, the WHO provided an update on the Classification of Tumours of the Central Nervous System to include the molecular parameters in addition to histology to define many tumour entities. For MB, the genetically defined subcategories included the Wnt activated, Shh activated and *TP53* mutant, Shh activated and *TP53* wild-type, and Non-Wnt/non-Shh Group 3 (G3) and Group 4 (G4) (Louis et al., 2016).

As advances in technology set the pace for our understanding of human cancers, developments associated with next generation sequencing (NGS) further enhanced our understanding of the genes, pathways and molecular processes associated with commonly diagnosed human cancers. MB has been subject to several NGS studies by large groups such as the MB Advanced Genomics Consortium (MAGIC), International Cancer Genome Consortium (ICGC) and the Pediatric Cancer Genome Project (PCGP). The advances in technology were soon followed by development in the field of bioinformatics to analyze large volumes of data from multiple quantitative platforms. In 2017, Cavalli *et al* applied similarity network fusion to genome-wide DNA methylation and gene expression data of >700 primary samples to further support the existence of subtypes within

the 4 consensus subgroups (Cavalli et al., 2017); While, Northcott *et al* described 24 mutational signatures that are operative in MB, in addition to subgroup-specific drivers and pathways with the data available from NGS (Northcott et al., 2017) (Figure 1.3). Pushing our understanding of MB further, in 2018, Forget *et al* and Archer *et al* integrated proteomic and phosphoproteomic data to lend further credence to the existence of subtypes within the consensus subgroups and to suggest the significant role of post-transcriptional regulation in the non-Wnt/Shh MB (Archer et al., 2018; Forget et al., 2018; Zomerman et al., 2018).

In efforts to bridge the molecular data to the clinical observations in pediatric MB, a thorough review of recent published and unpublished data, by Ramaswamy and colleagues updated the consensus highlighting subgroups of MB associated with exceptionally high-risk features. Risk groups were defined based on current survival rates: low risk (>90% survival), average (standard) risk (75-90% survival), high risk (50-75% survival) and very high risk (<50% survival). The Wnt subgroup and non-metastatic G4 MB (G4 MB) with chromosome 11 loss or whole chromosome 17 gain were recognized as low risk and may qualify for reduced therapy. High-risk patients encompassed patients with metastatic Shh, G4 tumours, and *MYCN*-amplified Shh MB. And, G3 with metastasis or Shh with *TP53* mutations were identified as very high risk (Ramaswamy et al., 2016). Ramaswamy urges researchers to consider investigating alternative treatments for the high- and very high-risk patients for whom there is currently no alternative targeted treatments but the standard regime.

1.7 MYC and group 3 medulloblastoma

G3 MB tumours express genes involved in GABAergic/photoreceptor signaling and has generally been observed to take an aggressive clinical course with a particularly poor prognosis (i.e., 5-year survival rates of 39% for infants <3 years and 58% for children >3years) despite aggressive therapy (Kool et al., 2012). G3 MB is more common in males, observed in infancy, childhood, rarely in teenagers. Why pediatric MB is restricted to the pediatric population is unknown. Metastasis is found in a high percentage of patients at presentation. Upregulation of *MYC*, a known oncogene are exclusively observed in G3 MB and are mutually exclusive with amplicons containing *OTX2*, a transcription factor that has been shown to cooperate with *MYC* (Northcott et al., 2012a). Ramaswamy *et al* had identified that at least 2 subgroups within G3 MB exists where lack of *MYC* amplification and non-metastatic confers a standard risk (75-90% survival) and metastatic presentation upgrades your risk stratification to very high (<50% survival). While the amplification of *MYC* has been identified as a independent risk factor, phosphoproteomics analyses revealed post-translational modification of *MYC* or a non-*MYC* amplified/*MYC* signature was predictive of patient outcome shifting the cancer genomic paradigm to incorporate a proteomic component, paving the way to emphasize the importance of the study of the final effectors, the gene protein product (Archer et al., 2018; Zomerman et al., 2018). Due to the pharmacological difficulty of targeting a transcription factor, small molecule inhibitors of MYC, have not yet achieved wide success or acceptance in the clinic (Northcott et al., 2012a). The identification of a pathway responsible for

driving G3 MB for which treatments are currently available have been welcomed by clinical trial groups for the last 5 years. A recent clinical trial, SJYC07, published their results of risk-adapted therapy in young children with medulloblastoma unfortunately reported no improvement in event-free survival with poor outcomes associated with G3 B despite aggressive upfront adjuvant therapy (Robinson et al., 2018). Currently as part of the St. Jude's MB12 (SJMB12) trial, the addition of the folate antimetabolite pemetrexed, and the nucleoside analog gemcitabine is currently being evaluated in select G3 and G4 patients. Neither of these agents however target specific pathway mediators in G3 MB nor G4 MB.

Though a number of molecularly targeted clinical trials are underway, at present, the vast majority of patient receive the same age dependent combination therapy (Polkinghorn and Tarbell, 2007). This approach however fails to cure two-thirds of patients with G3 MB and most probably over-treats children with Wnt MB who invariably survive with long-term cognitive and endocrine side effects (Robinson et al., 2012). The introduction of de-escalation therapy in the clinical trial setting has addressed the question of whether we are over-treating our Wnt MB, however an appreciable knowledge gap remains to inform effective molecularly guided treatment strategy for G3 MB.

1.8 Summary of intent

The heavy focus on transcription factors and transcriptional control within the field of brain tumour stem cell biology research has provided a narrow perspective on a vastly complex field of cancer stem cell behaviour.

The primary theme of this thesis can be described as testing the hypothesis that the normal function of post-transcriptional modifiers of normal neural development, such as the RNA binding protein Musashi-1, may be altered to confer onto G3 MB the poor clinical prognosis and has to potential to provide a novel paradigm to consider candidates for targeted therapy.

If this hypothesis holds true, and MSI1 aberrantly controls the fate of transcripts associated with CSCs, its inhibition would set off a detrimental cascade of downstream signaling pathways to abrogate the CSC's capacity for tumour propagation. The experimental aims to test the primary hypothesis are as follows:

1) To confirm the expression of MSI1 in G3 MB

2) Validate the MSI1 KD in multiple human models of G3 MB to test its effects on the functional stem cell properties *in vitro* and *in vivo.*

3) Determine the mechanism by which MSI1 globally contributes to the maintenance of normal stem cell and CSC properties in normal neural stem cell and *MYC* amplified G3 MB cell lines by performing eCLIP and differential pathways analysis.

4) Determine the mechanism by which MSI1 contributes to the maintenance of cancer stem cells in *MYC* amplified G3 MB cell lines by employing a systems biology approach and performing and executing a comparative

integrative computational analysis of RNA-sequencing (RNA-seq), TMT-MS (tandem mass tag mass spectrometry), and polysome-sequencing (polysome-seq) of control vs. sh*MSI1* inhibited G3 MB cells.

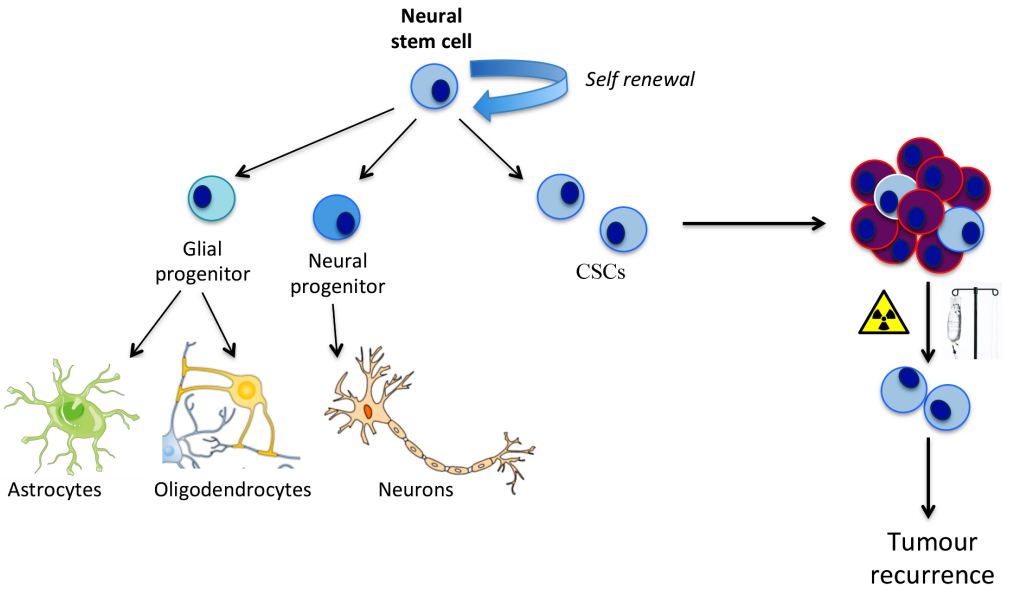


Figure 1.1: Cancer stem cell theory applied to the central nervous system: At the top of the hierarchy is the neural stem cell giving rise to an exact phenocopy of itself and a more differentiated daughter cell, a neural or glial progenitor on the left. The glial progenitor gives rise astrocytes and oligodendrocytes and the neural progenitor to neurons. As applied to the cancer stem cell theory, on the right, the neural stem cell undergoes a mutagenic event becoming a cancer stem cell (CSC) giving rise to exact phenocopies of itself along with more differentiated cells comprising the bulk tumour. The bulk tumour is treated by the upfront cytotoxic radiation and chemotherapy, however the CSC is able to evade non-targeted therapies to promote tumour recurrence.

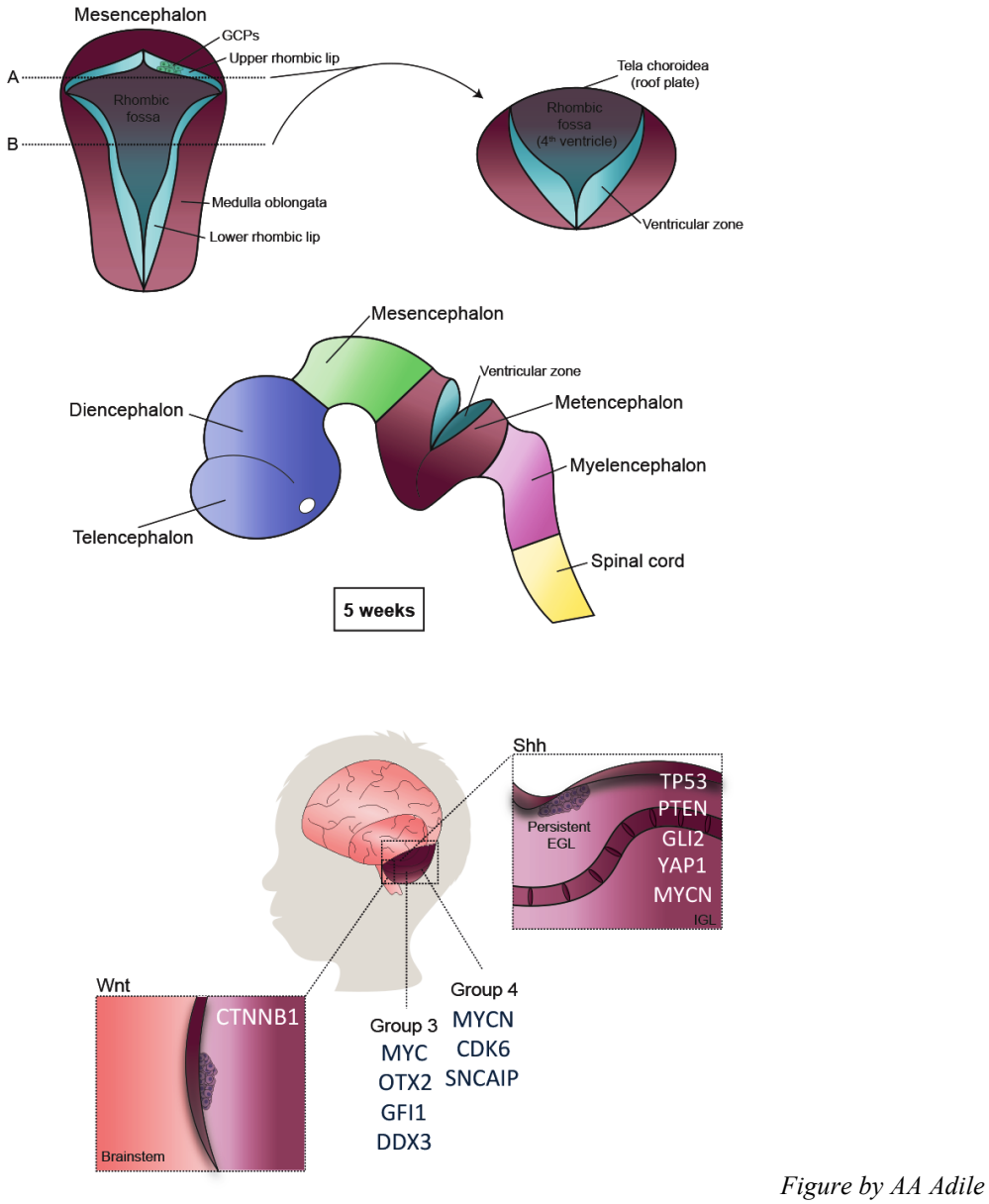


Figure by AA Adile

Figure 1.2: Medulloblastoma cell of origin: As a developmental tumour, medulloblastoma (MB) serves as a perfect model for the study of cancer stem cells. The germinal matrix is a reservoir or neural stem cells as is the upper (A) and lower (B) rhombic lips which will later form the embryonic 4th ventricle with its surrounding ventricular zone (i.e., where granule cell precursors reside). These stem cells have been studied to suggest distinct developmental cells of origin. Wnt from the dorsal brainstem, Shh from the persistent granule cell precursors (GCP) within a persistent external granular layer, and some studies have suggested early exposure to mutations of driver genes associated with G3 and G4 MB.

Subgroup	**Wnt**		**Shh**				**G3**			**G4**		
Subtype	Alpha	Beta	Alpha	Beta	Gamma	Delta	Alpha	Beta	Gamma	Alpha	Beta	Gamma
Incidence	10%		30%				25%			35%		
Gender	1M:1F		1M:1F				2M:1F			3M:1F		
Age	3-17	>10	3-17	0-3	0-3	>17	0-10	3-7	0-10	3-17		
Mets	9%	21%	20%	33%	9%	9%	43%	20%	40%	40%		
5 year survival	97%	100 %	70%	70%	90%	90%	65%	55%	40%	65%	75%	80%
Copy number variants	6		MYCN + GLI2+ YAP1 +	PTEN -		10q22-, 11q23.3-	7+, 8-, 10-, 11-, i17q-	OTX2+, DDX3-	MYC+	MYCN+, CDK6+	SNCAIP duplication	CDK6 +
Other events			*TP53* mutations			TERT promotor mutations		High GFI1/1B expression				
Histology	Classic, rarely LCA		Desmoplastic, nodular, Classic, LCA				Classic, LCA			Classic, LCA		

Figure 1.3: Pediatric medulloblastoma, molecular subgroups, genomic alterations and clinical outcomes (Adapted from Pollack et al, 2019).

CHAPTER 2

THE ROLE OF TECHNOLOGY AND METHODS TO STUDY THE MULTIPLE FACETS OF BIOLOGY

2.0 Preamble

This chapter is an original peer-reviewed review article, and is a prelude to the large-scale multi-platform systems biology approach to the investigation of MSI1 in G3 MB, The article is presented in its final format.

"This article was published in Future Neurology. Kameda-Smith MM, Manoranjan B, Bakhshinyan D, Adile AA, Venugopal C, Singh SK. Brain tumour initiating cells: With great technology will come greater understanding. 10 November 2017, doi:https://doi.org/10.2217/fnl-2017-0011. Copyright ©Future Medicine Publishing Group"

MMK-S researched and synthesized the current state of methodology and techniques to examine the brain tumour-initiating cell. BM and DB contributed to the parts of the transcriptional regulation section. AAA assisted with original figure production and all authors revised the manuscript in its final accepted form.

BRAIN TUMOUR INITIATING CELLS: WITH GREAT TECHNOLOGY WILL COME GREATER UNDERSTANDING

2.1 Abstract

The discovery of the brain tumour-initiating cells (BTICs) resulted in a paradigm shift within the cancer research community to consider brain tumours as an outcome of developmental mechanisms gone awry. This review will guide the

reader through the technological advances that hold the powerful potential to allow brain cancer researchers to develop an intimate understanding of the dynamic and complex mechanism governing brain tumour behavior.

2.2. Main

The paradigm shift resulting from the discovery of brain tumour-initiating cells (BTICs) invited the field of cancer research to consider brain tumours as an outcome of developmental mechanisms gone awry (Reya et al., 2001). We previously described the low frequency of BTICs across primary brain tumours, diversity and dynamic state of BTIC progeny, and intratumoural heterogeneity as challenges encountered when modeling and investigating the functional genomics of BTICs (Singh, 2013). More recent advances in biochemical techniques have identified key factors linking the BTIC's molecular composition and associated clinical phenotype, thereby further narrowing the gap between the laboratory bench top and clinical bedside.

The central dogma of molecular biology, first described by Francis Crick in 1958, was possible after the discovery of "nuclein" by Friedrich Miescher. Nuclein, later known as deoxyribonucleic acid was hypothesized as the source material for the building blocks of life. In view of the advancements observed in just over a decade within molecular biology by 1970, Crick correctly predicted the combined efforts of biology, chemistry, and physics will synergistically advance experimental techniques and improve our understanding of molecular biology (Crick, 1970). He felt very strongly that an important factor, "sometimes not

sufficiently appreciated has been the tremendous power of modern experimental techniques, mostly springing from physics and physical chemistry" (Crick, 1970). Crick keenly reflects, "if it be accepted that most of the problems with which we are today rather closely concerned are likely to be solved by the year 2000, it is worthwhile to consider what problems are likely to remain unsolved" (Crick, 1970). He foretold, "questions concerning the behavior of the cell and further the organism as a whole are likely to remain unsolved." Though molecular biology has been successful concentrating its efforts to experimentation of a small part of a biological system, Crick felt that "in the long run, problems involving the complex interactions cannot be avoided" (Crick, 1970). These very issues plague the present day experimental cancer biologist.

Advancements in cancer research have aided in identification of BTICs in brain tumours as cells responsible for the tumourigenesis, maintenance, dissemination, and therapeutic evasion of a multitude of brain tumours that affect both children and adults (Clarke et al., 2006; Lowry and Temple, 2009; Manoranjan et al., 2012; Singh et al., 2004; Wu et al., 2012). Genes that have been upregulated and downregulated in BTICs have been studied (Manoranjan et al., 2013) and drugs that could be used to modulate these pathways have been investigated (Gajjar and Finlay, 2015). But as Crick anticipated, the intricate interplay between cells within a tumour is far from understood.

The two hallmarks of cancer stem cells are multi-lineage differentiation and self-renewal, giving the BTIC the capacity to originate, propagate and even disseminate the brain tumour depending on the tumour microenvironment,

subtype or grade (Singh, 2013). A stem cell or BTIC's decision to either maintain its stem cell identify or differentiate into a specific cell type is ultimately determined by the outcome of a complex interplay between extracellular signaling pathways, transcriptional regulatory networks, and post-translational modifications. Advances in experimental biochemistry have allowed researchers to characterize differences between normal neural stem cells and BTICs to identify potential mechanisms of tumourigenesis, maintenance, dissemination and treatment evasion. In this review, we describe the technological advances within the major fields of stem cell biology, biochemistry and biophysics that may hold the key to understanding how BTICs arise, evade treatment and are conferred a growth advantage over the resident neural cells.

Transcriptional regulation

The expression of genes is a highly regulated process at multiple interdependent levels. In eukaryotes, gene expression is initiated at the level of transcription where transcription factors bind to DNA to regulate RNA synthesis of specific genes in response to varying external and internal stimuli. In mammals, transcription factors represent approximately 10% of all protein-coding genes, making up the largest single class of proteins encoded in the genome (Levine and Tjian, 2003).

At the transcriptional level, stemness is largely governed by the pluripotency factors OCT 4 (Seymour et al., 2015b), SOX2 (Seymour et al., 2015b; Suva et al., 2014; Vanner et al., 2014), NANOG (Seymour et al., 2015b), and FoxG1

(Manoranjan et al., 2013), which function to maintain the stem cell phenotype. These transcription factors form a core stem cell transcriptional network by binding to both promoter-proximal DNA elements and to more distal regions that can be 100s of kilobases away (Guo et al., 2011b; Jackson et al., 2015; Seymour et al., 2015a). However, there is a paucity of evidence to link OCT4 and NANOG to governing BTIC stemness. SOX2 appears to be a key player, with OCT4 and NANOG's importance in brain tumourigenesis remaining yet to be explored (Favaro et al., 2014; Gangemi et al., 2009). Interestingly, recent work by Asadi *et al* examined splice variants of Oct4 in brain tumour tissues by RT-PCR and IHC and observed increased expression of the Oct4B1 splice variant in high grade brain tumours (Asadi et al., 2016). Further, Goncalves da Silva *et al* demonstrated that high OCT4A levels drive tumourigenicity and metastatic potential of OCT4A overexpressing Daoy and D283 MB cell lines (Goncalves da Silva et al., 2017). Targeted therapies through silencers of potency factors in brain tumour research have demonstrated promising results in animal models. SOX2 peptide vaccination in immunodeficient mice transplanted with high-grade oligodendroglioma cells delayed tumour development, increased survival rates, and with the combination of temozolamide, further doubled survival time compared to vehicle-treated controls (Favaro et al., 2014).

Development of DNA microarray technology and subsequently RNA-sequencing (RNA-seq) has made it possible to efficiently measure the expression of each annotated gene at the transcript level. RNA-seq is a whole transcriptome shotgun sequencing approach utilizing next-generation sequencing to identify the quantity

of RNA present at a specific time point of several organ systems in humans. (Figure 2.1). New technologies developing in breast cancer and ovarian research includes (1) engineered zinc-finger based artificial transcription factors to selectively silence SOX2 gene expression in cell lines causing SOX2 mRNA downregulation (Stolzenburg et al., 2012), (2) RNA interference technology against NANOG resulted in reduced expression of cyclin D1 and MYC with associated reduction in key stem cell properties (Han et al., 2012), and (3) RNA interference promotes apoptosis and reduces cancer cell viability (Peng et al., 2010). To address concerns regarding intratumoural heterogeneity, single-cell transcriptome sequencing has also been applied, yielding interesting leads for potential biomarkers (Ramskold et al., 2012).

With their enormous tumourigenic potential, the relative frequency of BTICs is thought to be quite limited. This is in part due to the few stemness genes that have been identified from genomic profiling of the bulk tumour in two of the most malignant primary brain tumours, adult glioblastoma (Verhaak et al., 2010) and pediatric medulloblastoma (Northcott et al., 2012b). More recent studies have performed single-cell sequencing to further characterize the clonal heterogeneity within these tumours (Patel et al., 2014; Tirosh et al., 2016). The identification of a core set of neurodevelopmental transcription factors essential for the initiation and maintenance of brain tumours (Liau et al., 2017; Patel et al., 2014; Tirosh et al., 2016) has provided the transcriptional evidence required to support a continuum of stemness-related expression states, with functional relevance. Specifically, BTICs demonstrate phenotypic plasticity in their ability to transition

between states of active proliferation and slow-cycling through their regulation of these primitive transcriptional programs (Liau et al., 2017). Therefore, slow-cycling BTICs with enhanced developmental transcriptional machinery are able to evade current therapeutic strategies and promote subsequent disease recurrence and relapse (Vanner et al., 2014). These observations have been characterized with the use of elegant transgenic mouse models in which a neural stem cell gene expression profile was identified as the source of tumour initiation and maintenance in the most aggressive subgroup of childhood medulloblastoma (Kawauchi et al., 2012; Pei et al., 2012). Of greater clinical significance has been the demonstration of *in vivo* therapy resistance and subsequent tumour progression by only those glioblastoma cells with an activated nestin promoter, a marker of neural stem cells (Chen et al., 2012).

With RNA microarray experimentation, the relative mRNA expression of a corresponding gene is assumed to reflect the amount of protein in the cell. In other words, there is little regulation at the post-transcriptional level. This notion has been challenged with evidence that there is a relatively weak correlation between number of transcripts and protein products of a gene suggesting post-transcriptional regulation is predominant (Gygi et al., 1999; Washburn et al., 2003). To further understand the mechanisms controlling self-renewal and differentiation, it is becoming clear that we need to look beyond transcriptional networks towards translational and post-translational events that modulate gene expression.

Translational regulation

While in eukaryotes, transcription factors regulate the synthesis of RNA of a specific gene in response to intracellular or environmental stimuli, the pathway to protein synthesis undergoes numerous checks and balances along the way. One of these check points is the post-transcriptional regulation of mRNA which tightly controls the fate of the RNAs. Critical regulators of this process are RNA-binding proteins (RBPs) and microRNA (miRNA).

In eukaryotes, transcription and translation occur in different compartments, allowing for a multitude of factors to control RNA at the post-transcriptional level. Though earlier studies revealed RBPs to function in transporting mRNA from the nucleus to the site of their translation, it is becoming ever clearer that these proteins are involved in the regulation of nearly all levels of post-transcriptional control (i.e., splicing, polyadenylation, transport and localization, mRNA stability, and translational control) (Keene, 2007). RBPs bind single or double stranded RNA and determine its fate from synthesis to decay (Castello et al., 2012; Jarvelin et al., 2016; Kechavarzi and Janga, 2014; Mittal et al., 2009). Given their essential role in the control of gene expression at the post-transcriptional level, alterations in their expression or mutation in either RBPs or their RNA targets have been shown to promote brain tumourigenesis (Chen et al., 2016b; de Araujo et al., 2016).

RBPs as a functional class are rapidly turned over at the transcript level and are tightly controlled at the protein level (Mittal et al., 2009). In general, mRNA of RBPs encoding central physiological pathways have longer half-lives, while those

encoding regulatory and signaling proteins have shorter half-lives. Data on mRNA half-life and abundance indicate the mRNA expression of RBPs to be transient but whenever they are transcribed, they are produced at high concentrations. The implications of this finding is yet to be explored in BTICs (Mittal et al., 2009). DNA mutations can bring about changes in post-translational modification, which affect protein stability and activity of key oncogenic and tumour suppressor pathways. Though our understanding of the transcriptional regulatory landscape continues to develop, the post-transcriptional regulatory control by RBPs and miRNA remains largely unknown.

Until recently, techniques to identify RBP targets have been limited as RNA immunoprecipitation (RIP) results were plagued with post-lysis RNA-protein interaction compromising the data when trying to identify true *in vivo* association. Cross-linking and immunoprecipitation (CLIP) addressed the concern of post-lysis RNA-protein interaction by irreversibly binding the RBP to the RNA with ultraviolet light and allowing analysis of the true RBP-RNA interaction. This technique has developed through numerous improvements such as photoactivatable-ribonucleoside-enhanced CLIP (PAR-CLIP) (Hafner et al., 2010), to individual-nucleotide-resolution CLIP (iCLIP) (Konig et al., 2010) to the current enhanced CLIP (eCLIP) (Van Nostrand et al., 2016) (Figure 2.2). The challenge encountered by researchers utilizing this technology is the inefficiency of cross-linking and the requirement of a highly efficient antibody for the RBP of interest. At present however, CLIP remains the gold standard of analyzing RBP-RNA interaction and has the powerful potential to identify novel regulators of

normal stem cell populations. Examples in the hematopoietic (Rentas et al., 2016) and gastrointestinal (Li et al., 2015) systems found RNA binding protein targets identified through CLIP identified pathways in which the RBP regulates cord blood hematopoietic stem cell (HSC) expansion and colonic cancer stem cell activation respectively. Identifying post-transcriptional modifiers of oncogenic pathways in BTICs has the potential to identify novel downstream therapeutic targets.

To address these challenges however, other techniques developed in parallel take advantage of the cell's endogenous enzymes to either tag or introduce an irreversible editing event to a region of the mRNA in the locus of RBP interaction. Two novel techniques have been described recently: RNA tagging (Lapointe et al., 2015) and Target of RNA Identification By Editing (TRIBE) (McMahon et al., 2016). In RNA tagging, an RNA binding protein of interest can be fused to an enzyme (PolyU polymerase) that adds uridines to the end of RNA, covalently marking the RNA target with a polyU tag, which is subsequently identified using high-throughput whole genome sequencing. TRIBE however, addresses the second challenge of CLIP reducing the number of cells required to the double digits. TRIBE was developed with the same concept as the DNA-oriented targeted DamID (van Steensel and Henikoff, 2000). *In vivo* expression of a fusion protein between the RBP of interest and only the catalytic domain of the RNA-editing enzyme ADAR is produced. The transcriptome is then sequenced to identify novel editing events by the fusion protein. Both RNA tagging and TRIBE have been demonstrated to correlate highly with CLIP.

In addition to RBP, microRNAs (miRNAs) are essential regulators of post-transcriptional gene regulation and have important roles in many fundamental biological processes (Huntzinger and Izaurralde, 2011). In animals, these small RNA molecules are transcribed from intronic regions by RNA polymerase II and through the Drosha-Dicer pathway where mature miRNA are incorporated into miRNA induced silencing complexes (miRISCs) to exert their post-transcriptional control of gene regulation (Chen and Kang, 2015). While the function of mature miRNA has been vigorously debated over the last decade, remarkable progress toward the understanding of the role of miRNA has resulted in the consensus that these small molecules predominantly mediate mRNA degradation, thereby gene silencing (Huntzinger and Izaurralde, 2011). Evidence for mRNA degradation is mainly from transcriptomics studies observing abundance of miRNA targets inversely correlate with levels of miRNA (Baek et al., 2008; Hendrickson et al., 2009; Huntzinger and Izaurralde, 2011; Selbach et al., 2008). Over the last decade it has become clear that miRNA expression are dysregulated in human malignancies (Hayes et al., 2014). Expression levels of miRNA can be altered in the brain through a variety of mechanisms including chromosomal changes, epigenetic defects and mutations in the machinery of their biogenesis (Barciszewska, 2016). D'Urso *et al* and Aldaz *et al* demonstrate that miRNA are integrally involved in the development and progression of brain gliomas (Aldaz et al., 2013; D'Urso et al., 2012) and Srinivasan *et al* utilized The Cancer Genome Atlas (TCGA) dataset and miRNA array to identify a 10-miRNA expression signature associated with outcome survival in glioblastoma (Srinivasan et al.,

2011). As miRNA have proven greater stability than mRNA in a range of specimens (e.g., blood, urine, formalin-fixed tissue blocks), in addition to demonstrating greater sensitivity than proteins, it has lead to considerable interest in development of miRNA as biomarkers for a range of biological diseases including cancer (Pritchard et al., 2012). For the researcher investigating the mechanisms of gene regulation, measuring miRNA expression can inform systems-level of gene regulation, particularly when combined with mRNA profiling and other whole genome scale data. Currently, the majority of large-scale miRNA target identification experiments involve differential expression of a single miRNA followed by downstream gene expression or proteomic analysis (Pritchard et al., 2012).

RNA-binding proteins and miRNA play an integral role in determining the fate of RNA, but there remains the question of how much protein associated with the BTIC is produced. If activation of oncogenes or inhibition of tumour suppressor genes produce proteins unique to BTICs, this may provide them with an evolutionary advantage over normal neural cells, allowing BTICs to evade therapies designed to eradicate malignant cells. Identifying the RBP and miRNA's binding sites is the first step in the determination of the mechanism by which these translational regulators promote a BTIC state. Next lies the challenge of determining which oncogenic pathways are regulated by them and how to reduce the BTIC drive to dominance.

Proteomics

Despite extreme genetic heterogeneity, tumours often demonstrate similar alterations in the expression, stability, and activation of proteins important in oncogenic signaling pathways (Oh, 2016). Until recently, systematic approaches to characterize the molecular makeup of brain tumours remain biased to DNA- and RNA-based platforms as evidenced by the enormous volume of transcriptomics data available with less known about its functional implications. The main challenges that proteomic research face include extreme diversity of physical and chemical properties of proteins where (1) no single method is capable of isolating all proteins, (2) lack of high-fidelity amplification of proteins for experimentation as compared to polymerase chain reaction (PCR) for DNA or RNA, and (3) difficulty of *de novo* peptide sequencing when compared to nucleic acid sequencing(Oh, 2016). By combining mass spectrometric-based (Figure 2.3) and antibody-based principles, researchers are now able to quantify the relative and absolute abundances of protein within a sample (Shi et al., 2012; Ullal et al., 2014).

To address intratumoural heterogeneity, single cell proteomic technologies provide the basic scientist with an important tool to study the cellular differences found in tumours with significantly higher resolution. A microfluidics-based single-cell proteomic chip has been applied to the study of glioblastoma to identify predominant signal transduction pathways under varying conditions. By assessing the relevant pathway proteins under a number of conditions, protein correlation networks can be constructed and demonstrate previously unappreciated network interconnectivity shedding new light on potential

therapeutic resistance (Shi et al., 2012). The challenges to this technique however include the lack of a sophisticated commercial platform, variations in primary tumour sample handling in the clinical setting, and lack of 3-D spatial information of the single tumour cells obtained (Tian et al., 2016). Overcoming these challenges are a possibility, even more so in combination with imaging analysis (i.e., in situ hybridization of tissue sections).

Cytometry by Time-of-Flight (CyTOF) is another recent technology for single cell analysis to investigate protein diversity across tumour cells. This technology applies both the high-throughput cell handling capability of a conventional flow cytometer and analytical capacity of atomic mass spectrometry. Limitations for the wide utilization of the CyTOF technology include (1) the number of available metal isotopes, the high cost of instrumentation and reagents and (3) the lack of high quality antibodies to probe for proteins of interest (Tian et al., 2016). CyTOF however is yet to be utilized in brain cancer research but has the potential to provide the researcher with the granularity of data similar to a recently reported single cell transcriptomics study of glioblastoma (Patel et al., 2014).

The holy grail of global proteomics is what every basic scientist has been waiting for: Sequential Window Acquisition of all Theoretical Mass Spectra (SWATH). This method promises to be the first truly global approach to proteomics with the ability to qualify and analyze 5000+ proteins from complex mixtures. While still under development, this technology will be combining the ability to extract signals from highly complex data with the rapid acquisition speed of the latest generation of mass spectrometers. SWATH promises to do for proteomics what

microarray and now RNA-seq has done for transcriptomics: provide a way to globally measure abundance of almost all proteins across samples (Tian et al., 2016).

The ultimate goal of these proteomic technological applications to brain cancer research is to identify novel biomarkers that can be clinically validated for diagnostics and eventually therapeutics. The emerging field of proteomic technology has the potential to augment our study of BTICs to identify biomarkers for early detection, diagnosis, stratification, progression monitoring, and candidate discovery for targeted therapy. Unraveling the mechanisms of intrinsic and acquired resistance to anti-cancer therapies at a protein level has the potential to improve DNA/RNA based classification of tumours, and may prove to be crucial to devising reversal strategies (Oh, 2016).

Epigenetics

Epigenetic alterations are stable, long-term changes in cellular phenotype that are not due to variations in DNA sequence (Holliday, 1987). The additional focus on mechanisms of regulation of gene expression at the transcriptional and translational levels is another emerging field in cancer research. Post-translational epigenetic modifications of histones come in the form of acetylation, methylation, phosphorylation, citrullination, ubiquitination sumoylation, ADP ribosylation, deimination, or proline isomerization. Accessibility of DNA regulatory elements is another level of controlling gene expression. These modifications, govern the interaction of transcription factor networks with other regulatory factors. These

events dictate a specific gene as either actively transcribed, poised, or silenced at any given time (Young, 2011).

The global finding of hypomethylation of DNA in human tumours was followed by the identification of hypermethylation of tumour-suppressor genes and inactivation of microRNA genes by DNA methylation (Esteller, 2008). DNA methylation occurs in the cytosines that precede guanines (AKA: dinucleotide = CpG). These CpG rich regions are known as CpG islands and span the 5' end of a gene's regulatory region. DNA methylation occurs in the context of chemical modification of histone proteins. These histone proteins package DNA, in addition to participating in the regulation of gene expression themselves. Histones mainly store epigenetic information such as arginine and lysine methylation, serine phosphorylation and lysine acetylation that have the potential to affect gene transcription (e.g., methylation of histone 3 at lysine 4 is closely linked to transcriptional activation) and DNA repair (Esteller, 2008). Altered epigenetics has been observed to play an important role in the development of brain tumours (Maury and Hashizume, 2017; Sturm et al., 2014).

Polycomb repressive complexes 1 and 2 (PRC1 and PRC2) represent two major multi-protein complexes involved in epigenetic silencing of genes driving self-renewal and proliferation across multiple cancer subtypes (Otte and Kwaks, 2003). The initial step of the canonical PRC-mediated gene repression is the tri-methylation of lysine 27 on histone 3 (H3K27me3) by PRC2 (Cao et al., 2002; Czermin et al., 2002; Muller et al., 2002). The H3K27me3 mark on the promoter of the target gene is then recognized by PRC1 complex, which further maintains

the repressive state through ubiquitination of lysine 119 on histone H2A (H2AK119ub) (Fischle et al., 2003; Min et al., 2003; Wang et al., 2004).

One of the key loci that are repressed by Polycomb-group (PcG) proteins is the *CDKN2A* locus encoding for $p16^{INK4A}$ and $p14^{ARF}$ tumour-suppressor proteins which contribute to cell cycle mediation (Sharpless and Sherr, 2015). The epigenetic repression of *CDKN2A* locus has been shown to increase self-renewal and proliferation of CSCs across multiple cancer subtypes (Bignell et al., 2010; Chandler and Peters, 2013). However, when upregulated, Bmi1 has been implicated in promoting self-renewal, proliferation and therapy evasion in multiple cancer types including neuroblastoma, glioblastoma (GBM) and medulloblastoma (MB) (Glinsky, 2007; Li et al., 2010; Vrzalikova et al., 2008; Wang et al., 2008b). Clinically, elevated levels of Bmi1 have been shown to correlate with poor overall patient survival and increased incidence of tumour recurrence (Guo et al., 2011a; Honig et al., 2010; Hosen et al., 2007; Leung et al., 2004; Li et al., 2009; Merkerova et al., 2007; Tabor et al., 2011; Vrzalikova et al., 2008). In GBM, Bmi1-assisted activation of DNA double strand breaks (DSB) has been shown to allow CD133 positive, brain tumour initiating cells, to escape radiotherapy (Facchino et al., 2010). Similar to Bmi1, GBM patients with increased levels of EZH2, a member of PRC2 complex, have been shown to correlate with poor patient prognosis (Wu et al., 2013). Studies have demonstrated upregulation of EZH2 in both mouse and human GBM BTICs. And, pharmacologic inhibition of EZH2 (Orzan et al., 2011) and target knockdown (KD) studies with short-hairpin RNAs against EZH2 (Abdouh et al., 2009; Suva

et al., 2009) resulted in decreased self-renewal of GBM BTICs *in vitro* and impaired tumour-initiating capacity *in vivo* (Suva et al., 2009).

Chromatin modification is an intricately orchestrated process that involves a variety of proteins and histones and capturing this assembly in an unbiased fashion at specific loci has been far from a trivial task. Techniques such as chromatin immunoprecipitation assays have the capability to only look at one factor (histone post-translational modification or protein) at a time (Figure 2.4). To achieve simultaneous profiling of multiple targets, chromatin affinity purification coupled with mass spectrometry (ChAP-MS) was developed (Byrum et al., 2012). The method of choice for studying the BTIC epigenetic landscape is highly dependent on a number of factors. If the candidate genes to be analyzed are known, digestion based assays followed by PCR or bisulfite conversion are appropriate. However, if the genes are unknown, one can search for differentially methylated regions using bisulfate conversion followed by microarray or bisulfite sequencing or whole genome methylation profiling (Kurdyukov and Bullock, 2016). Bisulfite treatment of DNA, first described by Frommer in 1992, takes advantage of the observation that it mediates a deamination of unmethylated cytosine to uracil (Frommer et al., 1992). These converted residues will be read as thymine, as determined by PCR-amplification and subsequent Sanger sequencing analysis. The 5-methylcytosine residue however are resistant to this conversion and therefore will continue to be read as cytosine. Therefore, comparison of the Sanger sequenced bisulphite treated sample to an untreated DNA sample enables the detection of the methylated cytosines adequately. This technique is important

in BTIC analysis because DNA methylation, which commonly occurs at the C5 position of cytosines within the CpG dinucleotide, plays an important role in essential biological functions (e.g., cellular proliferation, differentiation, chromosome stability). Cytosine methylation has long been recognized as an important factor in silencing genes of mammalian cells by blocking transcription factors binding to DNA thereby inhibiting transcription. Further, aberrant DNA methylation is often associated with loss of DNA homeostasis and genomic instability leading to cancer.

These comprehensive techniques have unraveled descriptions of the epigenomes of human cancer cells. More importantly, they have directed therapeutic utilization of the knowledge of the epigenetic landscape of cancer cells to treat brain tumours with alkylating agents such as temozolomide, lomustine and cyclophosphamide to name a few. Specific to studies with temozolomide, its' therapeutic benefit depends on the compound's ability to methylate DNA to trigger tumour cell death. However, some tumour cells are able to repair this type of damage and reduce the therapeutic efficacy of temozolamide by expressing the aklylguanine DNA alkyltransferase (AGT) protein encoded in humans by the *O*-6-methylguanine-DNA methyltransferase (MGMT) gene. If the MGMT gene is silenced by promoter methylation, DNA repair will be compromised conferring a survival benefit to patients with glioblastoma who are treated with radiotherapy and alkylating agents (Hegi et al., 2005). There have been numerous studies analyzing the clinical implications of MGMT promoter methylation status in glioblastoma and have found that overall survival of patients with methylated

MGMT status conferred a greater overall survival after chemoradiotherapy (Esteller et al., 2000; Glas et al., 2009; Hegi et al., 2004; Hegi et al., 2005; Herrlinger et al., 2006; Stupp et al., 2009). Accordingly, MGMT promoter methylation was identified as a prognostic marker. Studies that combine temozolamide with other alkylating agents such as lomustine have demonstrated further survival benefit in newly diagnosed glioblastoma (Glas et al., 2009). In relation to BTICs, a recent study demonstrated that inconsistently methylated MGMT promoter status is not uncommon and conferred a worse prognosis when compared to methylated cases (Xia et al., 2016). An explanation for this finding has been demonstrated by Sciuscio *et al* where they observed some glioblastomas with low percentage of overall MGMT promoter methylation can be cultured to produce glioblastoma-derived spheres, or a BTIC population, with high degrees of MGMT promoter methylation suggesting there may be an element of intratumoural heterogeneity at the epigenetic level, yet again emphasizing the importance of single cell studies (Sciuscio et al., 2011).

Living systems depend on the dynamics of gene expression to regulate cell survival, adaptation to stress, homeostasis, cell fate, and differentiation, in response to environmental signals. A multilayered transcriptional control mechanism serves as a system of checks and balances that allows for the fine-tuning and adaptability of the CSC genomic landscape. These advantages conferred to the BTIC are the ultimate challenge with which researchers grapple. The BTIC is not only a moving target, but as a product of being a dynamic system, multiple levels of gene regulation requires examination at multiple time

points with great precision. Heisenberg's uncertainty principle applied to brain tumourigenesis requires our understanding of more than one variable (i.e., transcriptional regulation, post-transcriptional regulation, epigenetics) at the same time, in order to truly conceptualize BTIC tumourigenesis, maintenance, dissemination and treatment evasion.

2.3 CONCLUSION

As we learn more through careful experimentation with BTICs, we start to understand that there are multiple regulatory pathways and multiple levels at which these pathways are continually modulated. As BTIC researchers it is important not only to embrace the constantly evolving technology in the different fields of molecular biology, chemistry, and physics, but to apply them to their model systems to obtain a whole system dynamic understanding of the environment in which the BTIC is maintained and thrives.

2.4 FUTURE PERSPECTIVE

If we can synergistically compile and collate data emerging from technical advances in physics, physical chemistry and biochemistry as Dr. Crick had predicted in 1970 for the year 2000, we are most certainly on the cusp of a deeper understanding of how BTICs confer their survival advantage. With the advances in techniques to study translational and proteomic characterization, we will certainly see an explosion of data to further our mechanistic understanding of BTIC cancer biology. The next 5-10 years will be an exciting time for the basic

scientist working within brain tumour neurobiology. More and more clinical trials will be guided from **the bench top and the gap between the two will soon be obsolete.**

2.5 EXECUTIVE SUMMARY

- BTICs shifted the experimental paradigm of brain tumours as the outcome of developmental mechanisms gone awry
- A stem cell or BTIC's decision to either maintain its stem cell identify or differentiate into a specific cell type is ultimately determined by the outcome of a complex interplay between extracellular signaling pathways, transcriptional regulatory networks, and post-translational modifications.
- The identification of a core set of neurodevelopmental transcription factors essential for the initiation and maintenance of brain tumour has provided the transcriptional evidence required to support a continuum of stemness-related expression states, with functional relevance.
- At the post-transcriptional level, RBP and miRNA play significant roles in modulating BTIC oncogenic pathways. The rapidly developing area of RNA binding site identification has provided a mechanism for BTIC researchers to gain a multi-dimensional understanding of how these cells disseminate, evade conventional therapy and maintain a stem cell state
- Despite extreme genetic heterogeneity, tumours often demonstrate similar alterations in the expression, stability, and activation of proteins important in oncogenic signaling pathways. The emerging field of proteomic

technology has the potential to identify biomarkers for early detection, diagnosis, stratification, progression monitoring, and candidate discovery for targeted therapy.

- Epigenetic modifications regulate gene expression at the transcriptional and translational levels. Discoveries from epigenetic profiling experiments have directed therapeutic utilization of the knowledge of the epigenetic landscape of cancer cells to treat BTICs with alkylating agents such as temozolomide, lomustine and cyclophosphamide.
- A multilayered transcriptional control mechanism serves as a system of checks and balances that allows for the fine-tuning and adaptability of the CSC genomic landscape
- As BTIC researchers it is important not only to embrace the constantly evolving technology in the different fields of molecular biology, chemistry, and physics, but to apply them to their model systems to obtain a whole system dynamic understanding of the environment in which the BTIC is maintained and thrives.

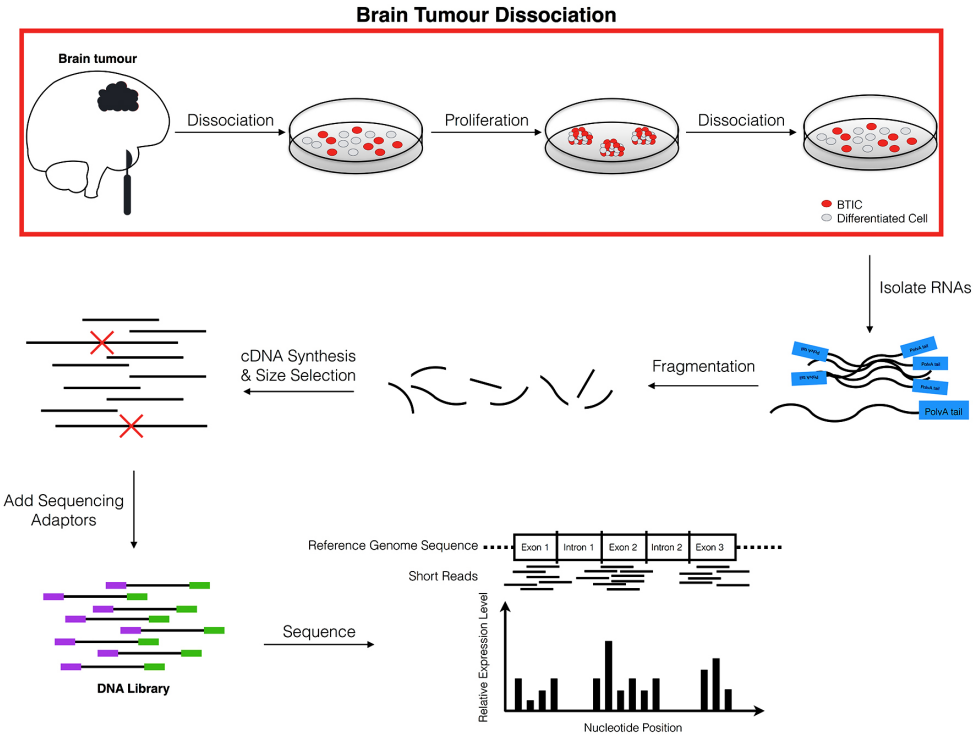


Figure 2.1: Depiction of RNA Sequencing of BTICs. Following tumour resection, brain cancer cells that are composed of both BTICs (red) and differentiated cells (gray) are allowed to proliferate and are subsequently dissociated for RNA isolation to occur. RNA is then fragmented and synthesized into cDNA, in which larger reads is selected out. Sequencing adaptors, or linkers are added to the ends of the short reads of the DNA library, which are then sequenced relative to the reference genome sequence to analyze the cellular transcriptome.

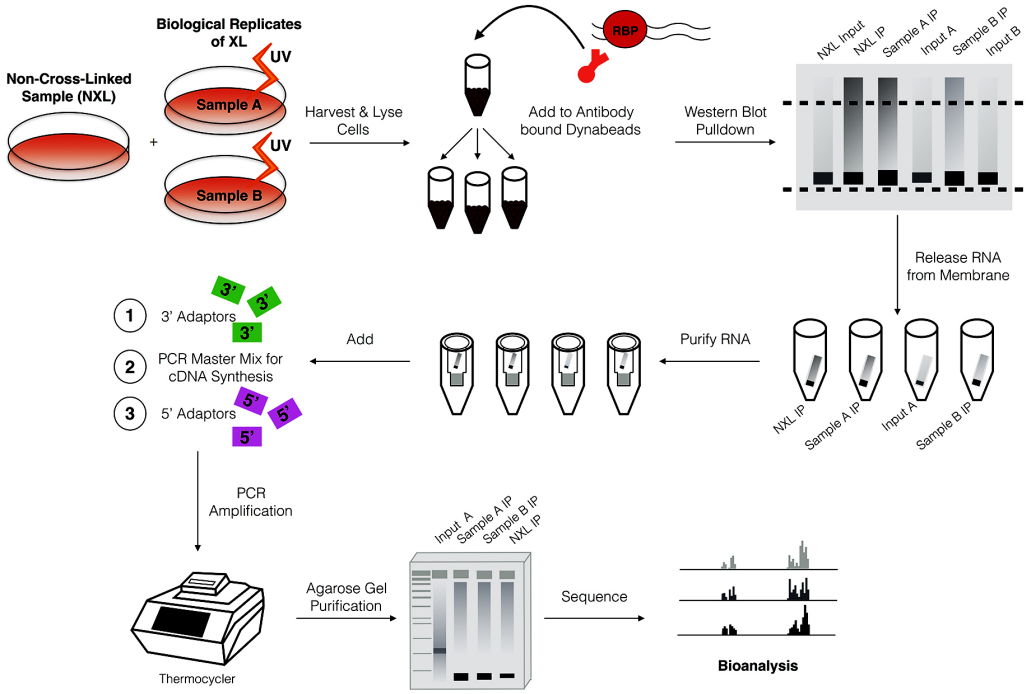


Figure 2.2: Illustration of how BTICs can be analysed at the post-transcriptional level by investigating the binding sites of RNA binding proteins by utilizing the enhanced CLIP (eCLIP) protocol. Cells are cross-linked using UV irradiation and then harvested and lysed. Dynabeads are bound to antibody of the RNA binding protein of interest and added to the lysates. The cells then processed to identify the RNA binding sites by sequencing and bioanalysis.

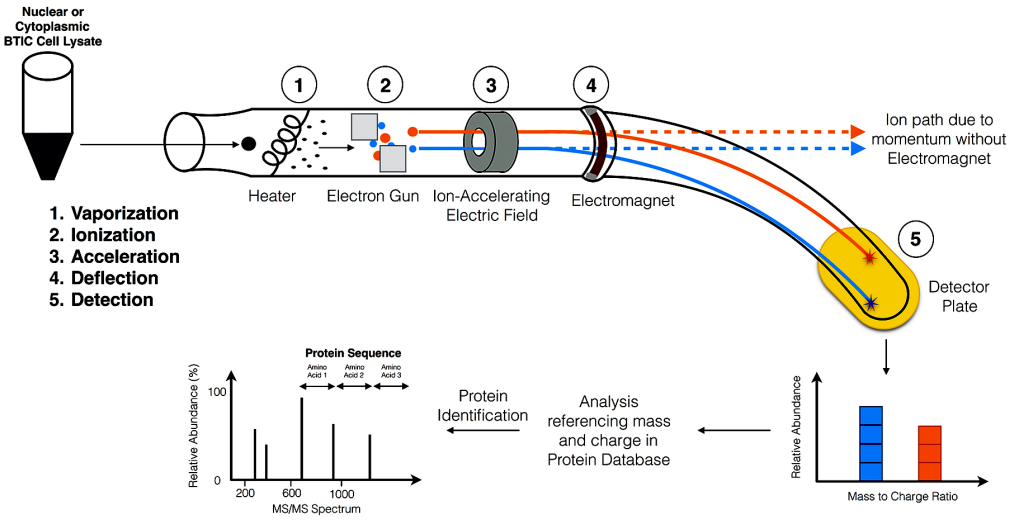


Figure 2.3: Schematic of BTIC protein identification via mass spectroscopy. Nuclear and cytoplasmic BTIC protein is extracted following gel electrophoresis, after-which fragmentation takes place. Cell lysates are separately input into the mass spectrometer to undergo five stages for protein identification - (1) the sample is subjected to vaporization, in which it become gaseous, (2) the electron gun generates a charged field in which the sample's atoms become ionized, (3) the ion-accelerating electric field enables the particles to travel through a narrow, linear path, while (4) the presence of the electromagnet results in a non-linear path, such that larger particles resist the electromagnet field relative to smaller particles, which (5) is picked up by the detector. This allows for protein identification based on the mass to change ratio and Tandem mass spectrometry (MS/MS) spectrum.

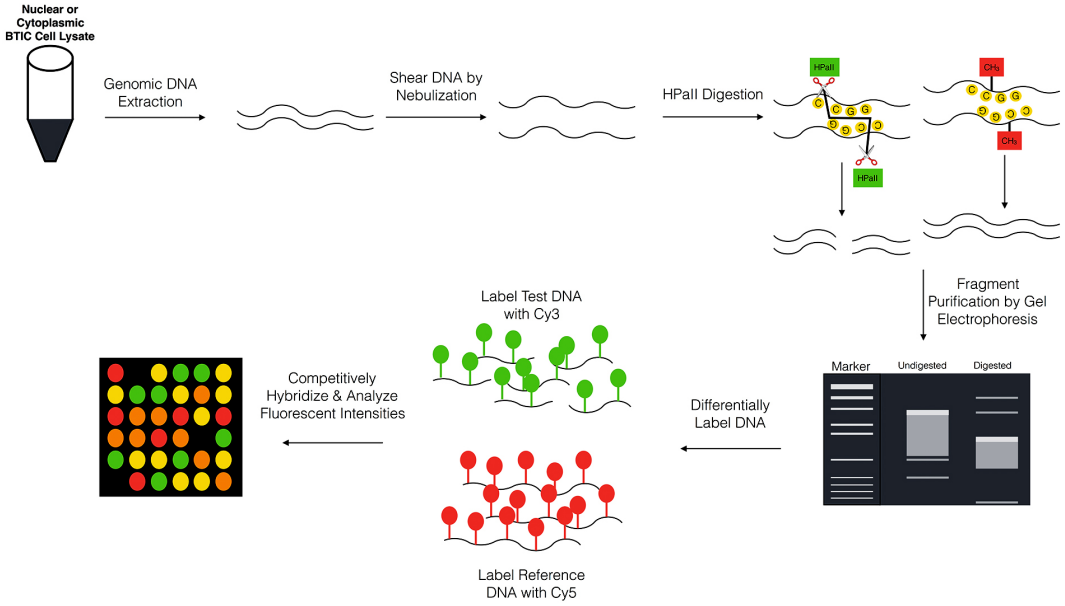


Figure 2.4: Illustration of BTIC methylation profiling. Genomic DNA is extracted from cell lysates, sheared with a nebulization agent and either digested with restriction enzymes (green), such methylation-sensitive Hpall, or undigested through the addition of methyl groups (red). Gel purification of fragments between 80-2500bp follows, after-which differential labeling of the test DNA with Cy3 (green), in comparison to Cy5 for the reference DNA (red), occurs. The labels are hybridized and further analyzed based on their fluorescent intensities.

CHAPTER 3

TECHNOLOGICAL ADVANCES APPLIED TO THE STUDY OF THE RNA BINDING PROTEIN MSI1 IN NEURAL AND CANCER STEM CELLS

3.0 Preamble

This chapter will explain the basic principles and utility of the experiments carried out in our systems biology, multi-platform approach to the characterization of the mRNA binding protein Musashi 1 (MSI1) in G3 MB. Herein, the gene or transcript of Musashi will be italicized with the capitalized form indicating the human ortholog (i.e., human: *MSI1;* mouse: *Msi1*). The same will hold true for the protein without italics (i.e., human: MSI1; mouse Msi1). The impedes for writing this chapter was to provide details to understand what is known about MSI1 and early insights of its functional role within the cell as well as provide a the rationale for the large data sets generated in order to apply a systems biology approach to the to the characterization of an RNA binding protein.

3.1 MSI1 in normal neural stem cells

Musashi (MSI) is a family of RNA binding proteins characterized by 2 RNA recognition motifs (RRMs) separated by a short linker region (Okano et al., 2005). The MSI RRMs, which are located in the N-terminal half of the protein, are highly evolutionarily conserved; In fact, the mouse and human RRMs are essentially identical (Good et al., 1998). MSI proteins are usually found in the

cytoplasm consistent with their function in translation regulation (Sakakibara et al., 2001). However, they have also been detected in the nuclear fraction (UniProtKB, 2017) suggestive of a role in alternative pri-mRNA modification (i.e., alternative splicing, polyadenylation, capping) and transport. In mammals, two isoforms of this family, MSI1 and MSI2 exists. MSI2 exhibits high sequence homology with MSI1, which is more than 90% at the amino acid level within the RNA binding domain (Nishimoto and Okano, 2010). The cytogenic location of *MSI1* is on chromosome 12q24.31 and *MSI2* is on 17q22 (Figure 3.1).

When MSI was first identified, it was discovered as a required factor for the asymmetric cell division of the sensory organ precursor cell of Drosophila adult external sensory organ (Nakamura et al., 1994). Because the two-bristled phenotype was reminiscent of the great samurai, Miyamoto Musashi's swords, the gene was named *Musashi*. Since its discovery, *MSI* has proven to be critical in controlling stem and progenitor cell function across organisms. Where the *Msi* transcripts were present in stem cells that have the potential to yield neurons and glia, its expression could not be detected in fully differentiated neuronal and glial cells (Sakakibara et al., 1996).

Differences in the spatial expression of the 2 isoforms has been observed with *Msi2* transcribed over all the tissues, and *Msi1* enrichment particularly noted in neural stem cells or progenitor cells of the peri-VZ area in embryonic (Sakakibara et al., 1996) and postnatal (Sakakibara and Okano, 1997) mammalian brains. Furthermore, Msi1 and Msi2 were observed to have independent roles in the nervous system. *Msi1* loss alone led to the development of hydrocephaly in the

majority of Msi1 deficient mice (Sakakibara et al., 2002). By contrast, *Msi2* was found to be uniquely expressed in the parvalbumin-containing GABA neurons in the neocortex and neurons in the basal ganglia, which may be suggestive of an *Msi2*-specific function in these cell lineages (Sakakibara et al., 2001).

It is worth noting that the central role of MSI in gene regulation, which is to modulate translation through its RNA binding domain, was first identified in the nervous system and has most extensively been studied in neural stem cells. Potential mechanisms include repression and activation of translation (Figure 2). One of the most cited studies providing evidence for the mechanism by which MSI1 maintains the stem cell state was the observation by Imai *et al* elegantly showing *NUMB* to be a the target of MSI1. They proposed MSI1 directly binds the 3' untranslated region (UTR) of the *NUMB* mRNA and downregulates its expression thereby triggering a rise in Notch signaling (Imai et al., 2001). Activation of Notch signaling induces the transactivation of the promoter of the *Hes*-1 gene and is also known to regulate the self-renewal of NSCs positively (Hitoshi et al., 2002). Multiple other genes of potential interest were identified but of specific interest, similar to Uren *et al* (Uren et al., 2015) and Sakakibara *et al* (Sakakibara et al., 2002), we found that while the extensively studied quintessential target of MSI1, *NUMB*, was expressed in our system as evidenced in the RNA-seq, it was not bound by the endogenous MSI1 on SU_MB002 nor NSC201cb as previously described (Imai et al., 2001). It is becoming clear that the changes in *NUMB* transcript and its associated NOTCH pathway to help maintain cellular self-renewal is indeed more complicated than a direct inhibition of the *NUMB* as previously reported (Uren et al., 2015). This in part can be explained by the original experiments by Imai *et al* performed in cells that do not

endogenously express MSI1, emphasizing the importance of the context specific functional study of a gene in its natural milieu.

3.2 MSI1 in cancer stem cells

Much of the experimental efforts to better understand the biological mechanism of solid cancers has been directed towards the identification of gene transcription mechanisms determining stem and progenitor cell differentiation. However, little is known about the complimentary mRNA translation mechanisms that contribute to the control of stem and progenitor cell differentiation. Nucleic acid binding proteins constitute 23% of the functionally annotated human genes, reflecting the vital role these proteins play in the control of gene expression (Venter et al., 2001). Post-transcriptional regulation of gene expression is mediated by group of proteins that binds to pri-mRNA and mRNA. Their activity is essential for correct splicing, localization, and translation of cellular components and their dysregulation is implicated in numerous human diseases. The RNA binding protein (RBP), *Musashi (Msi)* has been observed to play a crucial role in promoting stem cell self–renewal (MacNicol et al., 2015) and implicated in the central nervous system tumours including glioma and MB (Chen et al., 2016a; Cox et al., 2013; Dahlrot, 2014; Dahlrot et al., 2013a; Dahlrot et al., 2013b; de Araujo et al., 2016; Johannessen et al., 2009; Kanemura et al., 2001; Lagadec et al., 2014; Muto et al., 2012; Sanchez-Diaz et al., 2008; Toda et al., 2001; Vo et al., 2011; Vo et al., 2012). Proteins, which bind multiple RNA targets, can act as a master regulator of gene expression at the post-transcriptional level to co-ordinate cellular processes and alter the CSC. RBPs, such as *MSI,* are highly expressed in

cancer cells (Dahlrot et al., 2013a; Johannessen et al., 2009; Kanemura et al., 2001; Nakano et al., 2007) and have been found to correlate with poor clinical course (Chen et al., 2015; Dahlrot, 2014; Dahlrot et al., 2013a; Dahlrot et al., 2013b; Ma et al., 2015). Identifying the gene targets and RBP interactomes of MSI has the powerful potential to highlight essential pathways of tumourigenesis of MB that cannot necessarily be detected by the extensively studied of transcriptomics alone.

A complete functional understanding of any RBP requires the identification of its RNA targets. Though *NUMB* has been identified as a target for MSI1 in derepressing the Notch pathway, *PTEN* has also been suggested as a target of MSI1 in derepressing the PI3K-AKT pathway. Both of these pathways, when activated, promote a stem-like phenotype of the cancer cell. However, identifying biologically relevant RBP targets however has presented with numerous challenges to date. Newer techniques to determine the binding sites of RBP, such as MSI, has evolved from simple immunoprecipitation, to cross-linking and immunoprecipitation followed by sequencing (CLIP-seq) (Ule et al., 2005; Ule et al., 2003) and more recently a method utilizing RNA-editing moiety to identify targets of RNA binding proteins, (AKA: Targets of RNA-binding proteins identified by editing: TRIBE) (McMahon et al., 2016).

3.3 Systems biology approach

Systems biology is the computational modeling of complex biological systems focusing on complex interactions within biological systems using a holistic rather

than a reductionist approach. Cancer systems biology aims to understand cancer as an integrated system of genes, networks and interactions rather than an entity of isolated molecular and cellular components (Archer et al., 2016). Uncovering the many molecular features of cancer inevitably reveals new forms of complexity requiring novel computational models to help explain the complex and dynamic relationships observed in cancer. Computational tools developed to make sense of the large amount of data associated with the biology of cancer aids the understanding of this complex biological systems. Herein, we describe the available techniques to perturb the expression of an RBP followed by how RBP specific holistic experimental techniques have evolved to facilitate our understanding of their role within a complex biological system.

3.4 Techniques to perturb genes coding RNA binding proteins and their limitations

Comparison of the wild type and experimentally induced genetic variants have proven a valuable asset for determining the physiological and developmental contribution of a gene. Multiple techniques to carry out these gene mutations have been described. However, the essentiality of the gene within a system is a central consideration to facilitate the ability to make meaningful downstream observations. For essential genes, an approach to inhibit genes by RNA interference (i.e., shRNA, siRNA, CRE-inducible) may be appropriate as a complete knock out of the gene will be lethal and provides minimal substrate for downstream analysis. RNA interference (RNAi) refers to the double stranded

(dsRNA) induced modality of gene silencing. In this process, RNA homologous to the one strand of complimentary dsRNA is degraded (Fire et al., 1991; Fire et al., 1998; Hannon, 2002; Izant and Weintraub, 1984). In the mammalian system, the introduction of long dsRNA (>50bp) yielded systemic, non-specific inhibition of translation due to activation of interfereon and protein kinase R response (Kumar and Carmichael, 1998). This problem was circumvented by the use of synthetic short interfering RNA (siRNA) that can be either delivered exogenously or expressed endogenously from RNA polymerase pol III promoters (Lipardi et al., 2001; Sijen et al., 2001). Subsequently, to facilitate the stable expression of RNA inhibition, the expression of short hairpin RNAs that are subsequently processed into siRNA (Singer and Verma, 2008). And finally, although a powerful gene editing tool, Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated Cas9 gene editing technology (Cong et al., 2013) has advanced the field of genetic engineering and our overall understanding of the role of individual genes, MSI1 was considered to be an essential gene in our system and therefore gene editing to knock out the gene was not evaluated and detailed discussion of this technology outwith the scope of this chapter.

3.4.1 Small interfering RNA (siRNA) inhibition

The simplest form of RNAi is the cytosolic delivery of siRNA oligonucleotides. These double stranded RNA (dsRNA) molecules that are approximately 20-25 nucleotides in length can produce siRNA, short hairpin RNA (shRNA) or microRNA (miRNA) by RNase III ribonuclease enzyme, Drosha and exported

into the cytoplasm (Filippov et al., 2000). Once the dsRNA is in the cytoplasm, it is cleaved by a ribonuclease called Dicer into smaller segments called siRNA. RNA induced silencing complex (RISC) proteins including Slicer and Argonaut associate with the siRNA and complementarily bind to the mRNA and cleaves the mRNA with RNAse H, causing degradation (Bernstein et al., 2001; Ji, 2008) and genetic silencing. siRNA is not only encoded by the organisms' dsRNA but also can be encoded by transposons, viruses and synthetically produced to in the form of antisense oligonucleotides (ASO). Later it was found that imperfect complementarity of siRNA, similar to microRNA (miRNA) can result in gene silencing by translational repression or deadenylation of the cognate mRNA (Saxena et al., 2003) suggesting 2 mechanisms for siRNA associated gene silencing. The main drawback of siRNA is that their utility is limited to cells capable of transfection, their effects are transient in mammals, in addition to their potential for "off target" effects due to the unintended complementarity of the target sequence to the gene of interest and the off target gene (Ui-Tei et al., 2008a; Ui-Tei et al., 2008b). For the most part however, with the advancement in next generation sequencing (NGS), off targeting has been mitigated by the conscientious bioinformatics-informed design of the siRNA sequences. Furthermore, from a practical standpoint, the expense in the use in multiple large-scale experiments may be a barrier to its use.

3.4.2 Short hairpin RNA (shRNA) inhibition

An alternative to the use of transient siRNA is the use of shRNA which are short engineered RNA molecules with a tight hairpin turn used to silence target gene expression (Paddison et al., 2002). The drawback of the transiency of siRNA is overcome by manipulating the RNAi to become stable and heritable by enforced expression of the silencing trigger (e.g., an inverted repeat sequence forming a hairpin structure) (Smith et al., 2000; Tavernarakis et al., 2000) thereby eliminating the need for multiple rounds of transfection and improving reproducibility of results. These shRNA are synthesized within the host cell by DNA vector-mediated production. These shRNA are expressed *in vivo* from RNA polymerase III (Pol III) promoters to induce stable suppression in mammalian cells (Lipardi et al., 2001; Sijen et al., 2001). In contrast to siRNAs, shRNA are capable of DNA integration and is comprised of 2 complementary 19-22 base pair (bp) RNA sequences linked by a short loop of 4-11 nucleotides found in miRNA (Paddison et al., 2002). Following transcription, the shRNA is exported into the cytosol and processed by Dicer where the loop of the hairpin is processed off and forms siRNA duplexes. The siRNA formed then specifically bind to the target mRNA, which is subsequently loaded onto the RISC complex for target-specific degradation. While shRNA has its advantages over siRNA, the issues associated with off target effects remains a large caveat to the use of shRNA. Furthermore, disruption in the phenotype after transduction can result from the integration site dependent effects (Musiyenko et al., 2007). To mitigate these shortfalls of the technique, multiple shRNA constructs should be tested and optimized per experiment (Taxman et al., 2006).

3.4.3 CRE recombinase conditional mutagenesis

The use of CRE recombinase to produce conditional mutagenesis and insert DNA cassette into eukaryotic chromosomes is common practice. The simplicity of the CRE-loxP system is the requirement of a short recombination sequence and a single easily purified recombinase enzyme. CRE exists as a monomer even at high concentrations and binds with high affinity to loxP DNA sequences resulting in a dimer of CRE subunits bound to loxP sites (Austin et al., 1981). This dimer in turn binds to a dimer in another loxP site, forming a tetramer. The strands are then cut by CRE protein and rejoined by DNA ligase(Nagy, 2000). Developing a mouse with a CRE-lox system is generated with either CRE under a general or tissue-specific promoter. Mice can also have LoxP sites flanking your target gene or region of interest (i.e., Floxed mice), which are then bred with CRE mice to generate CRE LoxP mice (Sauer, 1998). The gene of interest can then be knocked out with the introduction of CRE recombinase.

The first description RNAi induced by CRE recombinase was by Tiscornia *et al* where they constructed a lentiviral vector carrying a mouse U6 promoter separated from the shRNA by a random sequence of DNA flanked by modified loxP sites (Tiscornia et al., 2004). The silencing cassette is not expressed until the addition of CRE recombinase delivered by a lentiviral vector. The cognate gene is silenced through site-specific recombination, excising loxP-"gene"-loxP cassette to generate a functional intron that contains a single loxP site. A caveat to the use of CRE recombinase-inducible RNAi however is the observation that high

multiplicity of infection (MOI) of the CRE lentivirus has been observed to be toxic to cells and therefore careful titration and experiments to include CRE only conditions have been recommended to be an essential part of study design (Tiscornia et al., 2004).

3.5 Techniques to characterize the CSC's transcriptome, translatome and proteome and their limitations

Characterizing the role of an RBP that has multiple roles within the cell from transcript stabilization to facilitating mRNA translation requires experiments designed to capture its multifaceted nature. Essential to the functional study of an RBP is to 1) identify its binding targets, 2) determine its role in promoting mRNA stability, 3) determine its role in facilitating translation of mRNA to its cognate protein product. An unbiased approach to this would require large-scale studies of RBP binding, whole transcriptome, translatome and proteomic analysis.

3.5.1 RNA-binding protein RNA target identification

The challenges that biologist and biochemists met with prior to the development of cross-linking and immunoprecipitation (CLIP) and high resolution profiling experiments was the uncertainty associated with post-lysis non-specific binding of experiments designed to identify RBP-RNA interaction such as RNA immunoprecipitation (RIP), RIP and sequencing (RIP-seq), and RIP and chromatin immunoprecipitation (RIP-ChIP). These caveats and the evolution of CLIP (i.e., CLIP, iCLIP, PAR-CLIP, eCLIP) have been detailed in the review

article in Chapter 2. Briefly, in CLIP, RBP-RNA interactions are irreversibly stabilized via UV cross-linking, followed by immunoprecipitation of the desired protein using factor specific antibodies and the associated RNA is extracted/isolated and converted into a DNA library for high-throughput sequencing (Ule et al., 2003). Since the publication of the article in Chapter 2 however, there have been further advancements in the eCLIP protocol to further improve the protocol's adaptor strategy (i.e., new 3' linker ligation strategy) to enable single-end enhanced CLIP (seCLIP) removing the requirement for paired end sequencing of eCLIP libraries (Van Nostrand et al., 2017).

CLIP however is not without caveats. PCR amplification can introduce artefacts that can be addressed with the use of unique molecular identifiers (UMI) (Chakrabarti et al., 2018). Furthermore, certain nucleotides and amino acids have been observed to preferentially cross-linked by UV light introducing variant bias with crosslinking efficiency varying between proteins, which can be addressed with cross-linking optimization steps prior to sample processing. An additional biological challenge is the quantification of spurious RNAs as highly abundant RNAs are more likely to produce more CLIP signal. The main challenge experienced by many biologist who employ CLIP in their laboratories, is the downstream analysis, requiring specific algorithms to filter signal from noise (Chakrabarti et al., 2018).

3.5.2 Whole transcriptome characterization

With over a decade of experience with sequencing the dynamic state of genomes, examining RNA quantitative and qualitative we have gained considerable insight into the dynamic nature of the human genome. We now appreciate that the human transcriptome is a complex landscape comprising of both coding and non-coding RNA species. In the era of next generation DNA sequencing (NGS), in a very short period of time, we have improved our ability to identify RNA on a massive scale. The primary goal of most RNA-seq experiments in the context of the study of RBPs is to quantify is and compare the gene expression under varying experimental conditions to be able to infer biological function to differential expression at the transcript level. The change in relative expression of the commonly studies steady state polyA mRNA can inform the investigator of the RBP's role in processing pre-mRNA. Despite the large amount of information we gain from the analysis of the whole transcriptome, RNA-sequencing harbours challenges from experimental design to data analysis. The sequencing depth of RNA-seq experiments varies as transcripts are expressed at different levels within the cell with their coverage differing substantially in any RNA-seq experiment. Deeper sequencing is required for lower abundant transcripts though on balance, deeper sequencing can increase your rate of false positives in differential expression calls (Ozsolak and Milos, 2011). For the majority of genes from human tissue, the amount required has been estimated at 15-50 million reads (Liu et al., 2014). Read length has been postulated to be conducive for reducing mapping bias and ambiguity in assigning reads to genomic elements (Ozsolak and Milos, 2011). However, for humans, no substantial improvements in

transcriptome assembly quality for 150 base pairs and differential expression analysis with reads over 50 base pairs has been found (Chhangawala et al., 2015). As with all large datasets, the challenge in the interpretation of the experimental conditions under study is entirely dependent on the biostatistical analysis applied.

3.5.3 Whole translatome characterization

Building on the quantification of gene expression using RNA-seq, high throughput sequencing to monitor translation *in vivo* with ribosome profiling has provided invaluable insights into the regulation of protein synthesis. In contrast to RNA-seq, ribosome profiling and sequencing (ribo-seq) directly measures the quantity of mRNA (i.e., 30 nucleotides of mRNA) that is protected by a translating ribosome from nuclease activity (Ingolia et al., 2011). Nevertheless, the quantification and processing of ribosome footprints (i.e., where they have been found to be stalled on the mRNA after the addition of an antibiotic to inhibit translation (i.e., anisomycin, cycloheximide) introduce their unique sets of challenges (Brandman and Hegde, 2016). While the translational insights provided by the observations of ribo-seq has been considerable and more closely reflects the rate of protein production, this technique is not without one significant limitation. One of the major caveats of ribo-seq is the inability to identify translationally inactive mRNA that are sequestered into ribonucleic particles and stalled or paused in polysomes. This phenomena has been shown to occur especially in neurons (Ishimura et al., 2016) and is an important consideration in the study of MSI1's role in translation facilitation. Ribo-seq is therefore unable to

discriminate between actively translating and RNA fragments protected by inactive or stalled ribosomes, which can lead to misinterpretations of the ribosome associated fraction of mRNA. To overcome this limitation, a protocol employing puromycylation (i.e., RiboLACE) has been reported to be able generate a global snapshot of active ribosome footprints at a single nucleotide resolution (Clamer et al., 2018). As with RNA-seq, the interpretation of the data is entirely dependent on the biostatistical analysis and often the bottle neck in biological interpretation of data.

3.5.4 Whole proteome characterization

As Chapter 2 has detailed quantitative proteomics and the advancements in the field. The ability to quantify differential protein abundance between experimental samples will determine the correlation between the translatome and the proteome and in part answer the concern for identifying mRNA associated with stalled ribosomes and not actively translated. The use of tandem mass tagged mass spectrometry (TMT-MS) has provided the field of proteomics with the capability of identifying and quantifying >8000 proteins in a given sample. Similar to eCLIP, RNA-seq and ribo-seq, data analysis is one of the major challenges of MS-based shotgun proteomic experiments, and has posed unique challenges that has been difficult to overcome prior to user friendly interfaces and software to process proteomic data. MaxQUANT is one such publically available proteomic pipeline improving peptide mass accuracy and in turn peptide and protein identification (Cox and Mann, 2008). Furthermore, the associated PERSEUS

software has facilitated the non-computational biologist to be able to analyse data to determine significantly differentially abundant proteins (Tyanova et al., 2016). As the final effector of the cell, accurate peptide and protein identification is the foundation from which meaningful data interpretation is achieved.

3.6 Advantages and challenges associated with large datasets and data integration

Much of the study of cancer to date has focused on understanding the disease mechanism through the study of a single level (e.g., genomics, transcriptomics, epigenetics). However the complexity of the human genome and its' multi-level regulation has made the multi-omic the favoured approach to the study of cancer. The challenges that come with the integration of multi-omics is far from insignificant. Each platform is unique in its ability to identify significantly differentially expressed or modified genes or proteins poses the question of where to set the significance cut-off. Purists will argue that the platforms should all have a cut off with identical false discovery rates (FDR) and adjusted *p*-values, and others will maintain that each platform is unique and comparisons across the modalities should be based on modality-specific and meaningful cut-offs. Bearing in mind the advantages and limitations of each platform, integration of multi-omics data should be not only significance driven but guided by biology.

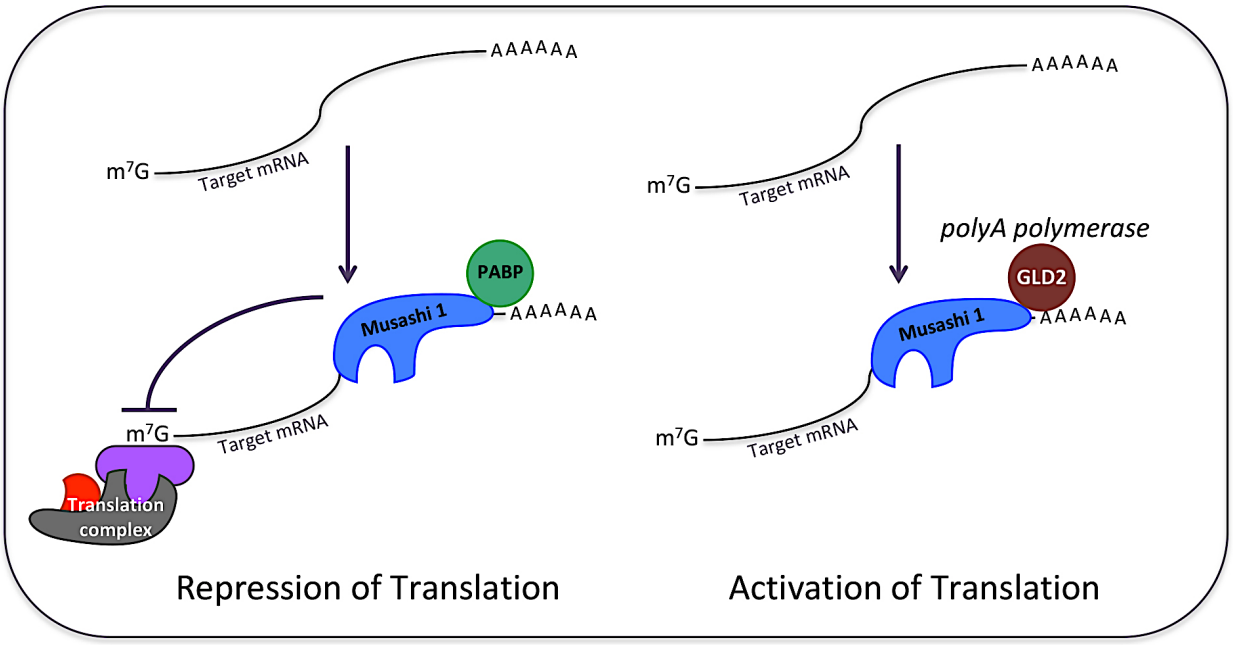


Figure 3.1: Observed mechanisms of action of Musashi family RBP in cancer (modified from Fox *et al*, 2015). A: Translational repression by binding the 3'UTR of the target transcript

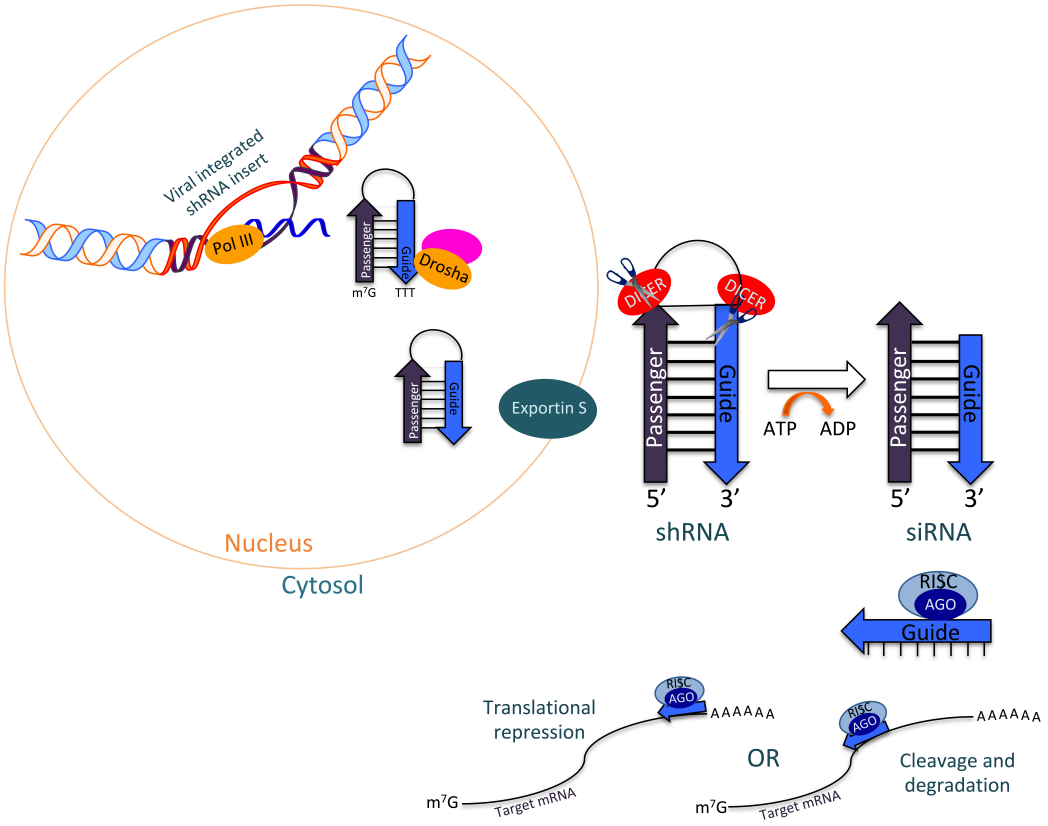


Figure 3.2: Schematic of shRNA gene inhibition. shRNA are synthesized within the host cell by DNA vector-mediated production. These shRNA are expressed *in vivo* from RNA polymerase III (Pol III) promoters to induce stable suppression in mammalian cells. Following transcription, the shRNA is exported into the cytosol and processed by Dicer where the loop of the hairpin is processed off and forms siRNA duplexes. The siRNA formed then specifically bind to the target mRNA, which is subsequently loaded onto the RISC complex for target-specific degradation.

CHAPTER 4

THE ROLE OF MUSASHI-1 IN MEDULLOBLASTOMA STEM CELLS

4.0 Preamble

This chapter is an original article in preparation for submission. It is presented in pre-submission format.

"This article is currently in preparation for submission with data analysis in progress. Kameda-Smith MM, Luo E-C, Zhu H, Xella A, Brown K, Yee B, Tan F, Xing S, Fox RG, Venugopal C, Bakhshinyan D, Adile AA, van Nostrand E, Subapanditha M, Provias J, Moffat J, Fleming A, Hope K, Reimand J, Lu Y, Reya T, Yeo G, Wechsler-Reya R, Singh SK. Musashi-1 is a master regulator of aberrant translation in group 3 medulloblastoma"

MMK-S designed the project, experiments, performed the experiments, analysed data and wrote the manuscript. AX designed and performed the majority of the mouse model of G3MB data and RGF developed the Floxed *Msi1* mouse. BY, E-CL, KB and HZ contributed to the bioinformatic integration of the multiple platforms generated by the project. SKS, GY, RWR co-supervised the project, interpreted results and edited the manuscript.

This project aimed to be the first to demonstrate MSI1's central role in G3 MB tumour propagation. It is also the first to use a systems biology approach to comprehensively characterize the role of MSI1 in a G3 MB using a multi-platform approach integrating eCLIP and sequencing, RNA-seq, ribo-seq and TMT-MS data sets.

Musashi-1 is a master regulator of aberrant translation in *MYC*-amplified Group 3 medulloblastoma

Pediatric medulloblastoma (MB) is the most common solid malignant brain neoplasm, with Group 3 (G3) MB representing the most aggressive subgroup. Despite *MYC* amplification representing an independent poor prognostic factor in G3 MB, efforts to target the MYC pathway have met with limited therapeutic success. As such, alternative mediators of G3 MB continue to be sought. Here we show that the RNA-binding protein, Musashi 1 (MSI1) is an essential moderator of G3 MB in both *MYC*-overexpressing genetically engineered mouse and patient-derived xenograft models. *Msi1* inhibition resulted in an abrogation of tumour initiation in both models, translating to a significantly prolonged survival. We then undertook a multi-omics enquiry of MSI1 function in human G3 MB, through an integrative analysis of global RNA-binding targets, and measurements of steady state and polysome-associated mRNA and protein abundance, revealing MSI1's role in modulating G3 MB-associated cancer driving genes. Our data suggests MSI1 is an essential master regulator that orchestrates the aberrant translational landscape of G3 MB.

(3861 words)

Medulloblastoma (MB) is the most common solid malignant brain neoplasm in children. Within the current consensus, MB comprises four subgroups: Wingless (Wnt), Sonic Hedgehog (Shh), Group 3 (G3) and Group 4 (G4) (Taylor et al., 2012). Group 3 (G3) MB is the most aggressive subgroup with the highest

frequency of children presenting with metastatic disease (Kool et al., 2012), and is associated with a poor prognosis (Forget et al., 2018; Ramaswamy et al., 2016). Multiple high-throughput genomic studies have confirmed that a subset of G3 MB uniquely harbors *MYC* amplification events, which independently confers a high risk for poor prognosis (Cavalli et al., 2017; Cho et al., 2011). An urgent call for novel therapies for this subgroup resulted in the consensus that inhibitors of bromodomains (Bandopadhayay et al., 2014), aurora kinases (Hill et al., 2015) and histone deacetylases (Ecker et al., 2015) be trialed as small molecule therapies to inhibit the MYC pathway (Gottardo et al., 2014). Although MYC is a highly unstable protein, there is a consensus that it has the potential to be an effective target for *MYC*-amplified cancers. Unfortunately, effective means to target MYC are not yet available (ClinicalTrials.gov, 2018a, b; Graff et al., 2016; Schoffski et al., 2011; Seymour et al., 2014) and the search for alternative mediators of the aggressive phenotype of G3 MB continues.

Recent advances in proteogenomics have suggested a requisite role of post-transcriptional modification in G3 MB (Archer et al., 2018; Forget et al., 2018; Zomerman et al., 2018). The high discordance between the transcriptome and proteome in G3 MB indicates a central role of post-transcriptional gene regulation such as alternative splicing, polyadenylation, nuclear degradation and export. While technologies in genomic, epigenomic and proteomic platforms continue to evolve to facilitate high throughput multi-platform analyses of primary tumour tissue, the ability to directly link the observations of the transcriptome and proteome remains a significant challenge. Overcoming these experimental barriers

becomes essential in the study of the brain, whereas a consequence of higher levels of post-transcriptional gene regulation compared to other organs (Wang et al., 2008a; Xu et al., 2002), dysregulation in post-transcriptional modifiers such as RNA binding proteins (RBPs) have been shown to result in neuronal dysfunction and aberrant brain development (Li et al., 2007). To facilitate the study of RBPs with higher fidelity than previous methods of RNA immunoprecipitation, experimental methods such as cross-linking and immunoprecipitation (CLIP) have emerged (Van Nostrand et al., 2016). These methodological advancements paired with the developments in the field of bioinformatics have facilitated the unbiased systematic investigation of this subset of potentially targetable trans-regulatory elements to better delineate their role in cancer.

To understand one of the potential mechanisms that confers an aggressive phenotype on pediatric G3 MB, we focused on the significantly up-regulated neural RNA binding protein, MSI1. While the role of MSI1, as a marker of neural stem and progenitor cells has been established (Sakakibara et al., 2002; Sakakibara and Okano, 1997), its breadth of impact in cancer (Fox et al., 2016; Li et al., 2015) and more specifically neural cancers is only beginning to emerge (Uren et al., 2015; Vo et al., 2012). Cancer is characterized by changes in cell symmetry and self-renewal capacity and therefore it comes as no surprise that genes regulating these processes in normal neural development, such as *MSI1*, may be implicated in developmental brain tumours such as MB. The initial work investigating the role of MSI1 and MB identified its association with cell proliferation (Sanchez-Diaz et al., 2008), regulated by tumour suppressor

miRNAs (Vo et al., 2011), and poor clinical prognosis (Vo et al., 2012). Using the clinically most aggressive genotype, the *MYC*-amplified G3 MB (Cho et al., 2011), we investigated the requirement of the RNA-binding protein, MSI1 in maintaining the stem cell properties of MB brain tumour initiating cells (BTICs). Moreover, to further our understanding of the role of MSI1 in this subset of G3 MB, we proceeded with an unbiased integrative analysis of eCLIP binding sites, with changes observed at both the traditionally studied transcriptome, in addition to proteome-wide changes after sh*Msi1* inhibition. We hypothesize that MSI1 plays a critical role in medulloblastomagenesis and that an integrative multi-omics approach will provide a systems biological perspective of its master regulatory role in *MYC*-amplified G3 MB.

RESULTS

Msi1 is overexpressed at in G3 MB and confers poorer overall survival.

To identify the clinical relevance of *MSI1* upregulation in the largest primary MB publically available database, we interrogated the transcriptomic landscape of *MSI1* in patient-derived MB data published by Cavalli *et al* (Cavalli et al., 2017). Significant upregulation of *MSI1* in G3 and 4 MB was observed as compared to Wnt and Shh MB ($p<0.0001$) however in contrast to Vo *et al*'s survival advantage observed in MSI1 negativity on IHC (Vo et al., 2012) we did not find that $MSI1^{high}$ tumours conferred a survival advantage, highlighting the concerns associated with transcriptomic based interpretations of cancer biology (Figure 4.1A, Supplementary Figure 4.1A). To determine the clinical significance of high

protein levels of MSI1, we stained primary MB slides from G3 and non-G3 subgroups and found, unique to G3 MB, areas of high MSI1 staining with corresponding positivity in Nestin staining and high Ki67 index (Figure 4.1B). Interestingly, G3 and G4 MB showed areas of nuclear MSI1 positivity on immunohistochemistry (IHC) staining whereas in the Wnt subgroup MB (BT853) MSI1 exclusively stained in the cytoplasmic fraction, suggesting a differentially active role for MSI1 in G3/4 MB compared to Wnt MB.

To determine the clinical significance of high MSI1 protein levels in childhood MB, we performed western immunoblot analysis of 2 pediatric metastatic MB cell lines that validated high expression of MSI1 (SU_MB002, HD-MB03) and NSCs cultured from an embryonic posterior fossa (NSC201cb) but not in the Wnt MB line (BT853) (Figure 4.1C). To identify the relationship between MSI1 and other stemness markers, we performed flow cytometric analysis for BMI1 and the putative BTIC marker, CD133 (Singh et al., 2004) as well as stained our xenografted tumours derived from transplantation of the aforementioned tumour cells (SU_MB002, HD-MB03, BT853) with Nestin and MSI1. Both lines were highly enriched for CD133 and BMI1 at the protein level (Figure 4.1D) confirming our patient -derived G3 MB cell lines, SU_MB002 and HD-MB03, to be highly enriched for markers associated with neural cancer stem cells. IHC of primary pediatric MB orthotopic xenografts shows areas of high MSI1 staining with corresponding areas with Nestin in G3 MB when compared to BT853 (Figure 4.1E). To follow up on our IHC findings of nuclear staining of MSI1 in G3 MB, we found that SU_MB002 also had substantial MSI protein within the

nuclear fraction (Supplementary Figure 4.1B). Initially testing our hypothesis in normal NSC, in contrast to Sakakibara *et al* (Sakakibara et al., 2002), we found that sh*MSI1* inhibition (Supplementary Figure 4.1C) lead to significant reductions in NSC properties including secondary neurosphere formation (p<0.001) and proliferation (p<0.001) (Supplementary Figure 4.1D-E) corroborating the high Baysian factor (i.e., essentiality score) for MSI1 in a genome wide CRISPR-Cas9 screen on NSCs (Toledo et al., 2015), implying MSI1 should be avoided as a direct targeted therapy in the setting of an actively developing central nervous system.

MSI1 is required for tumour propagation in murine G3 MB

To investigate the role of MSI1 in *MYC*-amplified G3 MB, we initially used a well-established G3 MB mouse model, driven by overexpression of Myc and dominant-negative p53 (hereafter called the MP model). Mouse MSI1 protein abundance was greater in MP tumours than in adult mouse cerebellum and olfactory bulb, but lower than that seen in embryonic brain (Figure 4.2A). *Msi1* transcript was also significantly upregulated in MP compared to adult mouse cerebellum and comparable to that seen in mouse neural stem cells (p=0.048) (Supplementary Figure 4.2A). Knockdown of mouse *Msi1* (*Msi1*) in MP cells by RNA interference (sh*Msi1*) (Figure 2B, p<0.0001) resulted in an abrogation of neurosphere formation (Figure 4.2C) and proliferation (Supplementary Figure 4.2B-C). Next, to determine whether *Msi1* was necessary for tumour propagation, we used the previously described $Msi1^{flox/flox}$ mice (Fox et al., 2016)(Figure 4.2D-

E). MP tumours were generated from *Msi1*$^{flox/flox}$ mice, and tumour cells were infected with lentiviruses encoding GFP or Cre recombinase. As shown in Figure 2G, Cre significantly diminished *Msi1* mRNA levels compared to control cells (p=0.0015). Deletion of Msi1 impaired the ability of tumour cells to proliferate and form secondary spheres (Figure 4.2H, Supplementary Figure 4.2D). Finally, Cre-infected *Msi1*$^{fl/oxflox}$ MP cells were injected *in vivo* to determine the effects of *Msi1* deletion on tumour growth (Figure 4.2F). Strikingly, whereas mice transplanted with control tumour cells all developed tumours within two months, none of the mice transplanted with Cre-infected tumour cells showed evidence of tumour formation (Figure 4.2I, Supplementary Figure 4.2E), and these animals exhibited 100% survival over 300 days (Figure 4.2J). As Cre-mediated cellular toxicity and induction of complete tumour regression has been described (Li et al., 2014), survival studies were also conducted using MP tumours from wild type (non-*Msi1*$^{flox/flox}$) mice. No significant difference in survival was seen between mice receiving Cre-infected and GFP-infected tumour cells, suggesting the survival benefit in the *Msi1*$^{flox/flox}$ experiments was not secondary to Cre toxicity (Supplementary Figure 4.2F)

MSI1 is critical for growth of human G3 MB

Building on the observation of *Msi1* deletion in a mouse model, we sought to determine if similar tumour suppressive effects would be achievable in human models of G3 MB. Given the essential role of MSI1 in maintaining the stem cell state, and a recent MB002 RNAi screen identifying MSI1 to be essential in SU_MB002 (McFarland et al., 2018), shRNA interference was employed to study

its role in G3 MB *in vitro* (Figure 4.3A). Western immunoblotting showed significant MSI1 protein level inhibition in SU_MB002 cells transduced with sh*MSI1* lentivirus, despite modest *MSI1* KD at the transcript level (Figure 4.3B). Functional assays including secondary sphere formation, proliferation and cell cycle analysis showed significant impairment of stem cell properties after sh*MSI1* inhibition in G3 MB BTICs (Figure 4.3C-E, Supplementary Figure 4.3A). Notably, despite *MSI1* being highly expressed in SU_MB002, further *MSI1* overexpression showed further enhancement of proliferative capacity (Supplementary Figure 4.3B) suggestive of a gene dose effect. To investigate the effect of sh*MSI1* inhibition *in vivo*, immunocompromised mice were injected intracranially with either control (n=12) or sh*MSI1* inhibited (n=12) SU_MB002 cells. MRI scan 5 weeks post-injection, when control mice were symptomatic and nearing endpoint, revealed large tumour growth in control mice where no tumour was identifiable in the matched sh*MSI1* mice (Figure 4.3E). Large tumours were found in control mice, as deduced by H&E staining of the xenografts, where sh*MSI1* mice exhibited substantially reduced tumour burden (Figure 4.3F). Deficits in tumour initiating capacity of sh*MSI1* mice translated into a significant survival benefit as visualized on Kaplan-Meier survival curve (p<0.0008) (Figure 4.3G), where 50% of mice in the sh*MSI1* cohort were asymptomatic at time point sacrifice, and was devoid of any histological evidence of tumour. SU_MB002 findings were confirmed in a second metastatic G3 MB cell line (HD-MB03) with cells that were significantly knocked down for *MSI1* at the transcript level and protein level (Supplementary Figure 4.3C). Large tumours were observed in

control group (n=12) as compared to the sh*MSI1* group (n=12) (Supplementary Figure 4.3D; bars represent 100uM) with sh*MSI1*-engrafted mice experiencing a significant survival benefit (p=0.0002) (Supplementary Figure 4.3E). These results are suggestive that MSI1 is required for tumour propagation in G3 MB.

MSI1 binding targets and whole transcriptome, translatome and proteome multi-omic characterization intersect on genes associated with transcription and translation.

To systematically characterize the functional role of MSI1, we performed a differential analysis of G3 MB BTICs and normal NSCs using enhanced crosslinking and immunoprecipitation and sequencing (eCLIP-seq) (Van Nostrand et al., 2016). We performed eCLIP-seq comparing SU_MB002 and human NSC cultured from and embryonic posterior fossa (NSC201cb) (Figure 4.4A). By performing control screen in fetal neural stem cells (NSC) cultured from the matched anatomical origin of MB, we hypothesized that candidate G3 MB specific down stream targets of an essential neurodevelopmental RBP such as MSI1 can be identified (Sakakibara et al., 1996; Sakakibara et al., 2001; Sakakibara et al., 2002; Sakakibara and Okano, 1997). Replicates were highly correlated *via* gene reads per kilobase of transcript per mapped reads (RPKM) with 1271 and 1382 binding sites mapping to 51 and 494 unique gene symbols in SU_MB002 and NSC201cb respectively (Figure 4.4B, Supplementary Table 4.1A and B). These 2 samples of cells of posterior fossa origin shared 110 MSI1 binding sites in common (Supplementary Figure 4.4A, Supplementary Table 4.2). Included in the shared genes was *CTNNB1* and *eIF4* subunits (i.e., *A2* and *G2*)

suggesting MSI1's role in moderating genes involved in essential processes such as transcription and translation initiation. To determine if eCLIP-bound genes shared between SU_MB002 and NSC201cb are statistically expected or larger than expected by chance alone, Fisher Exact test was applied yielding an OR 13.615 (95% CI: 11.08812, ∞, $p<2.2x10^{-16}$), therefore unlikely that shared genes overlap by chance and represents a set of related genes. A similar pattern of reproducible peak frequency was observed in SU_MB002 and NSC201cb as the bound sequences were primarily in the 3'UTR (i.e., followed by the coding sequence (CDS) and 5'UTR) (Figure 4.4C, Supplementary Table 4,3). Motif analysis verified the G[UAG]U trinucleotide was generally enriched around MSI1 eCLIP sites in both SU_MB002 and NSC201cb as previously reported (Ohyama et al., 2012; Uren et al., 2015) (Figure 4.4D-E). Differential analysis of the genic regions bound by MSI1 in SU_MB002 and NSC201cb identified unique binding transcripts in addition to shared targets that may represent the remnants of the normal stem cell properties (Supplementary Figure 4.4A). Pathway analysis of the 551 and 494 statistically significant MSI1-bound genes and their annotated pathways in SU_MB002 and NSC201cb (Supplementary Table 4.4) identified the enrichment of genes uniquely associated with Chromatin, Transcription and Translation in G3 MB as compared to their presumed cell of origin (NSC201cb), suggestive of MSI1-associated aberrant post-transcriptional modification of essential genes in G3 MB (Figure 4.4G; Supplementary Table 4.5A and B). To identify whether the inhibition of *MSI1* alters the steady state mRNA landscape, and yield some insight into MSI1's role in nascent mRNA processing and

stability, we performed RNA-sequencing (RNA-seq) of sh*MSI1* inhibited compared to control samples of SU_MB002. Despite applying a TREAT threshold, modest inhibition of *MSI1* resulted in 10,273 transcripts with detectable expression differences of which 1609 were significantly differentially expressed (p<0.05). Notably, eCLIP-bound genes were only modestly differentially expressed after sh*MSI1* inhibition, suggestive of an extensive cascading effect despite 22% loss of MSI1 steady-state expression in the system. Only 15 MSI1 eCLIP-bound genes were significantly differentially expressed after sh*MSI1* inhibition suggestive of their direct role in transcript stability or degradation (Figure 4.4G). This observation may also be explained by the binding of other RNA-binding proteins (e.g., *MATR3, HNRNPA1/PA3, NONO, PABPC1*), lncRNA (e.g., MALAT1, SNHG6, PTP4A2, RP11, OIP5) by MSI1 (Supplementary Table 4.1A).

As only 15 eCLIP-bound genes were significantly differentially expressed after sh*MSI1* inhibition in the RNA-seq anaysis with extensive downstream changes within the steady state transcriptome, we examined further downstream to characterize the SU_MB002 translatome. We explored the changes occurring in the polysome fraction after sh*MSI1* inhibition to determine if there was a significant change in the fraction of mRNA associated with ribosomes (i.e., MSI1-mediated accessibility of mRNA to ribosomes) (Figure 4.5A). The polysome profile of the sh*MSI1*-inhibited samples demonstrated a relative increase in polysome-associated mRNA as compared to the control sample

(Figure 4.5B). To gain further insight into which transcripts are differentially associated with the polysome fraction, we subsequently performed polyribosome profiling and sequencing (polysome-seq). Polysome-seq identified 11,385 transcripts that were associated with the polysome fraction with normalized RPKM of the sh*MSI1*-inhibited samples globally showing a reduction in polysome-associated mRNA that are annotated as cancer-associated genes as compared to the control samples (Figure 4.5C). Multiple cancer-associated genes are significantly downregulated in the translatome including *MAP3K13*, *MALAT1*, and the transcriptional enhancers of *MYC* and genes essential in early development, *NPM1* and *KMT2A* respectively.

To determine the final proteomic landscape after sh*MSI1* inhibition, we performed whole shotgun quantitative proteomic analysis (Figure 4.6A). MaxQUANT analysis of our isobaric compound tagged sample peptides showed a right shift in abundance of protein in the knockdown samples compared to controls, suggesting a global increase protein abundance after sh*MSI1* inhibition (Supplementary Figure 4.6A-B). Further analysis identified 350 proteins that were significantly differentially abundant, the majority of which (i.e., 76.57%) experienced fold change >1 after sh*MSI1* inhibition (Figure 4.6B). Protein set enrichment analysis (PSEA) (Lavallee-Adam et al., 2014) revealed downregulation of proteins associated with cancer aggressiveness and an upregulation of proteins associated with neuronal differentiation and apoptosis (Figure 4.6C-D).

Integrative multi-omic analysis identifies HIPK1 as a potential downstream target of MSI1 in G3 MB

To determine significant targetable processes in the transcriptomic, eCLIP, translatomic and proteomic datasets, we adopted an integrated modeling approach. ActivePathways performs multivariate pathway enrichment analysis to identify the functional changes in the experimental system (Paczkowska et al., 2018). Applying stringent significance criteria, we identified one gene to be eCLIP bound and significantly downregulated in the polysome-seq and proteome after sh*MSI1* inhibition (Figure 4.7A). Interrogation of the available reposited Gliovis database showed a positive correlation between MSI1 and HIPK1 expression (Figure 4.7B) with a less ubiquitous expression of HIPK1 throughout the body as compared to MSI1 (Figure 4.7C) suggestive of a potential downstream targetable gene for drug discovery. To add further credence to the validation of this downstream target, HIPK1 has been observed to phosphorylate the cMAP response element binding (CREB) protein resulting in the downregulation of the downstream binding targets of CREB and deactivation of their associated pathways (Hashimoto and Tsuji, 2017). Furthermore, an known MB associated syndrome, Rubinstein-Taybi has been shown to be highly associated with microdeletions of CREB (Blough et al., 2000), and a predisposition for G3 MB (Bourdeaut et al., 2014). In summary, the integrative analysis to interrogate the functional role of the neural RBP, MSI1 has shed light possible altered mechanisms of action as compared to its neural stem cell counterpart, and identified a downstream less ubiquitously expressed gene for targeted therapeutic drug discovery for the treatment of G3 MB.

DISCUSSION

While G3 MB as a molecular subgroup is clinically diverse, patients whose tumours exhibit MYC-amplification have the poorest prognosis (Archer et al., 2018; Cavalli et al., 2017; Cho et al., 2011). Understanding the basic mechanism by which specific gene products interact with DNA, RNA and other proteins requires a study of these regulatory levels not in isolation but in the context of the entire system. To gain further insight into the recent proteomic findings of Forget et al (2018) and Archer et al (2018), we describe how the RNA-binding protein, MSI1 exerts its vast influence at multiple levels of the post-transcriptional milieu, in multiple *MYC*-amplified models of G3 MB.

The striking tumour suppressive effect in both CRE-mediated and shRNA inhibition of a single gene in the mouse and human models respectively, is evidence of the central and vital role MSI1 in G3 MB propagation. Building on the work from Sachez-Diaz *et al* (2008) and Vo *et al* (2011, 2012), who first suggested a vital role for MSI1 in MB, we show extensive *in vitro* and the first *in vivo* confirmation of the requisite role of MSI1 as an essential post-transcriptional gene modifier in G3 MB. Furthermore, to characterize MSI1's role to explain the significant discrepancy observed between mRNA and protein levels in G3 MB, multi-platform analysis suggests a significant role for MSI1 in modulating the genes associated with regulation of RNA processing and translation initiation. The combined impact of the roles of 800-1000 known human RBPs identified using RBPome capture studies (Baltz et al., 2012; Castello et al., 2013) has been estimated to account for as much as 30% of protein expression variation (Castello

et al., 2012; Schwanhausser et al., 2011; Vogel et al., 2010). MSI1's binding of hundreds of transcript in SU_MB002, including the most frequently mutated gene in G3 MB, *SMARCA4* and characteristic G3 MB genes such as *MYC* and *OTX2*, suggests MSI1 upregulation is stabilizing these transcripts and enabling their translated proteins to undertake key driver roles in tumour maintenance and propagation.

The RNA-seq data, while informative, identified a significant number of genes differentially expressed after the perturbation of a single gene. The considerable number of differentially expressed genes after modest sh*MSI1* inhibition can be explained by the identification of a substantial number of transcripts endogenously bound by MSI1 (including other neural-associated RBPs and lncRNA). In SU_MB002, 119 annotated RBPs (Baltz et al., 2012; Castello et al., 2013) and 9 lncRNA were eCLIP bound by MSI1.

While the differential eCLIP analysis between NSC201cb and SU_MB002 provided some insight into which genes are post-transcriptionally modulated by MSI1 in normal NSC population as compared to the G3 MB BTICs, the significance of multiple binding regions of MSI1 to its transcript and how and when these transcripts are bound remain unknown. In SU_MB002, we found MSI1 to be localized to both the nuclear and cytoplasmic fractions in nearly equal ratios suggestive of an important role MSI1 plays in the localization and mRNA processing in G3 MB (Supplementary Figure 4.1B). Further investigation into the subcellular localization may have the powerful potential to further inform the role of MSI1 in differing cellular conditions (e.g., stress, hypoxia). In this study, a

dichotomous approach was taken with the assumption that binding results in modification of unknown functional significance of the target transcript. MSI1 is highly expressed in both fetal and adult neural stem cells. In contrast to Sakakibara *et al*'s (2002) observation that *Msi1* deletion alone did not decreased self-renewal ability of mouse NSCs, we found that with even with modest sh*MSI1* inhibition in human NSCs, significant reduction in stem cell properties such as proliferation and secondary sphere formation was observed (Supplementary Figure 4.1C-E). These finding were suggestive of a central functional role of MSI1 in maintaining the functional stem cell properties of NSCs and should therefore be avoided as a therapeutic drug target.

Moreover, the search for novel cancer driver genes for targeted drug discovery remains a significant challenge in a pediatric tumour such as MB, particularly because the genes associated with tumourigenesis are often implicated in normal development. Notch signaling has been associated with MB and a number of studies have suggested MSI1 binding of the Notch pathway inhibitor *NUMB* to be the overarching mechanism responsible for the downregulation of NOTCH pathway gene products (Imai et al., 2001). Similar to Uren *et al* (Uren et al., 2015) and Sakakibara *et al* (Sakakibara et al., 2002) however, we found that while the extensively studied quintessential target of MSI1, *NUMB*, was expressed in our system, as evidenced in the RNA-seq, it was not bound by the endogenous MSI1 in SU_MB002 nor NSC201cb as previously described. In fact the downregulation of the Notch pathway was not through the MSI1-*Numb* axis but through downregulation of key genes associated with the Notch pathway and genes

associated with the polysome fraction such as *DLL3*, a known suppressor of the Notch pathway in WT SU_MB002. These observations highlight the importance of direct measurements and the requirement of describing interactions through direct measure of changes within the relevant milieu.

Finally, in efforts to piece together the overall functional role of MSI1, we found an integrative analysis of the 4 large scale platforms applying stringent thresholds revealed a lack of overlapping genes between RNA-seq and not only quantitative proteomic analysis but polysome fraction as well. While an expected finding, the degree to which there was such discordance lays emphasis on the multiple post-transcriptional gene regulatory processes that are tightly regulated by a single RBP. HIPK1, a homeodomain interacting protein kinase was identified to be MSI1 eCLIP bound and significantly differentially expressed in two of three large scale unbiased large-scale platforms. The identification of a downstream gene that has a similar transcriptomic profile as MSI1 with a less ubiquitous protein expression throughout the body, and a mechanism of action that clinically has been shown to be the cause of a clinical syndrome associated with G3 MB, holds promise as a potential targeted therapy for further validation.

Overall, this study provides a multi-platform approach to the study of NSC determinant, MSI1, in *MYC*-amplified G3 MB, through characterization of eCLIP, transcriptomic, translatomic and proteomic landscapes, We demonstrate the functional relevance of our findings through multiple reproducible G3 MB models, with multi-omics analysis of SU_MB002 revealing mechanistic insight into the complex biology of MSI1's master regulatory role in aberrant translation

in G3 MB. Furthermore, integrative multi-platform analysis provided insight into a candidate downstream binding target, HIPK1 for further validation and therapeutic drug discovery.

STAR ⋆ METHODS

Detailed methods are provided in the online version of the paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell culture of G3 MB cell lines
  - Cell culture of MP tumours
- METHODS DETAILS
  - Lentiviral production
  - RTqPCR
  - Western immunoblotting
  - Functional stem cell assays
  - FACS sorting
  - Animal husbandry and *in vivo* experiments
  - Immunohistochemistry
  - eCLIP-sequencing (eCLIP-seq)
  - RNA-sequencing (RNA-seq)
  - Polyribo-sequencing (Polyribo-seq)
  - Tandem Mass Tag Mass Spectrophotometry (TMT-MS)
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - eCLIP-seq data processing
  - eCLIP motif analysis
  - eCLIP region-based fold-enrichment analyses
  - RNA-seq data processing and analysis
  - Polysome-seq data processing and analysis
  - Proteomic data processing and analysis
  - Gene set enrichment analyses (GSEA)
  - Pathway analysis, network diagrams and data visualization

SUPPLEMENTAL INFORMATION

Supplemental information includes 6 figures and 9 tables that can be found following the figures in this chapter

ACKNOWLEDGEMENTS

We are indebted to children and families of the pediatric MB community that serve as the sole motivation of this work. We would also like to thank all the members of the SCC-RI community who helped to trouble shoot experimental misadventures. We thank the members of the McMaster Illumina Facility for performing the sequencing to a very high standard, the MRI imaging facility for their expertise and care of mice during the imaging process, and the members of the Stem Cell Unit for their daily care of our mice. MMKS is supported by the Marta and Owen Boris Foundation, the Ontario Ministry of Health, Ontario Graduate Scholarship (OGS), the McMaster Department of Surgery, and Division of Neurosurgery. R.W.-R. is supported by funding from the National Cancer Institute (CA159859-07 and P30 CA30199), Alex's Lemonade Stand Foundation, William's Superhero Fund and the McDowell Charity Trust. S.K.S. is supported by the Canadian Institutes of Health Research Operating Grant, Neurosurgical Research and Education Foundation and American Association of Neurological Surgeons, Pediatric Section, the Ontario Institute for Cancer Research, and McMaster University Department of Surgery.

AUTHOR CONTRIBUTIONS

MKS, AX, CV, KH, YL, TR, RW-R GY and SKS conceptualized the experiments. MKS, AX, RF, FT, DB and AAA performed the experiments and acquired data. MKS, E-CL, HZ, BY, KB, AX, KH, JM, YL, JR, TR, GY, RW-R, SKS analyzed and interpreted the data. MKS wrote the manuscript with significant revisions contributed by JR, GY, RW-R, SKS. AF and SKS provided

guidance related to pediatric neurooncology pre-clinical trials. RW-R, GY, SKS supervised the study. All authors reviewed the results and commented on the manuscript.

DECLARATION OF INTERESTS

Dr. Sheila Singh is the Chief Executive Officer for Empirica Therapeutics.

REFERENCES

Archer, T.C., Ehrenberger, T., Mundt, F., Gold, M.P., Krug, K., Mah, C.K., Mahoney, E.L., Daniel, C.J., LeNail, A., Ramamoorthy, D.*, et al.* (2018). Proteomics, Post-translational Modifications, and Integrative Analyses Reveal Molecular Heterogeneity within Medulloblastoma Subgroups. Cancer Cell *34*, 396-410 e398.

Baltz, A.G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M.*, et al.* (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell *46*, 674-690.

Bandopadhayay, P., Bergthold, G., Nguyen, B., Schubert, S., Gholamin, S., Tang, Y., Bolin, S., Schumacher, S.E., Zeid, R., Masoud, S.*, et al.* (2014). BET bromodomain inhibition of MYC-amplified medulloblastoma. Clin Cancer Res *20*, 912-925.

Blough, R.I., Petrij, F., Dauwerse, J.G., Milatovich-Cherry, A., Weiss, L., Saal, H.M., and Rubinstein, J.H. (2000). Variation in microdeletions of the cyclic AMP-responsive element-binding protein gene at chromosome band 16p13.3 in the Rubinstein-Taybi syndrome. Am J Med Genet *90*, 29-34.

Bourdeaut, F., Miquel, C., Richer, W., Grill, J., Zerah, M., Grison, C., Pierron, G., Amiel, J., Krucker, C., Radvanyi, F.*, et al.* (2014). Rubinstein-Taybi syndrome predisposing to non-WNT, non-SHH, group 3 medulloblastoma. Pediatr Blood Cancer *61*, 383-386.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M.*, et al.* (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell *149*, 1393-1406.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2013). System-wide identification of RNA-binding proteins by interactome capture. Nat Protoc *8*, 491-500.

Cavalli, F.M.G., Remke, M., Rampasek, L., Peacock, J., Shih, D.J.H., Luu, B., Garzia, L., Torchia, J., Nor, C., Morrissy, A.S.*, et al.* (2017). Intertumoural Heterogeneity within Medulloblastoma Subgroups. Cancer Cell *31*, 737-754 e736.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

Cho, Y.J., Tsherniak, A., Tamayo, P., Santagata, S., Ligon, A., Greulich, H., Berhoukim, R., Amani, V., Goumnerova, L., Eberhart, C.G.*, et al.* (2011). Integrative genomic analysis of medulloblastoma identifies a molecular subgroup that drives poor clinical outcome. J Clin Oncol *29*, 1424-1430.

ClinicalTrials.gov (2018a). A Dose Exploration Study With MK-8628 in Participants With Selected Advanced Solid Tumours (MK-8628-006) (Bethesda (MD): National Library of Medicine (US).).

ClinicalTrials.gov (2018b). A Dose-Finding Study of MK-8628, a Small Molecule Inhibitor of the Bromodomain and Extra-Terminal (BET) Proteins, in Adults With Selected Advanced Solid Tumours (MK-8628-003) (Bethesda (MD) National Library of Medicine (US).).

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.

Ecker, J., Oehme, I., Mazitschek, R., Korshunov, A., Kool, M., Hielscher, T., Kiss, J., Selt, F., Konrad, C., Lodrini, M.*, et al.* (2015). Targeting class I histone deacetylase 2 in MYC amplified group 3 medulloblastoma. Acta Neuropathol Commun *3*, 22.

Forget, A., Martignetti, L., Puget, S., Calzone, L., Brabetz, S., Picard, D., Montagud, A., Liva, S., Sta, A., Dingli, F.*, et al.* (2018). Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. Cancer Cell *34*, 379-395 e377.

Fox, R.G., Lytle, N.K., Jaquish, D.V., Park, F.D., Ito, T., Bajaj, J., Koechlein, C.S., Zimdahl, B., Yano, M., Kopp, J.*, et al.* (2016). Image-based detection and targeting of therapy resistance in pancreatic adenocarcinoma. Nature *534*, 407-411.

Gottardo, N.G., Hansford, J.R., McGlade, J.P., Alvaro, F., Ashley, D.M., Bailey, S., Baker, D.L., Bourdeaut, F., Cho, Y.J., Clay, M.*, et al.* (2014). Medulloblastoma Down Under 2013: a report from the third annual meeting of the International Medulloblastoma Working Group. Acta Neuropathol *127*, 189-201.

Graff, J.N., Higano, C.S., Hahn, N.M., Taylor, M.H., Zhang, B., Zhou, X., Venkatakrishnan, K., Leonard, E.J., and Sarantopoulos, J. (2016). Open-label, multicenter, phase 1 study of alisertib (MLN8237), an aurora A kinase inhibitor, with docetaxel in patients with solid tumours. Cancer *122*, 2524-2533.

Hashimoto, K., and Tsuji, Y. (2017). Arsenic-Induced Activation of the Homeodomain-Interacting Protein Kinase 2 (HIPK2) to cAMP-Response Element Binding Protein (CREB) Axis. J Mol Biol *429*, 64-78.

Hill, R.M., Kuijper, S., Lindsey, J.C., Petrie, K., Schwalbe, E.C., Barker, K., Boult, J.K., Williamson, D., Ahmad, Z., Hallsworth, A.*, et al.* (2015). Combined MYC and P53 defects emerge at medulloblastoma relapse and define rapidly progressive, therapeutically targetable disease. Cancer Cell *27*, 72-84.

Imai, T., Tokunaga, A., Yoshida, T., Hashimoto, M., Mikoshiba, K., Weinmaster, G., Nakafuku, M., and Okano, H. (2001). The neural RNA-binding protein Musashi1 translationally regulates mammalian numb gene expression by interacting with its mRNA. Mol Cell Biol *21*, 3888-3900.

Kool, M., Korshunov, A., Remke, M., Jones, D.T., Schlanstein, M., Northcott, P.A., Cho, Y.J., Koster, J., Schouten-van Meeteren, A., van Vuurden, D.*, et al.* (2012). 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.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res *19*, 1639-1645.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A.*, et al.* (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res *44*, W90-97.

Lavallee-Adam, M., Rauniyar, N., McClatchy, D.B., and Yates, J.R., 3rd (2014). PSEA-Quant: a protein set enrichment analysis on label-free and label-based protein quantification data. J Proteome Res *13*, 5496-5509.

Li, N., Yousefi, M., Nakauka-Ddamba, A., Li, F., Vandivier, L., Parada, K., Woo, D.H., Wang, S., Naqvi, A.S., Rao, S.*, et al.* (2015). The Msi Family of RNA-Binding Proteins Function Redundantly as Intestinal Oncoproteins. Cell Rep *13*, 2440-2455.

Li, Q., Lee, J.A., and Black, D.L. (2007). Neuronal regulation of alternative pre-mRNA splicing. Nat Rev Neurosci *8*, 819-831.

Li, Y., Choi, P.S., Casey, S.C., and Felsher, D.W. (2014). Activation of Cre recombinase alone can induce complete tumour regression. PLoS One *9*, e107589.

Lovci, M.T., Ghanem, D., Marr, H., Arnold, J., Gee, S., Parra, M., Liang, T.Y., Stark, T.J., Gehman, L.T., Hoon, S.*, et al.* (2013). Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. Nat Struct Mol Biol *20*, 1434-1442.

McFarland, J.M., Ho, Z.V., Kugener, G., Dempster, J.M., Montgomery, P.G., Bryan, J.G., Krill-Burger, J.M., Green, T.M., Vazquez, F., Boehm, J.S.*, et al.* (2018). Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. Nat Commun *9*, 4610.

Milde, T., Lodrini, M., Savelyeva, L., Korshunov, A., Kool, M., Brueckner, L.M., Antunes, A.S., Oehme, I., Pekrun, A., Pfister, S.M.*, et al.* (2012). HD-MB03 is a novel Group 3 medulloblastoma model demonstrating sensitivity to histone deacetylase inhibitor treatment. J Neurooncol *110*, 335-348.

Ohyama, T., Nagata, T., Tsuda, K., Kobayashi, N., Imai, T., Okano, H., Yamazaki, T., and Katahira, M. (2012). Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. Nucleic Acids Res *40*, 3218-3231.

P'ng, C., Green, J., Chong, L.C., Waggott, D., Prokopec, S.D., Shamsi, M., Nguyen, F., Mak, D.Y.F., Lam, F., Albuquerque, M.A.*, et al.* (2019). BPG: Seamless, automated and interactive visualization of scientific data. BMC Bioinformatics *20*, 42.

Paczkowska, M., Barenboim, J., Sintupisut, N., Fox, N., Zhu, H., Abd-Rabbo, D., Group, P.N.a.P.A., Boutros, P., and Reimand, J. (2018). Integrative pathway enrichment analysis of multivariate omics data. BbioRxiv.

Pei, Y., Moore, C.E., Wang, J., Tewari, A.K., Eroshkin, A., Cho, Y.J., Witt, H., Korshunov, A., Read, T.A., Sun, J.L.*, et al.* (2012). An animal model of MYC-driven medulloblastoma. Cancer Cell *21*, 155-167.

Ramaswamy, V., Remke, M., Bouffet, E., Bailey, S., Clifford, S.C., Doz, F., Kool, M., Dufour, C., Vassal, G., Milde, T.*, et al.* (2016). Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. Acta Neuropathol.

Reimand, J., Kull, M., Peterson, H., Hansen, J., and Vilo, J. (2007). g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res *35*, W193-200.

Robertson, D., Savage, K., Reis-Filho, J.S., and Isacke, C.M. (2008). Multiple immunofluorescence labelling of formalin-fixed paraffin-embedded (FFPE) tissue. BMC Cell Biol *9*, 13.

Sakakibara, S., Imai, T., Hamaguchi, K., Okabe, M., Aruga, J., Nakajima, K., Yasutomi, D., Nagata, T., Kurihara, Y., Uesugi, S.*, et al.* (1996). Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. Dev Biol *176*, 230-242.

Sakakibara, S., Nakamura, Y., Satoh, H., and Okano, H. (2001). Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci *21*, 8091-8107.

Sakakibara, S., Nakamura, Y., Yoshida, T., Shibata, S., Koike, M., Takano, H., Ueda, S., Uchiyama, Y., Noda, T., and Okano, H. (2002). RNA-binding protein Musashi family: roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. Proc Natl Acad Sci U S A *99*, 15194-15199.

Sakakibara, S., and Okano, H. (1997). Expression of neural RNA-binding proteins in the postnatal CNS: implications of their roles in neuronal and glial cell development. J Neurosci *17*, 8300-8312.

Sanchez-Diaz, P.C., Burton, T.L., Burns, S.C., Hung, J.Y., and Penalva, L.O. (2008). Musashi1 modulates cell proliferation genes in the medulloblastoma cell line Daoy. BMC Cancer *8*, 280.

Schoffski, P., Jones, S.F., Dumez, H., Infante, J.R., Van Mieghem, E., Fowst, C., Gerletti, P., Xu, H., Jakubczak, J.L., English, P.A.*, et al.* (2011). Phase I, open-label, multicentre, dose-escalation, pharmacokinetic and pharmacodynamic trial of the oral aurora kinase inhibitor PF-03814735 in advanced solid tumours. Eur J Cancer *47*, 2256-2264.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.

Seymour, J.F., Kim, D.W., Rubin, E., Haregewoin, A., Clark, J., Watson, P., Hughes, T., Dufva, I., Jimenez, J.L., Mahon, F.X.*, et al.* (2014). A phase 2 study of MK-0457 in patients with BCR-ABL T315I mutant chronic myelogenous leukemia and philadelphia chromosome-positive acute lymphoblastic leukemia. Blood Cancer J *4*, e238.

Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. Nature *432*, 396-401.

Subapanditha, M.K., Adile, A.A., Venugopal, C., and Singh, S.K. (2019). Flow Cytometric Analysis of Brain Tumour Stem Cells. Methods Mol Biol *1869*, 69-77.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set

enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Taylor, M.D., Northcott, P.A., Korshunov, A., Remke, M., Cho, Y.J., Clifford, S.C., Eberhart, C.G., Parsons, D.W., Rutkowski, S., Gajjar, A.*, et al.* (2012). Molecular subgroups of medulloblastoma: the current consensus. Acta Neuropathol *123*, 465-472.

Toledo, C.M., Ding, Y., Hoellerbauer, P., Davis, R.J., Basom, R., Girard, E.J., Lee, E., Corrin, P., Hart, T., Bolouri, H.*, et al.* (2015). Genome-wide CRISPR-Cas9 Screens Reveal Loss of Redundancy between PKMYT1 and WEE1 in Glioblastoma Stem-like Cells. Cell Rep *13*, 2425-2439.

Uren, P.J., Vo, D.T., de Araujo, P.R., Potschke, R., Burns, S.C., Bahrami-Samani, E., Qiao, M., de Sousa Abreu, R., Nakaya, H.I., Correa, B.R.*, et al.* (2015). RNA-Binding Protein Musashi1 Is a Central Regulator of Adhesion Pathways in Glioblastoma. Mol Cell Biol *35*, 2965-2978.

van Nostrand, E.L., Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen J-Y, Cody NAL, Dominguez D, Olsen S, Sundararaman B, Zhan L, Bazile C, Bouvrette LPB, Bergalet J, Duff MO,Garcia KE, Gelboin-Burkhart C, Hochman M, Lambert NJ, Li H, Nguyen TB, Palden T, Rabano I, Sathe S, Stanton R, Su A, Wang R, Yee BA, Zhou B, Louie AL, Aigner S, Fu X-D, Lecuyer E, Burge CB, Graveley BR, Yeo GW. (2018). A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. bioRxiv.

Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K.*, et al.* (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods *13*, 508-514.

Vizcaino, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T.*, et al.* (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res *44*, 11033.

Vo, D.T., Qiao, M., Smith, A.D., Burns, S.C., Brenner, A.J., and Penalva, L.O. (2011). The oncogenic RNA-binding protein Musashi1 is regulated by tumour suppressor miRNAs. RNA Biol *8*, 817-828.

Vo, D.T., Subramaniam, D., Remke, M., Burton, T.L., Uren, P.J., Gelfond, J.A., de Sousa Abreu, R., Burns, S.C., Qiao, M., Suresh, U.*, et al.* (2012). The RNA-binding protein Musashi1 affects medulloblastoma growth via a network of cancer-related genes and is an indicator of poor prognosis. Am J Pathol *181*, 1762-1772.

Vogel, C., Abreu Rde, S., Ko, D., Le, S.Y., Shapiro, B.A., Burns, S.C., Sandhu, D., Boutz, D.R., Marcotte, E.M., and Penalva, L.O. (2010). Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. Mol Syst Biol *6*, 400.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470-476.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis, 2nd edn (Springer International Piublishing).

Wu, G., Feng, X., and Stein, L. (2010). A human functional protein interaction network and its application to cancer data analysis. Genome Biol *11*, R53.

Xu, Q., Modrek, B., and Lee, C. (2002). Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. Nucleic Acids Res *30*, 3754-3766.

Zomerman, W.W., Plasschaert, S.L.A., Conroy, S., Scherpen, F.J., Meeuwsen-de Boer, T.G.J., Lourens, H.J., Guerrero Llobet, S., Smit, M.J., Slagter-Menkema, L., Seitz, A.*, et al.* (2018). Identification of Two Protein-Signaling States Delineating Transcriptionally Heterogeneous Human Medulloblastoma. Cell Rep *22*, 3206-3216.

STAR ⋆ METHODS

KEY RESOURCE TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
MSI1	Abcam	ab52865
BMI1	R&D	MAB33342
Nestin	Millipore	AB5922
CD133/2-APC, human clone REA820	Miltenyi	130-112-196
β-tubulin	Abcam	ab6046
GAPDH	Abcam	ab8245
Biological Samples		
SU_MB002	Bandopadhayay *et al,* 2014	
Neural stem cells	This study	
HD-MB03	Milde *et al,* 2012	
MP tumours	Pei *et al*, 2011	
Critical Commercial Assays		
Super Script II cDNA synthesis Kit	Invitrogen	Cat: 18064-014
Quant –iT dsDNA BR Assay Kit	Invitrogen	Cat: Q32850
NEBNext Ultra II Directional RNA Library Prep Kit for Illumina	NEB	Cat: E7760S
Deposited Data		
MS data	This study	PRIDE: PXD012432 (ProteomeXchange PRIDE database)
Raw and processed RNA-seq data	This study	GEO: GSE126337
Raw and processed Polysome-seq data	This study	To be deposited
Raw and processed eCLIP data	This study	GEO: GSE126263
Software and Algorithms		
STAR	Dobin et al, 2013	https://github.com/alexdobin/STARRT/releases
Bowtie 2	Landmead and Salzberg, 2012	http://bowtie-bio-sourcefourge.net/bowtie2/index.shtml
R		https://r-project.org/
GSEA	Subramanian et al, 2005	http://software.broadinstitute.org/gsea/index.jsp
GO	Chen et al, 2013	http://geneontology.org/docs/go-enrichment-analysis/
EnrichR	Kuleshov et al, 2016	
Samtools merge v1.6	Li et al, 2009	http://samtools.sourceforge.net/
g:Profiler	Reimand et al, 2007	https://biit.cs.ut.ee/gprofiler/
ActivePathways	Paczkowska et al, 2018	https://github.com/reimandlab/ActivePathways/
Cytoscape (v.3.6.0)	Shannon et al, 2003	http://www.cytoscape.org
Boutroslab.plotting.general (v.5.9/2)	P'ng et al, 2019	https://labs.oicr.on.ca/boutros-lab/software/bpg

CONTACT FOR REAGENT AND RESOURCE SHARING

Requests for further information, resources and reagents should be addressed to the corresponding author, Sheila K. Singh (ssingh@mcmaster.ca)

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell culture of G3 MB cell lines

Primary human pediatric MBs, SU_MB002 and HD-MB03 were kind gifts from Dr. Yoon-Jae Cho (Harvard, MS) and Dr. Till Milde (Heidelberg) respectively. SU_MB002 is primary G3 MB cell line developed from a sample acquired at autopsy from a 4-year old boy who was metastatic at presentation and given his poor post-operative clinical status, parents had turned down radiotherapy and he received post-operative adjuvant cyclophosphamide. He initially responded to treatment only to experience disseminated disease within 3 months and palliated. HD-MB03 is G3MB line established from a fresh tissue section obtained at therapeutic intervention from a 3-year old boy who was metastatic at presentation as previously described (Milde et al., 2012). The primary human MB, BT853 is a Wnt MB cell line established from fresh tissue section at surgical resection from a 5-year old female. BT853 was established after informed consent from the family and as approved by the Hamilton Health Sciences/McMaster Health Sciences Research Ethics Board. The sample was dissociated in PBS containing 0.2 Wünsch unit/mL of Liberase Blendzyme 3 (Roche), and incubated at 37^0C in a shaker for 15 minutes. The dissociated tissue was filtered through a 70-µmcell strainer and collected by centrifugation (1200 rpm, 3 minutes). Tumour cells were resuspended in a serum-free BTIC enrichment media, and replated on ultra-low attachment plates (Corning). BTIC enrichment media was composed of NeuroCult complete media (StemCell Technologies, 10 ng/mL bFGF, 20 ng/mL EGF, 2 µg/mL heparin). Expansion media was used prior to

experimentation and BTIC enrichment. SU_MB002 and was expanded using the same BTIC enrichment media. HDMB-03 was expanded with the BTIC enrichment media supplemented with 10% fetal bovine serum (FBS). BT853 was expanded with Dulbecco's Modified Eagle's Medium high glucose (Life Technologies #11965-118) supplemented with 10% FBS. All samples were cultured in BTIC enrichment media for at least 48 hours prior to experimentation.

Cell culture of MP tumours

The development and culturing of MP tumours have been previously described (Pei et al., 2012). MP tumours were *in vivo* expanded *in vivo* in 6-8 week old NOD-SCID cerebellum. Cells from the mouse cerebellum were harvested after the animal reached endpoint. On the same day as tumour processing, $1.0x10^5$ cells were plated in 500µL Mouse Neurocult Complete™ (StemCell Technologies). The *Msi1*$^{f/f}$ (*Msi1*$^{flox/flox}$) mice were generated by conventional gene targeting by inserting LoxP sites around exons 1–4 (Genoway, France; Figure 4.2 D-E). The strategy results in the conditional deletion of 267bp of coding sequences encoding for the ATG. RNA interference knock down of the cells were performed via lentiviral transduction (mouse sh*Msi1*/CTRL-GFP) on the same day as tumour processing. Subsequently cells were sorted 48 hours post-transduction for mCherry-GFP-NearIR live cell positivity for *in vitro* functional assays.

METHOD DETAILS

Lentiviral production

Human Msi1 shRNA were developed in the Dr. Jason Moffat laboratory (Lentiviral TRC RNAi library, University of Toronto). Replication-incompetent lentiviruses were produced by co-transfection of the expression and packaging vectors pMD2G and psPAX2 in HEK293FT cells. Viral supernatants were harvested 72 hours after

transfection, filtered through a 0.45 µm cellulose acetate filter, and ultracentrifuged at 12,000rpm at 4°C for 2 hours. The viral pellet was resuspended into 200 µL of DMEM media and stored in -80 °C. Cells were analyzed for all *in vitro* and *in vivo* studies 96 hours after transduction and puromycin selection. Mouse *Msi1* shRNA were developed in the Dr. Tannishitha Reya laboratory. Viruses for cell transduction were produced with the same protocol as the Human *MSI1* shRNA.

RT-qPCR

Total cellular RNA was isolated using the Norgen Total RNA isolation kit and quantified using a NanoDrop Spectrophotometer ND-1000. Complementary DNA was synthesized from 0.5-1.0 µg RNA using iScript cDNA Super Mix (Quanta Biosciences) and a C1000 Thermo Cycler (Bio-Rad) with the following cycle parameters: 4 min at 25°C, 30 min at 42°C, 5min at 85°C and hold at 4°C. qRT-PCR was performed using Perfecta SybrGreen (Quanta Biosciences) and an CFX96 instrument (Bio-Rad). Gene expression was quantified using CFX Manager 3.0 software and expression levels were normalized to GAPDH/β-actin expression. Primers are listed in the table below.

Table 4.1: RT-qPCR primers

Gene	Forward Sequence	Reverse Sequence
MSI1	5' - CAC CAA TGG GTA CCA CTG AA – 3'	5' - ACT CGT GGT CCT CAG TCA GC – 3'
MYC	5' – AAT GAA AAG GCC CCC AAG GTA GTT ATC C -3'	5' – GTC GTT TCC GCA ACA AGT CCT CTT C -3'
GAPDH	5' - TGA ACC ACC AAC TGC TTA GC – 3'	5' - GGC ATG GAC TGT GGT CAT GAG – 3'
β-actin	5' –TAT CCC TGT ACG CCT CT – 3'	5' – AGG TCT TTG CGG ATG T – 3'

Western immunoblotting

Denatured total cell protein (10 µg) was separated using 10% Bis-Tris gel electrophoresis and transferred to nitrocellulose membranes. Western blots were probed with the following antibodies: β-tubulin (rabbit; 1:50,000 Abcam #ab6046), GAPDH (mouse;

1:2000; Abcam #ab8245), MSI1 (rabbit; 1:2000; Abcam #ab52865), mutant p53 (rabbit; 1:1000; Abcam #ab32049). Horseradish peroxidase conjugated with goat anti-rabbit IgG was employed as the secondary antibody (Bio-Rad). The bands were visualized with Chemidoc™ MP Imaging Systems (Bio-Rad) using the ImageLab version 15.2.1 software.

Functional stem cell assays

Cell proliferation assay

Single cells were plated in 96 wells, at a density of 100 cells/200 µL per well in quadruplicates for each sample and incubated for five days. 20 µL of a fluorescent cell metabolism indicator, Presto Blue (Life Technologies), was added to each well 4 hours prior to the readout. Fluorescence was measured with a FLUOstar Omega Fluorescence 556 microplate reader (BMG Labtech) at an excitation and emission wavelength of 540 and 570 nm respectively. Resultant readings were analyzed using the Omega software.

Secondary colony formation assay

SU_MB002 tumour aggregates were mechanically dissociated with a 1000 µL pipette tip whereas MP tumourspheres cells were enzymatically dissociated using TryplE™ (Thermo Fisher). Cells were live sorted at a density of 100 cells per well in 200 µL of BTIC enrichment media.

FACs sorting

sh*Msi1*-GFP transduced MP cells were dissociated and suspended in PBS+0.5M EDTA+1%FBS. Samples were sorted using MoFlo XDP cell sorter (Beckman Coulter). Dead cells were excluded using the viability dye near IR Live/Dead™ fixable staining kit (Life technologies, cat. L10119). Compensation was performed using mouse IgG

CompBeads™ (BD, Cat. 552843). External and internal staining were performed as previously described (Subapanditha et al., 2019). Analysis probed for CD133 with an anti-CD133 human clone REA820 (Miltenyi, #130-112-196) and BMI1 with an anti-BMI (Miltenyi, #130-106-736) stem cell markers. GFP expression was defined as positive or negative based on the analysis of regions established by the isotype control. Cells were sorted into 96 well plates containing 180µL of mouse Neurocult complete media (Stemcell Technologies). Small aliquots from each sort were analyzed to determine the purity of the sorted populations. Cells equilibrated a 37^{o}C overnight prior to experimentation. Cells were stained for 20 minutes with Near-IR live stain and mCherry-GFP-live sorted for *in vitro* experimentation. SU_MB002 cells were stained for 5 minutes with 7AAD live stain (Beckman Coulter, A07704) for *in vitro* experimentation.

Animal husbandry and in vivo experiments

All *in vivo* experiments were performed in accordance to the McMaster University Animal Research Ethics Board (AREB) approved protocols. Intracerebellar injections were all performed by the first author via free hand using anatomical landmarks (i.e., 5mm caudal to lambda and to the right 5mm laterally) and injected at a depth of 3mm using a 10uL Hamilton syringe for SU_MB002 and HDMB-03. The number of live cells was determine by using Trypan Blue (Thermo Fisher) exclusion and resuspended in 5uL of BTIC enrichment media. NOD-SCID mice were anesthetized using isofluorane gas (5% induction, 2.5% maintenance) and the cells were injected into right cerebellar hemisphere. Tumour-initiating capacities of SU_MB002 and HD-MB03 comparing control and sh*MSI1* knockdown constructs were performed by injecting 5.0x10^{5} and 1.0 x10^{3} cells respectively. Mice were assessed for histological differences of tumour burden (n=6 in each arm) and survival (additional n=6 in each arm).

In vivo imaging

Luciferase-based bioluminescence imaging was performed on MP tumours as previously described (Pei et al., 2012). Briefly, mice were given intraperitoneal injections of 150ng/g D-Luciferin (Caliper Life Sciences, #12279) and anesthetized with 2.5% isoflurane. At 7-8 min after injections, animals were imaged using the Xenogen Spectrum (IVIS®-200) imaging system. The mouse MRI imaging was performed on an automated 7Tesla wide-bore nuclear magnetic resonance (NMR) system (Bruker WB300). The protocol allows fro tumour visualization without the typical necessity of injectable gadolinium contrast agents. The mouse is anesthetized in an induction chamber at 5% isoflurane in pure O_2 and subsequent positioning in the imaging bed with continuous anesthesia with 1.5-2.5% isoflurane delivered with pure O_2 via nose cone. With a combination of stretches, rotations and translations the images are warped into a common alignment, which allows a direct spatial comparison between an animal and a set of health controls, or an animal at multiple time points allowing visual detection of tumours less than 0.5mm diameter confirmed by histopathology.

Immunohistochemistry

Unstained slides mounted with formalin fixed paraffin embedded tissue were de-paraffinized in xylene, blocked in 3% hydrogen peroxide, and antigen retrieval or unmasking procedure applied for Msi1, BMI1, and nestin staining were performed as previously described at the University Health Network Medical Laboratory Technologies(Robertson et al., 2008). Anti-rabbit-MSI1 primary antibody (Millipore, #AB5977) at 1/500 dilution, anti-rabbit-Nestin (Millipore, #AB5922) at 1:15,000 dilution, and anti-mouse-BMI1 (R&D, #MAB33342) at 1:500 dilution and incubated at room temperature for 1 hour. For the BMI1 IHC, MACH 4 reagents were then applied as directed (Intermedico, #BC-M4U534L) and developed using DAB (DAKO, #K3468).

And for MSI1 and Nestin, ImmPress® reagent were applied as per kit instruction. For all slides, Mayer's Hematoxylin was used to lightly counterstain and the slides were dehydrated. The slides were then mounted with MM 24 Leica mounting medium (Leica, #3801120).

eCLIP-seq library preparation

The eCLIP was performed on SU_MB002 and NSC201cb as previously described (Van Nostrand et al., 2016). $20x10^6$ cells were washed in ice-cold PBS and UV cross-linked at $600mJcm^{-2}$ on ice in a Stratalinker 2400. Cells were pelleted, lysed in lysis buffer and bound to Msi1 antibody (rabbit; 1:2000; Abcam #ab52865) bound magnetic Dynabeads™ M-280 sheep anti-rabbit beads (LifeTech, #11203D). RBP-RNA complexes were then captured on beads overnight in 4°C. The samples were then dephosphorylated and the 5'- and 3'-phosphate groups from RNA using FastAP (LifeTech, #EF0652) and T4 PNK (NEB, #M0314L). The samples were then washed and 3' linker ligated with barcoded RNA adaptors. The samples were then electrophoresed on a 1.5mm 4-12% Bis-Tris gel and transferred onto a nitrocellulose membrane at 30V at 4°C overnight. The diagnostic membrane was developed to ensure adequate pulldown of Msi1 associated RNA and the lanes cut and RNA purified using acid phenol:chloroform:isoamyl alcohol (125:24:1, v/v; pH 4.5; ThermoFisher; #AM9720). The resultant RNA was cleaned and concentrated in Zymo columns (Zymo; #R1016). The RNA was then reverse transcribed using the AR17 primer: 5'ACACGACGCTCTTCCGA3' and AffinityScript RT (Agilent, #600107) master mix. The resultant cDNA was cleaned with ExoSAPit treatment and the RNA removed. The cDNA was then 5' linker ligated with rand3Tr3 adapter using RNA Ligase (NEB, #M0437M). After Silane bead clean of the linker-ligation, the cDNA was quantified using qPCR with subsequent PCR amplification using 2x Q5 PCR master mix (NEB, #M0492L). The DNA was then electrophoresed in a 3% low melting agarose gel

and the library extracted using Qiagen MinElute gel extraction kit without heating and resuspended in nuclease-free water. The library was prepared and sent for paired-end 75bp Illumina sequencing at McMaster University. Of the best biological replicate, 8,117,399 and 17,778,371 reads from the SU_MB002 and NSC201cb IP libraries respectively passed quality filtering, of which 57.8% and 27.2% usable reads mapped uniquely to the human genome (hg19). eCLIP data reproducibility was verified through correlation between gene RPKM and statistically significant overlaps in clusters and genes within replicates.

RNA-sequencing (RNA-seq) library preparation

Total RNA was extracted from each sample and purified using the Norgen Total RNA isolation kit, and quantified using a NanoDrop Spectrophotometer ND-100 and quality control was assessed with bioanalyser total RNA Nanodrop. One µg of mRNA was fragmented to an average length of 200 bp by incubation for 5 min at 94^{o}C with 5X fragmentation buffer (Illumina, RS-100-0801). Efficiency of the fragmentation was defined on Bioanalyzer RNA Pico Chip. The fragmented mRNA was randomly primed and reversed transcribed using Super Script II cDNA synthesis kit (Invitrogen, 18064-014). After second-strand synthesis, the cDNA when through end-repair and ligation reactions according to the Illumina mRNA-Seq Sample Prep Kit protocol. The cDNA library was size-fractioned on a 2% TBE agarose gel. Material in the 350-400 bp range was excised and purified (Zymo Research, D4001). Half of the eluted cDNA library was used as a template for amplification according to mRNA-Seq Sample Prep Kit protocol. The PCR product was purified using PureLink PCR micro purification kit (Invitrogen, Q32850). The library was then used to build clusters on the Illumina flow cell and analysis was done using Illumina Hiseq 2000 platform (Illumina, San Diego, CA, USA) at McMaster University to a target depth of 6M reads per sample.

Polysome-sequencing (Polysome-seq) library preparation

$20x10^6$ cells incubated for 10min with 0.1mg of cycloheximide (Sigma)/mL of BTIC media and incubated for 5min at 37°C and 5% CO_2. Cells were then washed and harvested in ice cold PBS and flash frozen in liquid nitrogen and stored until ready to lyse pellets. Polysome buffer (20mM Tris HCl pH7.4, 150mM NaCl, 5mM MgCl2, 1mM DTT) and lysis buffer (Polysome Buffer with 1% Triton-X + Protease Inhibitors + RNase inhibitors + cycloheximide (100ug/ml) were prepared fresh and kept on ice. Pellets were lysed in 400uL of Lysis Buffer per $10x10^6$ cells. Lysate is then incubated on ice for 30 min then centrifuged at 15,000 rpm at 4°C for 2 minutes to remove debris. The supernatant is collected and loaded onto a 10-50% sucrose gradient in Polysome Buffer with freshly added 1X Halt™ protease inhibitor (Thermo), RNAse inhibitors and 100ug/mL of cycloheximide (Sigma). 100-200uL of cell lysate is loaded to the top of each sample column and ultracentrifuged at 35,000rpm at 4°C for 3 hours. The total RNA is then extracted from the remaining lysate using Trizol LS (Invitrogen). The polysome fractions are collected and pooled for polyribosome associated RNA extraction using Trizol LS. Extracted polysome associated RNA is resuspended in 50uL of sterile ultrapure water.

1ug of total RNA and an equivalent volume of RNA from the polysome fraction were used for library preparation as per Illumina standard RNA-seq protocol. An equal amount of Spike ins RNA Variant Controls (SIRVs, Lexogen SIRV-Set3 and ERCC) were added to each sample prior to library preparation to normalized for sequencing bias and to determine the threshold for measurable statistics. The library was then used to build clusters on the Illumina flow cell and analysis was done using Illumina Hiseq 4000 platform (Illumina, San Diego, CA, USA) at the UCSD Institute for Genomic Medicine to a target depth of 30M reads per sample.

Tandem Mass Tag Mass Spectrophotometry (TMT-MS)

Approximately 100ug total protein was extracted from each of the six samples (replicates of SU_MB002 control -scramble cells, sh*MSI1*-1 and sh*MSI1*-2 cells) using 8M urea and 100 mM ammonium bicarbonate. The protein samples were reduced, alkylated, and digested by trypsin (Promega) overnight at 37 °C. The resulting peptides were desalted with 10 mg SOLA C18 Plates (Thermo Scientific), dried, labeled with 6-plex Tandem Mass Tag reagents (Thermo Scientific), before being pooled together. 60 ug of the pooled sample was separated into 38 fractions by reverse phase liquid chromatography (RPLC) at pH=10 using a Thermo Acclaim PA2 C18 column (300 µm x 15cm bed volume). Half of each fraction was then loaded onto a home-made trap column (200 µm x 5 cm bed volume) packed with POROS 10R2 10µm resin (Applied Biosystems), followed by a home-made analytical column (50 µm x 50cm bed volume) packed with Reprosil-Pur 120 C18-AQ 5 µm particles (Dr. Maisch). LC-MS experiments were performed on a Thermo Fisher UltiMate^{TM} 3000 RSLCNano UPLC system that ran a 3hr gradient at 70nl/min, coupled to a Thermo QExactive HF quadrupole-Orbitrap mass spectrometer. A parent ion scan was performed using a resolving power of 120,000 and then up to the 20 most intense peaks were selected for MS/MS (minimum ion count of 1000 for activation), using higher energy collision induced dissociation (HCD) fragmentation. Dynamic exclusion was activated such that MS/MS of the same m/z (within a range of 10ppm; exclusion list size=500) detected twice within 5s were excluded from analysis for 30s.

QUANTIFICATION AND STATISTICAL ANALYSIS

Kaplan-Meier survival plots

Kaplan-Meier plots were generated for the reposited Cavalli *et al* dataset (using the author's annotations for Msi1^{high} and Msi1^{low}), PDX and MP data with Prism8 (GraphPad^{TM}). Log rank Hazard ratio was and both the Both Log rank (Mantel-Cox) and the Gehan-Breslow-Wilcoxan test (i.e., gives more weight to deaths at early time points) p-value are reported.

eCLIP data processing

eCLIP reads were processed and QC was performed according to the ENCODE data processing protocol for eCLIP reads as previously described (Van Nostrand et al., 2016). First, reads were demultiplexed according to their inline barcodes (MB002_Msi1: A01, B06; NSC201cb_Msi1: X2A, X2B) using a custom script, which also modifies each read name to include the read's unique molecular identifier (UMI) (demux.py). Next, reads were trimmed using cutadapt (v1.14) and filtered of any read mapped to RepBase (v18.05) sequences using STAR (v2.4.0j). Surviving reads were then mapped, again with STAR, to hg19 assembly to obtain genome alignments. PCR duplicate removal was then performed with a custom script based on UMI sequences placed inside each read name (barcodecollapsepe.py). De-duplicated mapped BAM files from each barcode were then combined (samtools merge v1.6), forming a single BAM file for each single IP and size-matched INPUT dataset. Read2 for each IP merged BAM file were used to call enriched peak clusters with Clipper (v1.2.1). These clusters were then normalized against size-matched INPUT reads and neighboring/overlapping clusters merged. Regions passing a -log10(p) significance of at least 3 and a log2(fold change) cutoff of 3 were deemed as significantly Msi1-bound for each replicate. The demultiplex script can be found at:

https://github.com/YeoLab/eclipdemux. The pipeline definitions and barcode collapse script can be found at: https://github.com/YeoLab/eclip

To obtain reproducible regions between two replicates, we used the modified IDR pipeline as previously described (van Nostrand, 2018). Using the outputs from the processing pipeline, input normalized peaks were ranked according to information content (p_i*log2(p_i/q_i). These ranked peaks were passed to IDR (v2.0.2) to determine regions of reproducibility. Full definitions for each tool and workflow can be found at: https://github.com/YeoLab/merge_peaks.

eCLIP motif analysis

Motif analysis was performed using HOMER (v4.9.1) wrapped inside a custom script (analyze_motifs.py found here: https://github.com/YeoLab/clip_analysis_legacy). The methodology was first described by Lovci *et al* (Lovci et al., 2013); briefly peaks were assigned to their corresponding regions of binding (CDS, 3'UTR, 5'UTR, proximal and distal intron +/- 500bp of an exon), then compared against a randomized background (random assignments of peak coordinates across each corresponding region).

eCLIP region-based fold-enrichment analyses

Region-based fold-enrichment was calculated as previously described (Van Nostrand et al., 2016). Briefly, mapped reads were counted along all transcripts in Gencode v19 ('comprehensive'). Reads were assigned to all transcripts annotated in Gencode v19. For reads overlapping >1 annotated region, each read was assigned to a single region with the following descending priority order: CDS, 5'UTR, 3'UTR. For each gene, reads were summed up across each region to calculate final region counts. A minimum of 10 observed reads were required for a gene to be considered in region-based fold-enrichment analyses. The MSI1 eCLIP data in SU_MB002 and NSC201cb has been deposited in the

NCBI Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo). The accession number of the eCLIP data reported in this paper is GEO: GSE126263.

RNA-sequencing data processing and analysis

An average of 5.5M reads from each sample passed quality filtering. Filtered reads were mapped to the human genome (hg19) using the STAR short-read aligner (v.2.4.2a) with the following command: STAR --genomeDir /path/to/GRCh37 --readFilesIn <file1.fastq.gz> <file2.fastq.gz> --readFilesCommand zcat --runThreadN 8 --outSAMstrandField intronMotif --outSAMtype BAM SortedByCoordinate --quantMode GeneCounts --sjdbGTFfile /path/to/gtf. The gencode.v19.annotation.gtf from the GENCODE database and the primary assembly of GRCh37 was used. Approximately 92% of the filtered reads mapped uniquely, and the read counts from each sample were merged into a single matrix using R. The raw and processed data has been deposited to the GEO database (accession: GSE126337).

The merged read count matrix was used to compute differential expression using the Bioconductor package limma (v3.38.3) as follows. First, transcripts were filtered using filterByExpr(min.count = 10, min.total.count = 15) (edgeR, v3.25.3) and normalized using calcNormFactors(method = "TMM"). A counts per million matrix was created from the normalized count matrix. Differential gene expression was conducted using the lmFit function and ranked using treat. Significant genes were identified using an FDR < 0.05 and absolute log2(fold-change) > 1.

Polysome-sequencing processing and analysis

An average of 30M reads from each sample passed quality filtering. Filtered reads were mapped to the human genome (hg19) using the START short-read aligner (v.2.4.2.a). The gencode.v19.annotation.gtf from the GENCODE database and the primary assembly of

GRCh37 was used. Approximately 92% of the filtered reads mapped uniquely, and the read counts from each sample were merged into a single matrix using R. The merged read count matrix was used to compute differential expression using the Bioconductor package limma (v3.38.3) as follows. First, transcripts were filtered using filterByExpr(min.count = 10, min.total.count = 15) (edgeR, v3.25.3) and normalized using calcNormFactors(method = "TMM"). A counts per million matrix was created from the normalized count matrix. Significant genes were identified using an FDR < 0.05 and absolute log2(fold-change) > 1. The raw and processed data will be deposited to the GEO database.

Proteomic data processing and analysis

LC-MS data generated was analyzed against a UniProt human protein database (42,173 entries) for protein identification and quantification by MaxQuant software (v.1.6.5) From 2,379,345 MS/MS spectra acquired in all 38 fractions, 136,833 unique peptide groups (with Peptide FDR<0.01) and 8,547 proteins (Protein FDR < 0.01) were identified and quantified (Cox and Mann, 2008). The Significant B values were calculated using the PERSEUS (v.1.6.5) software. Significance B value preset with a FDR<0.01 was used to identify proteins that are significantly differentially abundant and used for downstream integrative analysis. All raw data have been deposited in the ProteomeXchange Consortium via Proteomics Identification (PRIDE) (Vizcaino et al., 2016). The accession number of the proteomics data reported in this paper is PRIDE: PXD012432.

Gene ontology (GO) and Gene set enrichment analysis (GSEA)

GSEA (Subramanian et al., 2005) was applied using a combination of MSigDB C2 curated gene sets (v6.2), C5 Gene Ontology gene sets (v6.2), and C6 oncogenic signatures (v6.2) (Chen et al., 2013). GSEA was run using the fgsea Bioconductor

package (v1.2.1). Additionally, enrichment analysis was performed on sets of significant genes/proteins using the EnrichR database (Kuleshov et al., 2016).

Protein Set Enrichment Analysis (PSEA)

PSEA-QUANT (Lavallee-Adam et al., 2014), a protein enrichment analysis algorithm was used for our label-based mass-spectrometry-based quantitative proteomics to identify protein sets from Gene Ontology and Molecular Signatures databases that are statistically enriched with abundant proteins. Abundance ratios were used, with 10,000 samplings for statistical significance assessment, annotated using the Molecular Signature Database with the assumption that protein abundance dependence in the dataset (coefficient of variation = 0.5). Literature bias assumed protein annotation bias.

Triprotocol data integration

Significant genes were selected from each of the four large-scale datasets as follows. Genes bound by MSI1 were filtered from the eCLIP binding data using an FDR<0.001 and log2FC (shMSI1/control)> 3 (8-fold). Genes differentially expressed upon *MSI1* knock-down were selected as having an FDR<0.05 and an absolute FC >2, adjusted *p*-value<0.1. Significantly differentially expressed polysome associated mRNA was selected using an absolute FC >1.5 and adjusted *p*-value<0.1. Finally, significant proteins were selected as having a FDR<0.01, absolute FC >1.5 and adjusted *p*-value<0.1 in the sh*MSI1* vs control.

Pathway Analysis, network diagrams and data visualization

Pathway analysis for the comparison between eCLIP datasets in SU_MB002 and NSC201 cell lines was conducted using g:Profiler (Reimand et al., 2007). Genes were ranked by decreasing fold change. Gene sets from Reactome (v64, released 2018-10-02)

and Gene Ontology databases (version Ensembl v93/ Ensembl Genomes v40, released 2018-08-03) were included. Gene sets were limited to between 5 and 500 genes and pathways were filtered for a statistical threshold of $p < 0.05$. Pathway analysis for integration of mRNA, protein and eCLIP datasets were conducted using ActivePathways (Paczkowska et al., 2018). The corrected *p*-values from the mRNA differential expression analysis and the eCLIP annotation pipeline, and the uncorrected p-values from protein differential expression analysis were used as input to ActivePathways using the same gmt, pathway size and significance threshold as described above for g:Profiler. P-values were merged using Brown's method and corrected using Holm's method. Visualization was done in Cytoscape (v.3.6.0) and Inkscape. Data visualization was done using Boutroslab.plotting.general (v.5.9.2) (P'ng et al., 2019) and ggplot2 (v3.1.0) (Wickham, 2016). Data for the ribbon plot for the network diagram was extracted from the Reactome Functional Interaction Database (Wu et al., 2010). Data was visualized using Circos (Krzywinski et al., 2009).

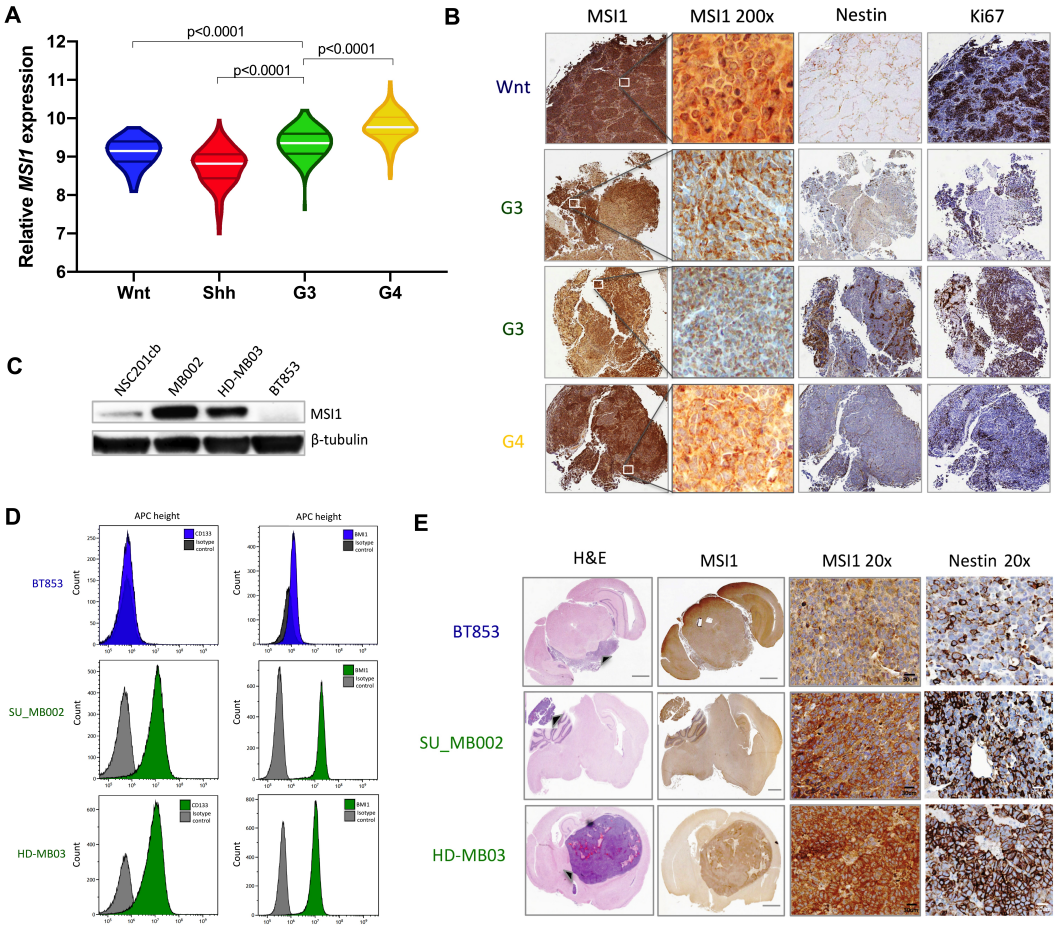


Figure 4.1: *MSI1* is overexpressed at in G3 MB. A. Cavalli *et al* transcriptomic data showing significant upregulation of *MSI1* in G3 MB as compared to Wnt and Shh MB ($p<0.0001$), **B.** Primary MB patient samples showing nuclear staining in G3.4 tumours with co-localization of Nestin and Ki67, **C.** Western immunoblot of MB cell lines demonstrating high protein expression of MSI1 in G3 MB lines (SU_MB002, HD-MB03) *vs* a cerebellar neural stem line (NSC201cb) and non-G3 MB lines (BT853), **D.** Primary patient-derived orthotopic xenograft immunostained for MSI1 and Nestin demonstrated greater protein expression of MSI1 in G3 MB (SU_MB002, HD-MB03). Bar represents 1000um unless otherwise indicated, **E.** Flow cytometric analysis demonstrating higher CD133 and BMI1 in G3 MB (SU_MB002 and HD-MB03) as compared to non-G3 MB (BT853) samples.

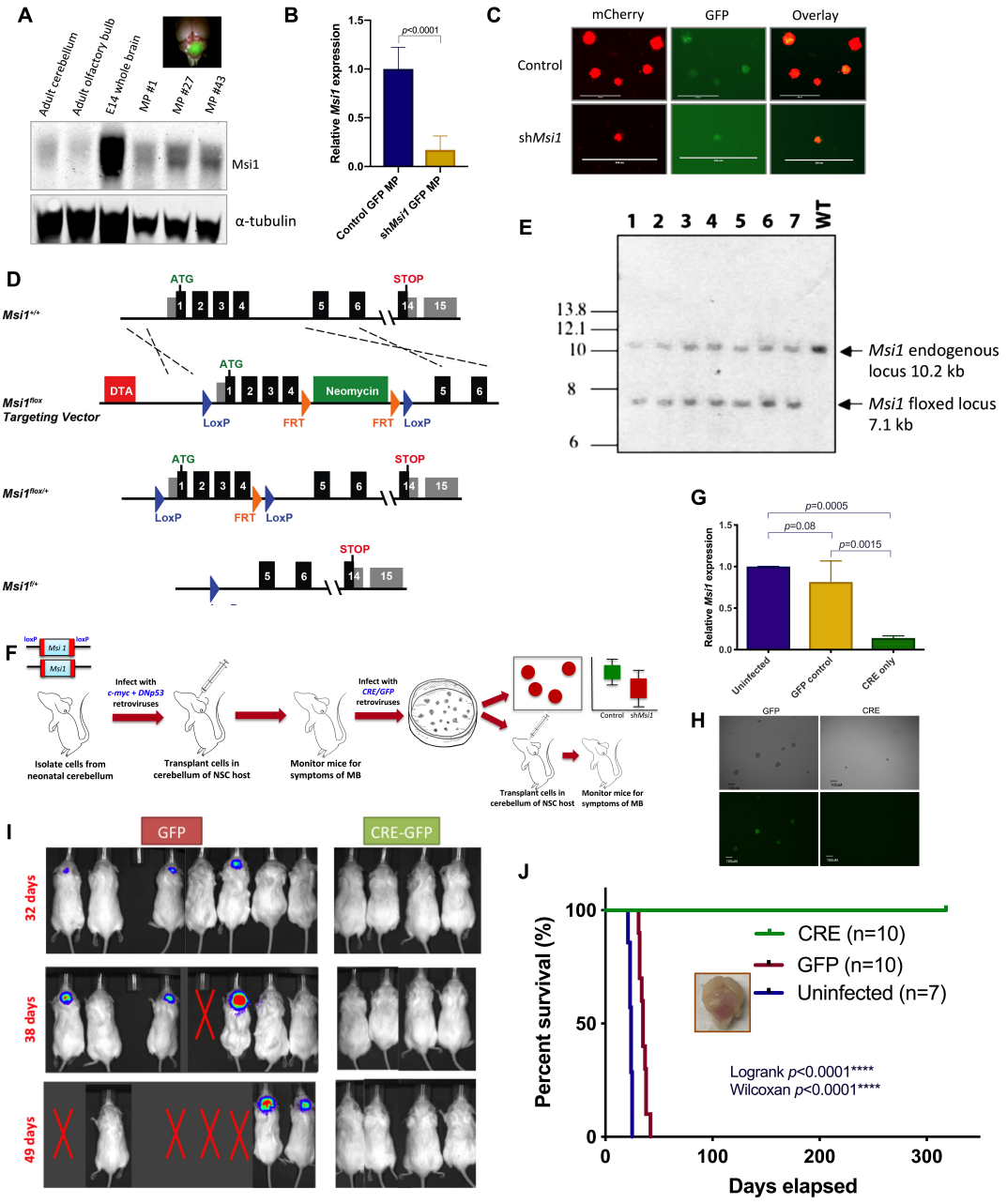


Figure 4.2: Cre mediated *Msi1*$^{fl/fl}$ KO in MP cells impair the key stem cell property of self-renewal and abrogates the MP's capacity for tumour formation. A. Western immunoblot showing greater MSI1 abundance in MP tumours as compared to adult mouse cerebellum and olfactory bulb but not embryonic whole brain. **B.** RTqPCR showing a dramatic reduction in *Msi1* expression after sh*Msi1* inhibition (p<0.0001), **C.** Secondary sphere formation assay showing obliteration of stem cell function in sh*Msi1* MP cells. **D.** Schematic of targeting strategy for the generation of *Msi1* conditional knockout mice (*Msi1*$^{fl/fl}$), **E.** Southern blot analysis of heterozygous mice carrying the floxed allele after genomic DNA digestion. **F.** Schematic of MP model adapted from Pei *et al*, 2012. **G.** RT-qPCR *Msi1*$^{fl/fl}$ KO validation of *Msi1*$^{fl/fl}$ KO vs control showing Cre-mediated excision lowers Msi1 mRNA levels compared to control GFP cells (p=0.0015), **H.** Secondary sphere formation after *Msi1*$^{fl/fl}$ KO of MP tumour cells showing greater neurosphere forming capacity of control GFP Msi1 MP than Cre-mediated *Msi1* KO MP cells. **I.** Tumour

burden analysis of showing no luminescence in Cre mediated *Msi1 KO* MP tumours indicating impaired tumour formation compared to GFP control mice, **J.** Kaplan-Meier survival analysis of Cre-mediated *Msi1* KO mice (n=10), control GFP mice (n=10) and untransduced MP mice (n=7) illustrating a clear survival benefit for the Cre-mediated *Msi1* KO cohort ($p<0.001$).

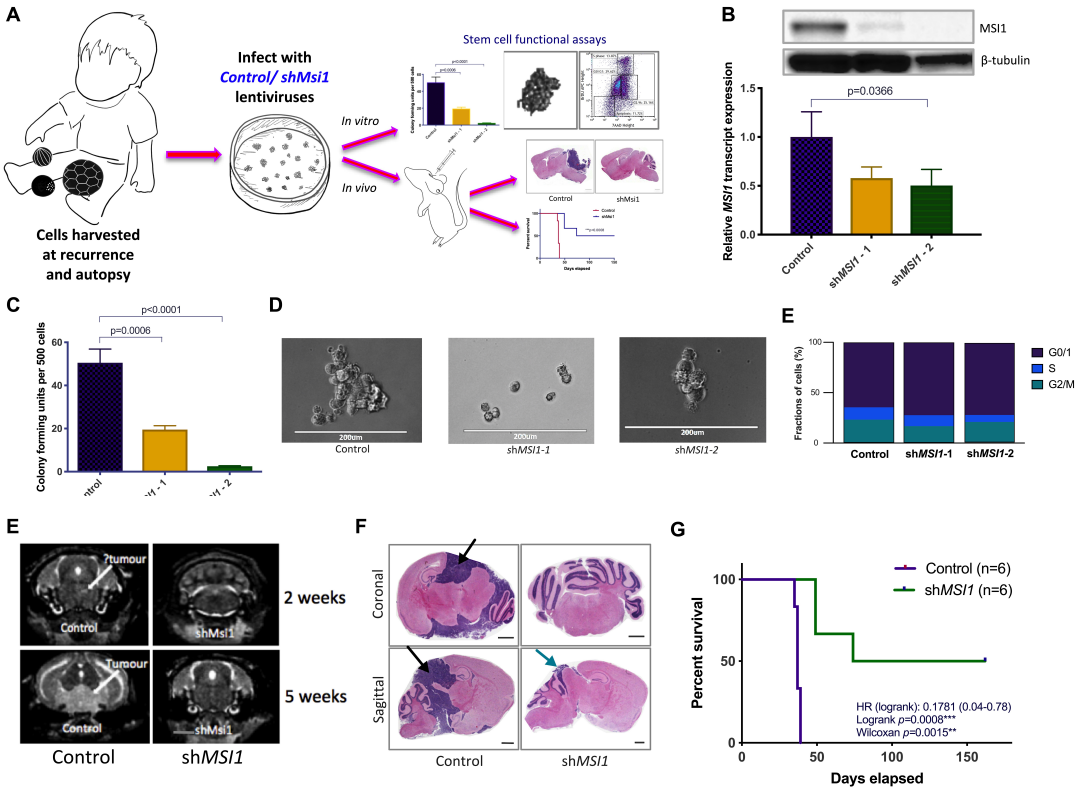


Figure 4.3: sh*MSI1* inhibition in G3 MB impairs key functional stem cell properties and tumour formation. A. Schematic of primary G3 MB cell line *in vitro* and *in vivo* sh*MSI1* experiments. **B.** Western immunoblot and RTqPCR sh*MSI1* KD validation of SU_MB002, **C-D.** Secondary colony formation after sh*MSI1* KD of SU_MB002 demonstrating significant impairment of stem cell properties in sh*MSI1*-2. **E.** Cell cycle analysis showing increased cells in G0/1 and higher ratios of G2/M to S in sh*MSI1* constructs, **F.** 24 NOD-SCID mice were engrafted with SU_MB002 at a density of $2.5x10^4$ cells with T2 MRI brain sequence mages of control vs sh*MSI1* KD SU_MB002 orthotopic xenografted mice showing minimal difference at 2 weeks, but subsequent large tumour growth (purple arrow) in control vs sh*MSI1* mouse at 5 weeks, **F.** H&E stain of control vs sh*MSI1* KD SU_MB002 transplanted mice in the tumour burden arm demonstrating large tumours in the control vs sh*MSI1* harvested brains (bars represent 100um), **G.** Kaplan-Meier survival curve of control vs sh*MSI1* KD SU_MB002 transplanted mice ($p<0.0008$).

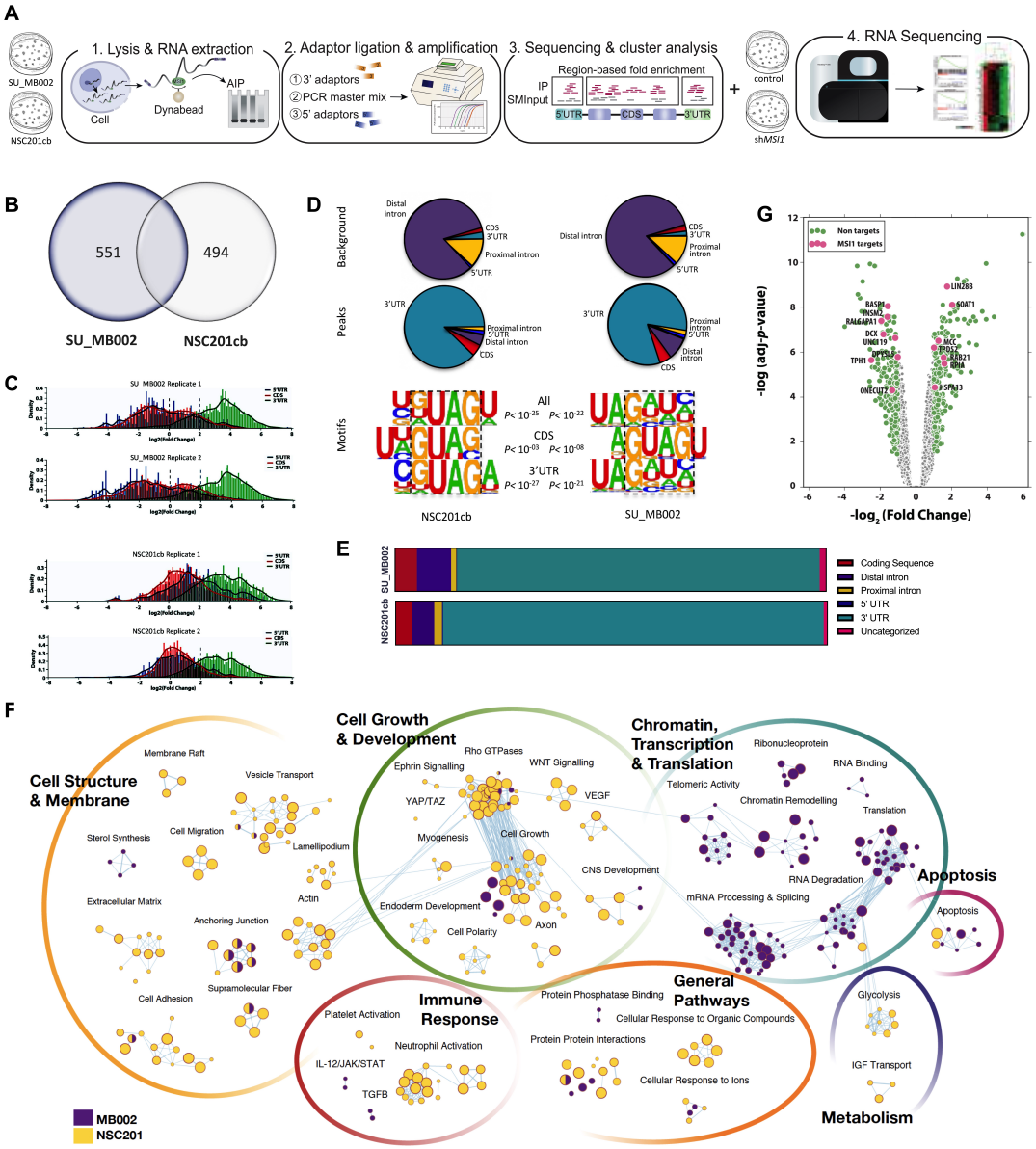


Figure 4.4: eCLIP analysis reveals key MSI1 bound transcripts involved in processes associated with Chromatin, Transcription and Translation. A. Schematic of experimental plan to identify Msi1 binding in SU_MB002 and NSC201cb and examination of changes to steady state transcriptome with RNA-seq, **B.** Venn diagram showing moderate overlap of binding genes between SU_MB002 and NSC201cb (>3logFC), **C.** eCLIP peak calling identifies 3'UTR (green peaks) as highest confidence binding sites for both SU_MB002 and NSC201cb (>3logFC), **D.** Motif analysis identifying MSI1 consensus binding sequence as G(UAG) with highest confidence in the 3'UTR ($p<10^{-48}$) in SU_MB002, **E.** Bar graph showing proportion of binding regions in NSC201cb and SU_MB002, **F.** Pathways analysis identifying key processes that are aberrant in SU_MB002 as compared to NSC201cb (>2logFC), **G.** Volcano plot showing genes that are bound by MSI1 are modestly differentially expressed (mean logFC 0.19±0.52, Range=-3.69-2.10).

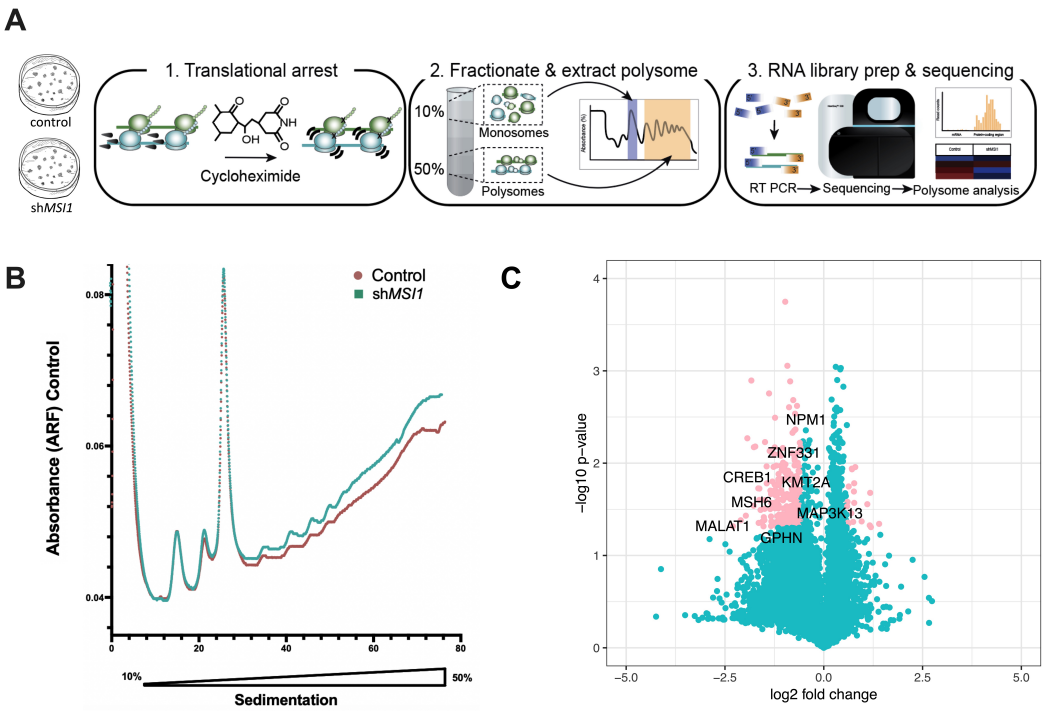


Figure 4.5: Polysome profiling and sequencing suggests MSI1 is downregulating the translation of known cancer associated genes. A. Schematic of experimental plan to investigate the ribosome associated fraction of mRNA, **B.** Polysome profile of control vs sh*MSI1* SU_MB002 shows an slight increase in mRNA within the polysome associated fraction in sh*MSI1* inhibited samples, **C,** Volcano plot of FC after shMSI1 KD showing the main effect of downregulation of cancer associated mRNA within the polysome fraction.

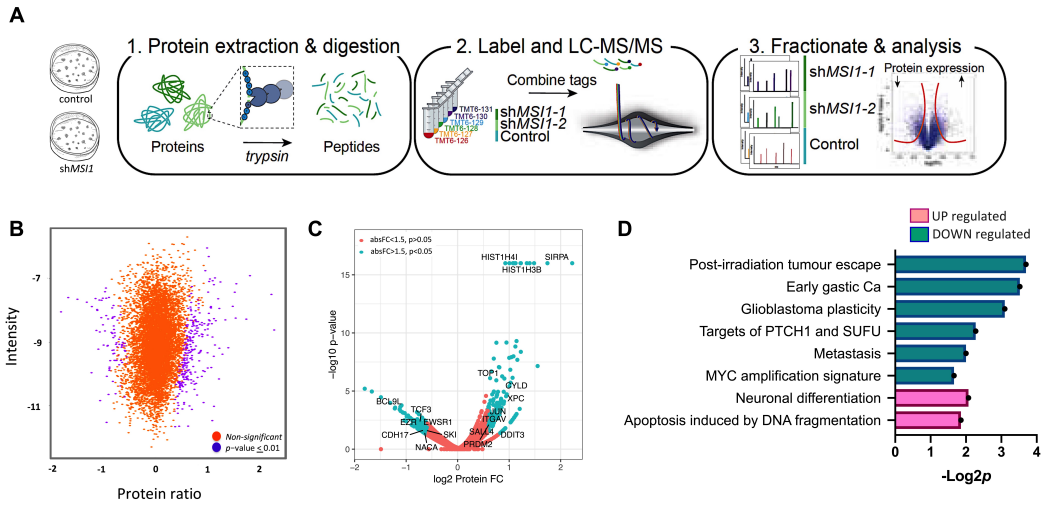


Figure 4.6: Whole shot gun proteomic analysis identifies downregulation of proteins associated with cancer aggressiveness **A:** Schematic of experimental plan to quantitatively investigate the proteomic landscape after sh*MSI1* inhibition, **B.** MaxQUANT and PERSEUS analysis of TMT-MS data showing 350 significantly differentially abundant proteins in the sh*MSI1* samples compared to the control (Significance B <0.01, FDR<0.01), **C.** Volcano plot with annotation of key genes that are up and down regulated at the proteomic level with ~50% inhibition of the MSI1 protein, **D.** PSEA analysis identifying downregulation of proteins associated with cancer aggressiveness and upregulation of genes associated with neuronal differentiation and DNA fragmentation.

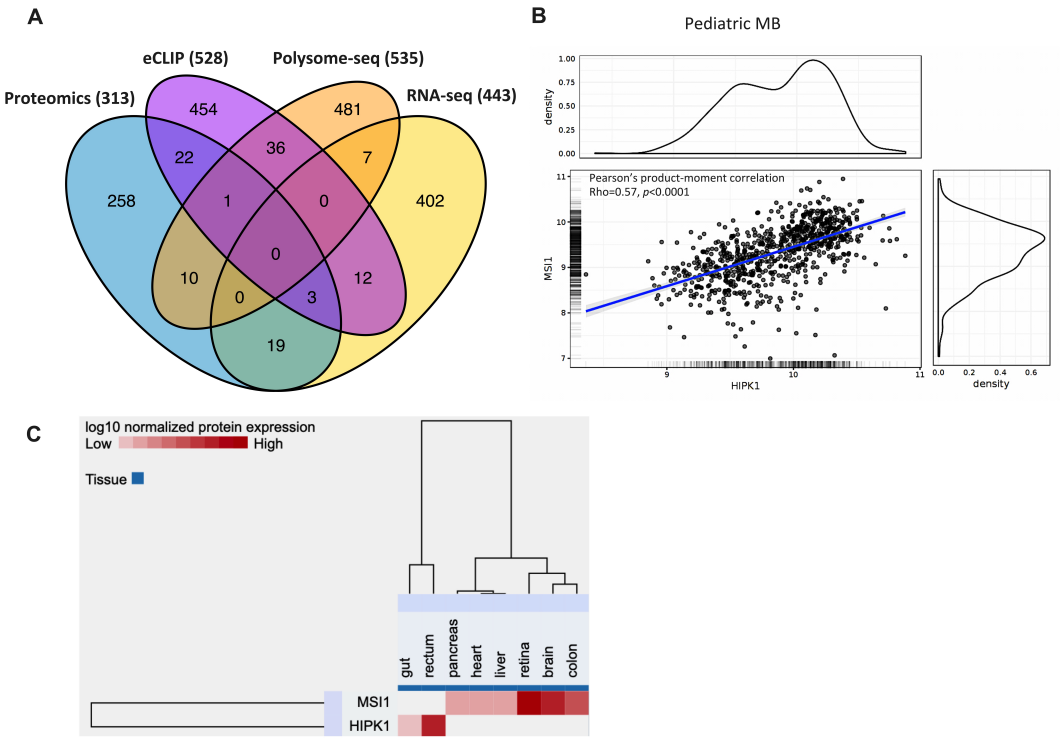
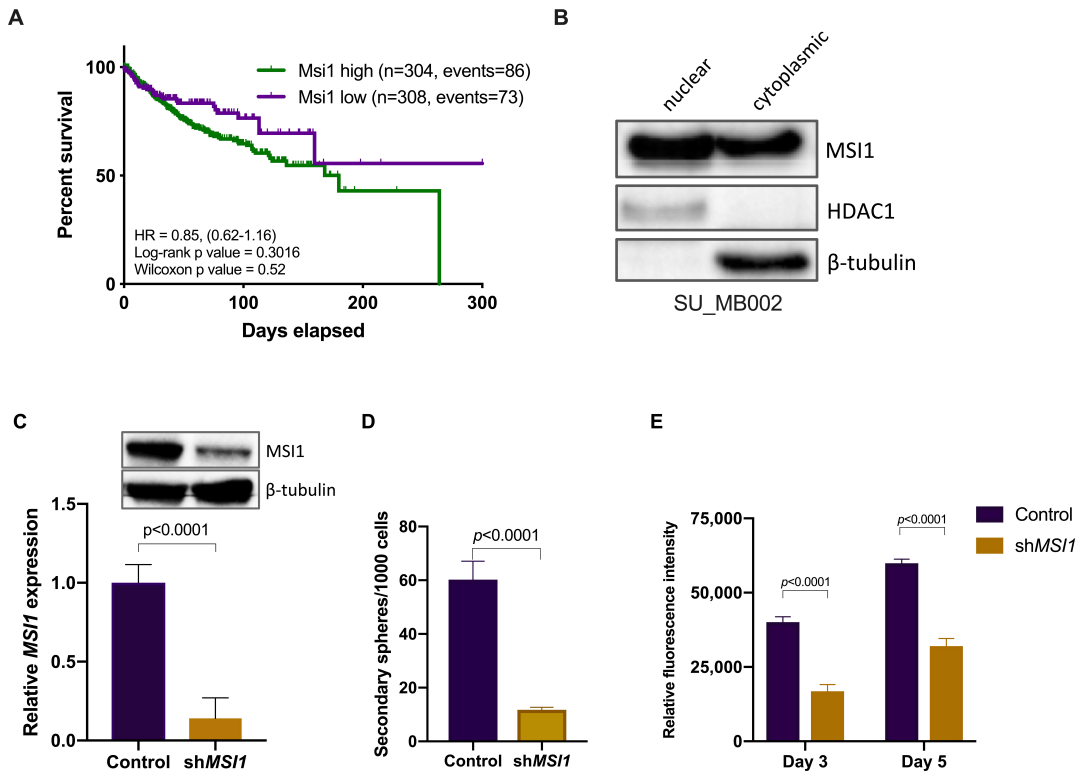
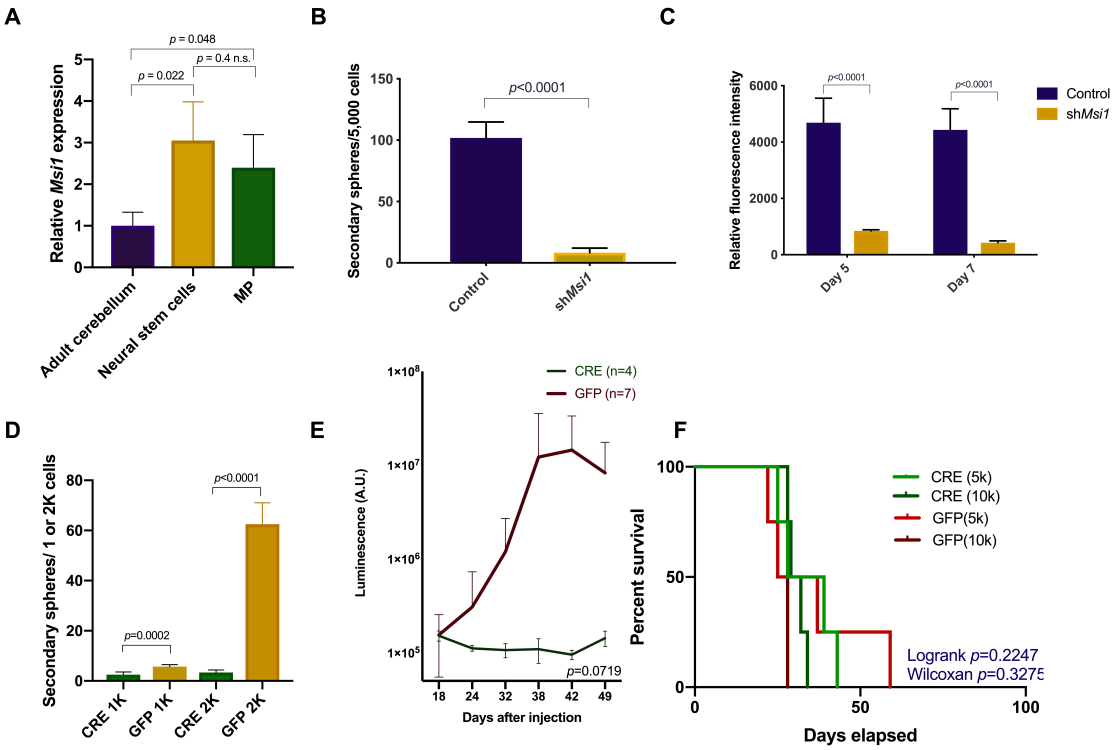


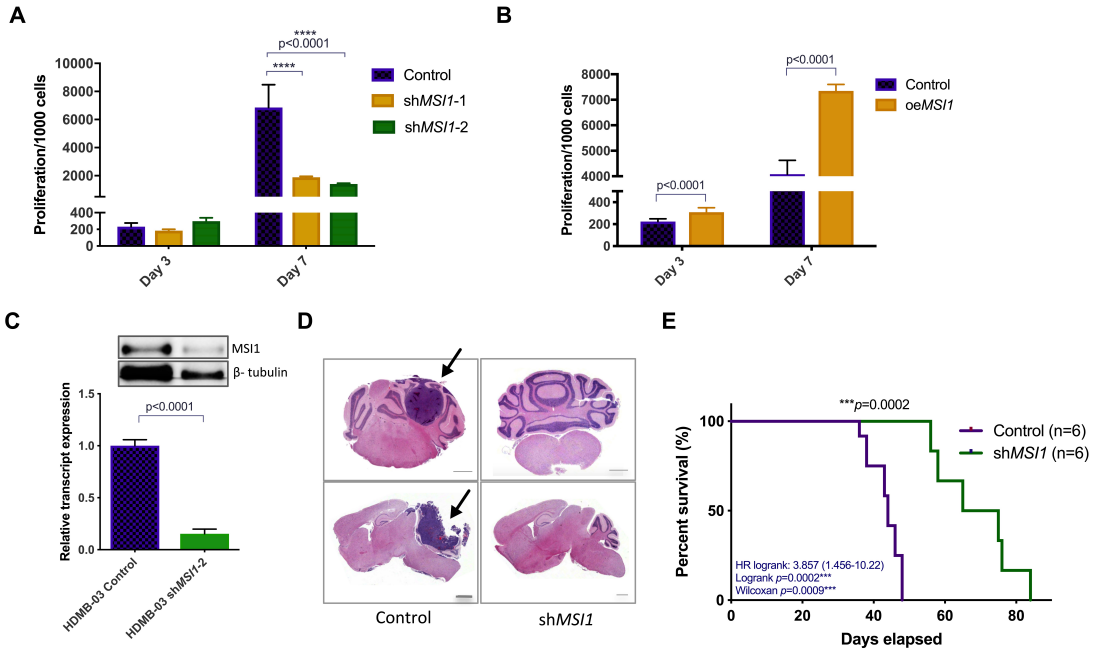
Figure 4.7: Integrative analysis identifies HIPK1 as a potential downstream target of MSI1 in G3MB for validation and therapeutic drug discovery **A.** Integrative analysis of the significantly differentially expressed genes/proteins after sh*MSI1* inhibition identifies one gene, *HIPK1*, that is MSI1 eCLIP bound and significantly down regulated at the translatomic and proteomic levels, B. Interrogation of the Cavalli et al transcriptomic database (Cavalli et al, 2017) shows *HIPK1* and *MSI1* have a positive correlation (Rho=0.57, p<0.0001), **C.** Heatmap generated from proteomicsDB.org illustrating HIPK1's limited expression as compared to MSI1.



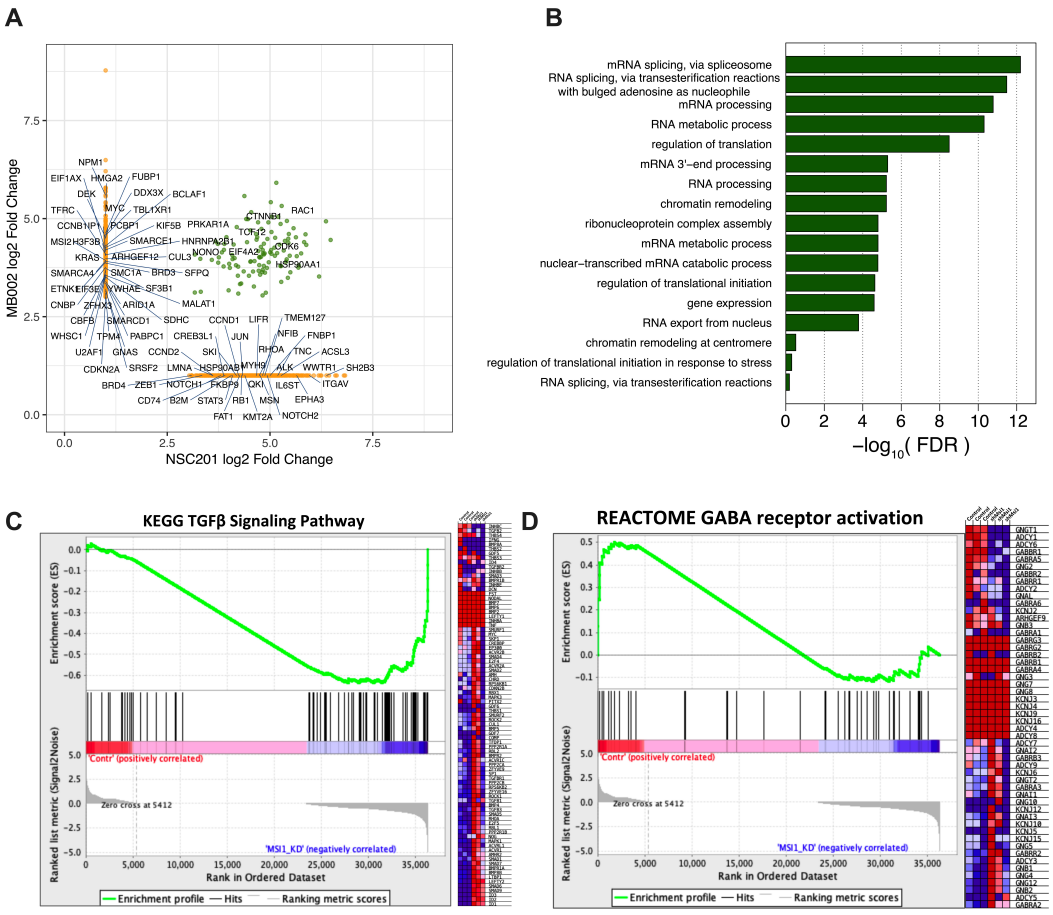
Supplementary Figure 4.1: *MSI1* is overexpressed at in G3 MB. A: Survival data from the Cavalli dataset for *MSI1* ^{high} (n=304) transcript expression of primary pediatric samples illustrating a non-significant increase in survival when compared to and *MSI1* ^{low} (n=308) samples (HR: 0,85 (0.62-116; Logrank *p*=0.3016), **B:** Western immunoblot of SU_MB002 nuclear and cytoplasmic fractions showing substantial nuclear staining of MSI1 in addition to the expected cytoplasmic fraction, **C:** MSI1 expression at the transcript and protein levels after sh*MSI1* inhibition in NSC, **D-E:** Significant reduction in secondary neurosphere formation and proliferation observed after sh*MSI1* inhibition (p<0.0001).



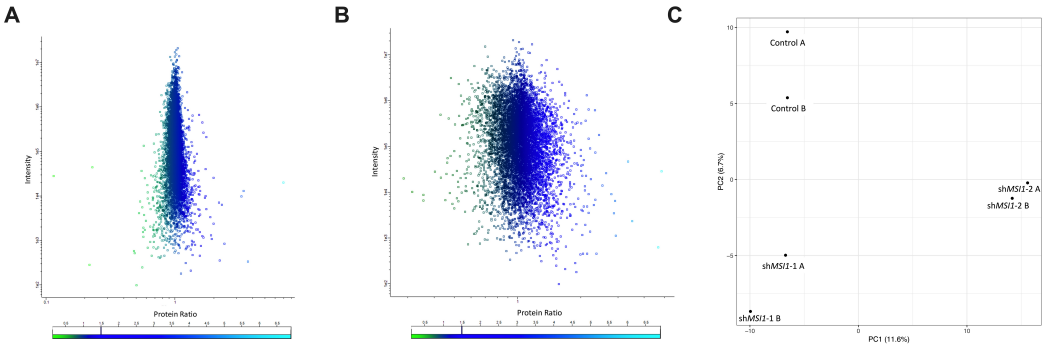
Supplementary Figure 4.2: Cre mediated *Msi1*$^{fl/fl}$ KO in MP cells impair the key stem cell property of self- renewal and abrogates the MP's capacity for tumour formation. A. RTqPCR showing significant overexpression of *Msi1* in MP as compared to the adult mouse cerebellum, **B.** Quantification of secondary sphere forming assay at day 7 (p<0.00001), **C.** Proliferation assay at day 5 and 7 showing significant reduction in proliferation in sh*Msi1* inhibited MP cells (p<0.0001), **D.** Quantification of secondary sphere formation between CRE and GFP infected MP cells (p<0.0001), **E.** Luminescence graph of Cre vs GFP mice showing reduced Luciferin signal in the GFP control mice compared to the Cre-Msi1-excised mice (Paired t-test, p=0.0719), **F.** Kaplan-Meier survival plot showing no survival benefit in Cre-treated cells and therefore Cre toxicity does not explain survival benefit (p=0.2247).



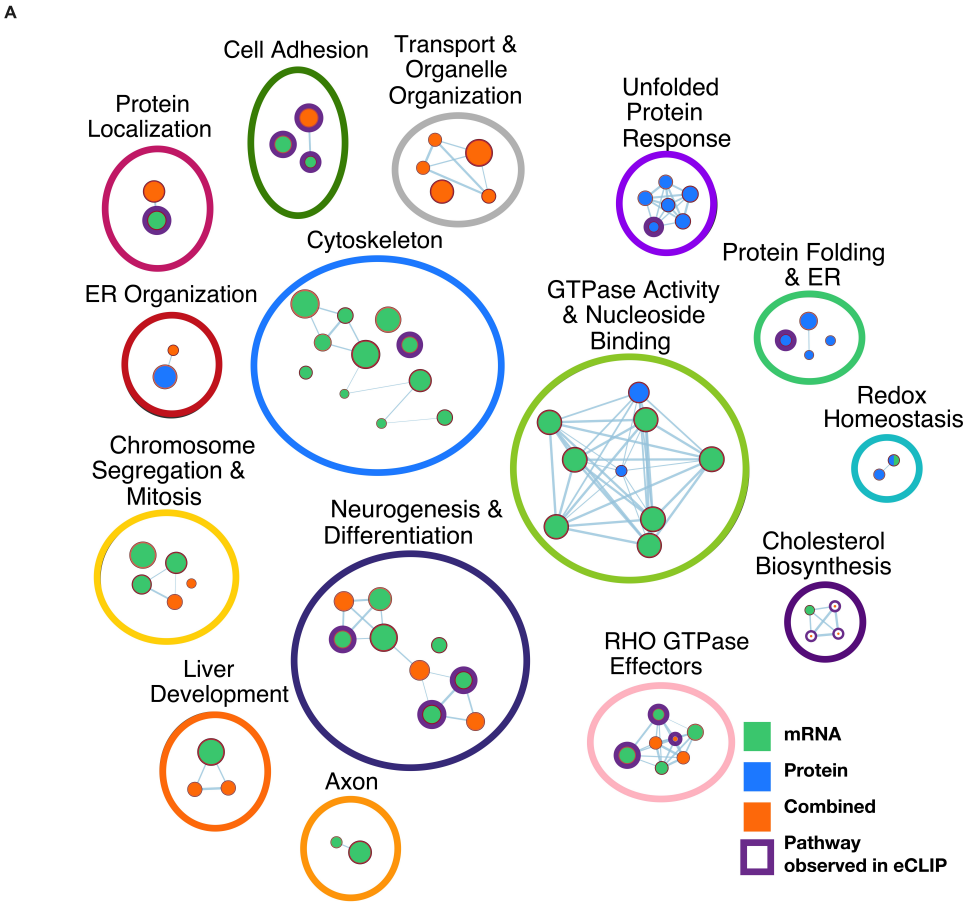
Supplementary Figure 4.3: sh*MSI1* inhibition in G3 MB impairs key functional stem cell properties and tumour formation. A. Proliferation assay at day 3 and 7 analysis after sh*MSI1* KD of SU_MB002 demonstrating significant impairment of proliferation in sh*MSI1*-2 (p<0.0001), **B.** Over-expression of *MSI1* in an already overexpressed system pushes the cell to further increase proliferation capacity (p<0.0001). **C:** *In vivo* experimentation repeated in a second G3 MB patient derived cell line, HD-MB03 showing significant *MSI1* KD at both the transcript and protein levels after sh*MSI1* inhibition (p<0.0001), and **D.** large tumours in the control arms as compared to the sh*MSI1* brains (bars represent 100uM), **E.** sh*MSI1* KD engrafted mice conferring a significant survival benefit *(p*=0.0002).



Supplementary Figure 4.4: eCLIP analysis reveals key MSI1 bound transcripts involved in TGF-β pathway, GABA activation and processes associated with Chromatin, Transcription and Translation A. Analysis of the eCLIP binding targets identifying key cancer associated genes shared by both SU_MB002 and NSC201cb, **B.** EnrichR analysis identifying dysregulation of RNA processes within the eCLIP bound differentially expressed genes in RNA-seq (FC>1.5, p<0.05), **C-D.** GSEA showing dysregulation of genes known to be associated with G3 MB, the TGF-β and GABA activation pathways after sh*MSI1* inhibition.



Supplementary Figure 4.5: MaxQUANT and PERSEUS analysis suggests an increase in steady state protein after sh*MSI1* inhibition. A-B. MaxQUANT and PERSEUS analysis showing extensive changes in the proteome after sh*MSI1* inhibition with a rightward shift of protein abundance (FDR<0.01, p<0.01), **C.** PCA analysis illustrating hairpin with greatest MSI1 KD separates further from control samples.



Supplementary Figure 4.6: Integrative pathways analysis of affected processes associated with differentially expressed genes in the transcriptomic, translatomic and proteomic landscape after sh*MSI1* inhibition in G3 MB A. Integrative Pathways analysis illustrating the predominant pathways regulated by MSI1 either directly (outlined in purple) or indirectly (all other) in multiple processes with direct regulation of cell adhesion, neurogenesis and differentiation.

Supplementary Tables

Supplementary Table 4.1A: Top 50 reproducible peaks per IDR identified and quantified in SU_MB002

Chromosome	$Log_{20}p$	log_2FC	Strand	Gene ID	Gene name	Binding location
chr12	130.9069785	8.777033777	+	ENSG00000111678.6	C12orf57	5utr
chr12	85.29091126	7.902410526	-	ENSG00000060982.10	BCAT1	3utr
chr5	81.07294875	6.837477208	+	ENSG00000015479.13	MATR3	3utr
chr20	59.91804093	5.840674085	-	ENSG00000200494.1	Y_RNA	noncoding exon
chr16	52.57756841	7.25922924	-	ENSG00000170540.10	ARL6IP1	3utr
chr7	43.64426868	6.247806948	+	ENSG00000242265.1	PEG10	3utr
chr1	42.91109846	5.029559045	-	ENSG00000173726.6	TOMM20	3utr
chrM	42.58447578	5.321081504	-	ENSG00000210196.2	MT-TP	noncoding exon
chr12	40.45456608	6.985242013	+	ENSG00000179195.11	ZNF664	3utr
chr4	38.53043893	5.423984001	-	ENSG00000118816.5	CCNI	3utr
chr11	37.7841246	3.556087385	+	ENSG00000251562.3	MALAT1	noncoding exon
chr21	37.58937087	6.769725775	+	ENSG00000198743.5|| ENSG00000272962.1	SLC5A3	3utr
chr12	37.534259	6.783979979	+	ENSG00000149948.9	HMGA2	3utr
chr12	36.87039582	4.937627988	+	ENSG00000135486.13	HNRNPA1	3utr
chr15	35.15934931	6.779854277	-	ENSG00000140391.10	TSPAN3	3utr
chrX	34.96593011	5.784971341	+	ENSG00000102144.9	PGK1	3utr
chr16	34.59372627	5.786678363	+	ENSG00000091651.4	ORC6	3utr
chr15	33.25813174	4.433604362	-	ENSG00000247809.3	NR2F2-AS1	noncoding exon
chr12	33.05641151	4.857046447	+	ENSG00000008394.8	MGST1	3utr
chr3	32.55258876	5.055059351	+	ENSG00000156976.10	EIF4A2	3utr
chr17	32.18120884	6.646291422	-	ENSG00000136450.8	SRSF1	3utr
chr6	31.84807882	5.022863505	+	ENSG00000196230.8	TUBB	3utr
chrX	31.5259295	4.410543501	+	ENSG00000147140.11	NONO	3utr
chrX	29.91107758	5.51403978	+	ENSG00000131171.8	SH3BGRL	3utr
chr10	29.71255077	5.561559347	+	ENSG00000099194.5	SCD	3utr
chr11	29.59140392	4.938285787	-	ENSG00000110321.11	EIF4G2	3utr
chrM	28.99250656	6.49170087	+	ENSG00000210164.1	MT-TG	noncoding exon
chr2	27.53936702	5.455572826	+	ENSG00000176887.5	SOX11	3utr
chr11	26.79778259	4.871195681	-	ENSG00000152558.10	TMEM123	3utr
chr14	25.96881056	4.777721156	-	ENSG00000174373.11	RALGAPA1	3utr
chr15	25.84084233	6.339717622	-	ENSG00000140319.6	SRP14	3utr
chr16	25.65569263	6.375022115	+	ENSG00000230989.2	HSBP1	3utr
chrX	25.61407125	5.406349608	+	ENSG00000102144.9	PGK1	3utr
chr22	24.90027276	6.23451097	+	ENSG00000189060.4	H1F0	3utr
chr12	24.83009878	5.384932835	-	ENSG00000187109.9	NAP1L1	3utr
chr7	24.31613848	6.244869532	-	ENSG00000105810.5	CDK6	3utr
chr11	24.0579986	4.941801413	+	ENSG00000134333.9	LDHA	3utr
chr14	23.33098965	4.727164562	-	ENSG00000080824.14	HSP90AA1	3utr
chr2	23.29552509	6.115726965	+	ENSG00000170144.14	HNRNPA3	3utr
chr8	23.23874893	3.352849272	-	ENSG00000070756.9	PABPC1	CDS
chr14	22.97695342	4.21446932	+	ENSG00000100941.4	PNN	CDS
chr6	22.94225784	3.025370111	+	ENSG00000112306.7	RPS12	CDS
chr17	22.91814631	6.284106912	+	ENSG00000129657.10	SEC14L1	3utr
chr2	22.11627158	4.257377881	-	ENSG00000115758.8	ODC1	3utr
chr16	21.94445387	6.010600185	-	ENSG00000167005.9	NUDT21	3utr
chr22	21.89832131	4.292426182	-	ENSG00000184117.7	NIPSNAP1	3utr
chr10	20.96114468	5.994796425	+	ENSG00000148798.5	INA	3utr
chr5	20.57943831	4.990571912	+	ENSG00000131507.9	NDFIP1	3utr
chr18	20.56341885	5.887381237	+	ENSG00000101557.10	USP14	3utr

Supplementary Table 4.1B: Top 50 of reproducible peaks per IDR identified and quantified in NSC201cb

Chromosome	$\log_{10}p$	log2FC	Strand	Gene ID	Gene name	Binding location
chr13	217.7647656	8.017232101	+	ENSG00000136156.8	ITM2B	3utr
chr3	114.1054669	7.135879065	+	ENSG00000044524.6	EPHA3	3utr
chr21	112.7230695	7.091495696	-	ENSG00000142192.16	APP	3utr
chr16	92.69099682	7.539088307	-	ENSG00000170540.10	ARL6IP1	3utr
chr17	90.75969817	6.398451659	+	ENSG00000129657.10	SEC14L1	3utr, CDS
chrX	85.6133828	6.297131384	-	ENSG00000046653.10	GPM6B	3utr, CDS
chr3	85.25096298	6.689735973	+	ENSG00000044524.6	EPHA3	3utr
chr5	84.74867819	4.711242582	-	ENSG00000113140.6	SPARC	3utr
chr10	84.74438984	5.014445832	+	ENSG00000026025.9	VIM	3utr
chr5	84.17973284	4.621682212	+	ENSG00000127022.10	CANX	3utr
chr1	83.22261668	4.31953998	-	ENSG00000132688.10	NES	3utr
chr10	80.08848551	3.836734099	-	ENSG00000197746.9	PSAP	3utr
chr10	78.04203853	3.921922292	+	ENSG00000099194.5	SCD	3utr
chr4	76.29918028	6.849455618	-	ENSG00000150625.12	GPM6A	3utr
chr2	74.17232709	3.208316853	-	ENSG00000115461.4	IGFBP5	3utr
chr7	74.07030771	8.019971227	+	ENSG00000106278.7	PTPRZ1	3utr
chr17	72.02050626	4.184159196	-	ENSG00000035862.8	TIMP2	3utr
chr16	66.89720728	6.438424985	+	ENSG00000087245.8	MMP2	3utr
chr5	64.89389141	7.499908527	+	ENSG00000131507.9	NDFIP1	3utr
chr2	64.76386598	5.371633971	-	ENSG00000115310.13	RTN4	3utr
chr4	62.57174239	7.757839672	-	ENSG00000145284.7	SCD5	3utr
chr2	59.90751481	6.499205975	+	ENSG00000138448.7	ITGAV	3utr
chr3	57.7510219	6.440111785	-	ENSG00000249992.1	TMEM158	3utr
chr7	56.37072295	4.367158893	-	ENSG00000146674.10	IGFBP3	3utr
chr4	56.36812144	7.620212461	-	ENSG00000150625.12	GPM6A	3utr
chr5	55.42376234	3.673434928	-	ENSG00000019582.10	CD74	3utr
chr5	55.04095317	6.93175982	+	ENSG00000131711.10	MAP1B	3utr
chr4	53.51981046	5.968547023	+	ENSG00000118785.9	SPP1	3utr
chrX	53.24762415	6.144752641	+	ENSG00000071553.12	ATP6AP1	3utr
chr8	52.25320783	7.709703391	+	ENSG00000168615.7	ADAM9	3utr
chr6	52.17422996	3.019355171	+	ENSG00000204580.7	DDR1	3utr
chr15	51.0747137	4.097832924	-	ENSG00000067225.13	PKM	3utr
chr5	49.88087911	6.489699082	+	ENSG00000131711.10	MAP1B	3utr
chr8	49.5919976	4.691394721	-	ENSG00000164733.16	CTSB	3utr
chr15	48.98493379	4.106230029	+	ENSG00000166710.13	B2M	3utr
chr16	48.85010379	4.952091381	+	ENSG00000087245.8	MMP2	3utr
chr12	48.18395494	5.745688464	+	ENSG00000174437.12	ATP2A2	3utr CDS
chr12	47.29393116	4.155004143	+	ENSG00000111640.10	GAPDH	3utr
chr15	46.05393111	4.025920892	-	ENSG00000067225.13	PKM	3utr
chr1	45.7614308	7.389914631	+	ENSG00000117600.8	LPPR4	3utr
chr2	43.85614242	5.036419588	+	ENSG00000213639.5	PPP1CB	3utr
chr11	42.72397501	3.390281252	+	ENSG00000134824.9	FADS2	3utr
chr6	40.54225356	6.030715387	-	ENSG00000112276.9	BVES	3utr
chr10	40.29638693	5.680760829	+	ENSG00000107984.5	DKK1	3utr
chr1	39.7511933	4.808361811	-	ENSG00000171603.12	CLSTN1	3utr
chr5	39.74508304	7.120078512	-	ENSG00000171617.9	ENC1	3utr
chr14	39.21487853	6.1380383	-	ENSG00000080824.14	HSP90AA1	3utr
chr17	38.6256614	5.395194207	-	ENSG00000131095.7	GFAP	3utr
chr18	38.19222827	3.891465185	-	ENSG00000170558.4	CDH2	3utr
chr3	37.00605891	5.801302771	+	ENSG00000168036.12	CTNNB1	3utr

Supplementary Table 4.2: Shared genes between SU_MB002 and NSC201cb (N=110)

GENE ID	GENE ID
ACTR2	NHLRC2
ANP32A	NONO
AP2B1	NORAD
APP	NUCKS1
ARL6IP1	NUFIP2
ARPP19	NXN
ATP2A2	PARP1
ATP6AP2	PFN2
ATP6V1G1	PGAM1
ATPAF1	PGK1
AZIN1	PPP1CB
CADM1	PREPL
CALD1	PRKAR1A
CALM1	PTPRG
CALU	PTTG1IP
CANX	RAB11A
CCNI	RAC1
CDC42	RGMB
CDK6	RNF187
CKAP5	RPN2
CKB	RTN4
CLIC4	SCARB2
CLSTN1	SCD
CTNNB1	SEC14L1
DCBLD2	SEPT11
DDAH2	SERINC1
DRAP1	SESN3
DYNLL2	SFXN3
EIF4A2	SLC16A1
EIF4G2	SLC38A2
ELOVL6	SLC39A14
EPT1	SLC39A6
FAM168B	SLC5A3
FSCN1	SON
GLUL	SPCS1
GNG12	SPCS3
GSE1	SRP14
GSPT1	STOM
HIPK2	TCF12
HMGCS1	TMBIM6
HNRNPU	TMED7
HSP90AA1	TMEM123
IGF1R	TMEM248
LAPTM4B	TNPO1
LDHA	TOMM20
LIPA	TRIM9
MAP1B	TROVE2
MARCH6	TSPAN3
MARCKS	TTC3
MARCKSL1	TUBB
MLEC	WNK1
NARS	YWHAG
NCKAP1	ZDHHC9
NDFIP1	ZNF106
NGRN	ZNF664

Supplementary Table 4.3: MSI1 eCLIP binding region counts for NSC201cb and SU_MB002

Binding Region	NSC201cb	SU_MB002
Coding Sequence	2979	5021
Distal intron	3994	7842
Proximal intron	1407	1178
5' UTR	173	160
3' UTR	68795	83213
Uncategorized	678	1506

Supplementary Table 4.4: Associated MSI1 eCLIP bound gene pathways

Term ID	Pathway	*p*-value	NSC201cb	SU_MB002
GO:0006397	mRNA processing	3.96E-13	0	1
GO:0008380	RNA splicing	4.25E-13	0	1
GO:0043062	extracellular structure organization	4.57E-13	1	0
GO:0000377	RNA splicing, via transesterification reactions with bulged adenosine as nucleophile	1.65E-12	0	1
GO:0000398	mRNA splicing, via spliceosome	1.65E-12	0	1
GO:0030198	extracellular matrix organization	2.29E-12	1	0
GO:0000375	RNA splicing, via transesterification reactions	2.30E-12	0	1
REAC:R-HSA-72163	mRNA Splicing - Major Pathway	4.75E-10	0	1
REAC:R-HSA-72172	mRNA Splicing	1.43E-09	0	1
GO:0050839	cell adhesion molecule binding	4.04E-09	1	0
REAC:R-HSA-72203	Processing of Capped Intron-Containing Pre-mRNA	5.68E-09	0	1
GO:0098589	membrane region	1.53E-08	1	0
GO:0044437	vacuolar part	3.20E-08	1	0
GO:0098857	membrane microdomain	4.85E-08	1	0
GO:1903311	regulation of mRNA metabolic process	5.01E-08	0	1
GO:0060284	regulation of cell development	5.06E-08	1	0
REAC:R-HSA-1474244	Extracellular matrix organization	5.54E-08	1	0
GO:0034655	nucleobase-containing compound catabolic process	6.23E-08	0	1
GO:0031252	cell leading edge	6.62E-08	1	0
GO:0006401	RNA catabolic process	8.55E-08	0	1
GO:0006402	mRNA catabolic process	1.23E-07	0	1
GO:0006403	RNA localization	1.95E-07	0	1
GO:0045121	membrane raft	2.71E-07	1	0
GO:0050767	regulation of neurogenesis	3.55E-07	1	0
GO:0030175	filopodium	4.40E-07	1	0
GO:0006413	translational initiation	5.21E-07	0	1
GO:0006417	regulation of translation	8.23E-07	0	1
GO:0051960	regulation of nervous system development	1.16E-06	1	0
GO:0031589	cell-substrate adhesion	1.61E-06	1	0
GO:0043312	neutrophil degranulation	2.23E-06	1	0
GO:0030667	secretory granule membrane	2.46E-06	1	0
GO:0002283	neutrophil activation involved in immune response	2.64E-06	1	0
GO:0036464	cytoplasmic ribonucleoprotein granule	2.75E-06	0	1
GO:0042119	neutrophil activation	4.13E-06	1	0
GO:0002446	neutrophil mediated immunity	4.37E-06	1	0
GO:0000904	cell morphogenesis involved in differentiation	5.22E-06	1	0
GO:0006446	regulation of translational initiation	5.62E-06	0	1
GO:0043202	lysosomal lumen	5.71E-06	1	0
GO:0036230	granulocyte activation	5.74E-06	1	0
GO:0034248	regulation of cellular amide metabolic process	7.60E-06	0	1
GO:0035770	ribonucleoprotein granule	8.79E-06	0	1
GO:0005775	vacuolar lumen	9.33E-06	1	0
GO:0071826	ribonucleoprotein complex subunit organization	1.21E-05	0	1
GO:0022618	ribonucleoprotein complex assembly	1.73E-05	0	1
GO:0031344	regulation of cell projection organization	1.90E-05	1	0
REAC:R-HSA-9006934	Signaling by Receptor Tyrosine Kinases	2.04E-05	1	0
GO:0050657	nucleic acid transport	2.06E-05	0	1
GO:0050658	RNA transport	2.06E-05	0	1
GO:0070937	CRD-mediated mRNA stability complex	2.10E-05	0	1
GO:0003727	single-stranded RNA binding	2.10E-05	0	1
GO:0042063	gliogenesis	2.12E-05	1	0
GO:0043487	regulation of RNA stability	2.34E-05	0	1
GO:0120039	plasma membrane bounded cell projection morphogenesis	2.40E-05	1	0
GO:0048858	cell projection morphogenesis	2.86E-05	1	0

GO:0051236	establishment of RNA localization	2.87E-05	0	1
GO:0098852	lytic vacuole membrane	2.88E-05	1	0
GO:0005765	lysosomal membrane	2.88E-05	1	0
GO:0015629	actin cytoskeleton	3.09E-05	1	0
GO:0120035	regulation of plasma membrane bounded cell projection organization	3.13E-05	1	0
GO:0031124	mRNA 3'-end processing	3.34E-05	0	1
GO:0098858	actin-based cell projection	3.46E-05	1	0
GO:0005788	endoplasmic reticulum lumen	3.50E-05	1	0
GO:0010494	cytoplasmic stress granule	3.76E-05	0	1
GO:0015931	nucleobase-containing compound transport	4.71E-05	0	1
REAC:R-HSA-8957275	Post-translational protein phosphorylation	4.72E-05	1	0
GO:0032990	cell part morphogenesis	5.05E-05	1	0
GO:0043488	regulation of mRNA stability	5.23E-05	0	1
GO:0031123	RNA 3'-end processing	5.48E-05	0	1
GO:0006913	nucleocytoplasmic transport	5.49E-05	0	1
REAC:R-HSA-6798695	Neutrophil degranulation	6.55E-05	1	0
GO:0043235	receptor complex	6.64E-05	1	0
GO:0030027	lamellipodium	6.77E-05	1	0
GO:0051169	nuclear transport	6.82E-05	0	1
REAC:R-HSA-381426	Regulation of Insulin-like Growth Factor (IGF) transport and uptake by Insulin-like Growth Factor Binding Proteins	6.82E-05	1	0
GO:0030424	axon	8.26E-05	1	0
GO:0006405	RNA export from nucleus	9.93E-05	0	1
GO:0098797	plasma membrane protein complex	0.000103	1	0
GO:0005518	collagen binding	0.000108	1	0
GO:0051592	response to calcium ion	0.000115	1	0
GO:0071277	cellular response to calcium ion	0.000126	1	0
REAC:R-HSA-72766	Translation	0.000138	0	1
GO:0048812	neuron projection morphogenesis	0.00015	1	0
GO:0005635	nuclear envelope	0.000153	0	1
GO:0030133	transport vesicle	0.000153	1	0
GO:0007160	cell-matrix adhesion	0.000163	1	0
GO:0000785	chromatin	0.000172	0	1
GO:0005774	vacuolar membrane	0.000234	1	0
GO:0051052	regulation of DNA metabolic process	0.00026	0	1
REAC:R-HSA-216083	Integrin cell surface interactions	0.000274	1	0
GO:0046916	cellular transition metal ion homeostasis	0.000279	0	1
GO:0033627	cell adhesion mediated by integrin	0.000288	1	0
REAC:R-HSA-2682334	EPH-Ephrin signaling	0.000289	1	0
GO:0061564	axon development	0.000291	1	0
GO:0072594	establishment of protein localization to organelle	0.000315	0	1
GO:0071564	npBAF complex	0.000316	0	1
REAC:R-HSA-4420097	VEGFA-VEGFR2 Pathway	0.00032	1	0
GO:0030155	regulation of cell adhesion	0.000366	1	0
GO:1903293	phosphatase complex	0.000369	0	1
GO:0008287	protein serine/threonine phosphatase complex	0.000369	0	1
GO:0050840	extracellular matrix binding	0.000435	1	0
GO:0034330	cell junction organization	0.000439	1	0
REAC:R-HSA-373755	Semaphorin interactions	0.000461	1	0
GO:0071013	catalytic step 2 spliceosome	0.000462	0	1
GO:0045664	regulation of neuron differentiation	0.000526	1	0
GO:0032870	cellular response to hormone stimulus	0.000581	1	0
REAC:R-HSA-195258	RHO GTPase Effectors	0.000595	1	1
GO:0031346	positive regulation of cell projection organization	0.00061	1	0
GO:0009611	response to wounding	0.000629	1	0
GO:0061013	regulation of mRNA catabolic process	0.000649	0	1

REAC:R-HSA-194138	Signaling by VEGF	0.000678	1	0
GO:0045296	cadherin binding	0.00073	1	1
GO:0031258	lamellipodium membrane	0.00081	1	0
GO:0042382	paraspeckles	0.000853	0	1
GO:0060627	regulation of vesicle-mediated transport	0.000863	1	0
GO:0097435	supramolecular fiber organization	0.000882	1	0
GO:0055076	transition metal ion homeostasis	0.000933	0	1
GO:0048667	cell morphogenesis involved in neuron differentiation	0.000936	1	0
GO:0032241	positive regulation of nucleobase-containing compound transport	0.00103	0	1
GO:0046833	positive regulation of RNA export from nucleus	0.00103	0	1
GO:0070934	CRD-mediated mRNA stabilization	0.00103	0	1
REAC:R-HSA-114452	Activation of BH3-only proteins	0.00103	0	1
REAC:R-HSA-5628897	TP53 Regulates Metabolic Genes	0.00104	0	1
GO:0051168	nuclear export	0.00109	0	1
GO:0003682	chromatin binding	0.00112	0	1
GO:0070161	anchoring junction	0.00123	1	1
GO:1903312	negative regulation of mRNA metabolic process	0.00126	0	1
GO:0043489	RNA stabilization	0.00133	0	1
GO:0007409	axonogenesis	0.00133	1	0
GO:0120034	positive regulation of plasma membrane bounded cell projection assembly	0.00136	1	0
REAC:R-HSA-156827	L13a-mediated translational silencing of Ceruloplasmin expression	0.00137	0	1
GO:0032204	regulation of telomere maintenance	0.00147	0	1
REAC:R-HSA-5627123	RHO GTPases activate PAKs	0.00147	1	0
GO:0044770	cell cycle phase transition	0.00148	0	1
GO:0098687	chromosomal region	0.00148	0	1
GO:0060560	developmental growth involved in morphogenesis	0.00169	1	0
GO:0010720	positive regulation of cell development	0.00171	1	0
GO:0000139	Golgi membrane	0.00172	1	0
GO:0098553	lumenal side of endoplasmic reticulum membrane	0.00172	1	0
GO:0071556	integral component of lumenal side of endoplasmic reticulum membrane	0.00172	1	0
GO:0005911	cell-cell junction	0.00175	1	0
GO:0051491	positive regulation of filopodium assembly	0.00179	1	0
GO:0008143	poly(A) binding	0.00185	0	1
GO:0070820	tertiary granule	0.00185	1	0
GO:0030658	transport vesicle membrane	0.00189	1	0
GO:0005793	endoplasmic reticulum-Golgi intermediate compartment	0.00192	1	0
REAC:R-HSA-6803529	FGFR2 alternative splicing	0.00193	0	1
GO:0033116	endoplasmic reticulum-Golgi intermediate compartment membrane	0.00197	1	0

Supplementary Table 4.5A: Pathway analysis SU_MB002 MSI1-bound genes associated with Chromatin, Transcription and Translation, JAK-STAT and TGF-β pathways

Chromatin, Transcription, Translation						JAK-STAT	TGF- β
ACTR2	DCP2	HLTF	MCM4	PTBP3	SQLE	CANX	CCNT1
AGO1	DCTN4	HMGA2	MDN1	PTGES3	SRP14	CDC42	DRAP1
AKIRIN1	DCTN5	HMGB2	MORF4L2	PTTG1IP	SRSF1	HNRNPA2B1	FKBP1A
ANP32A	DCTPP1	HMGCR	MPZL1	RAB11A	SRSF2	HNRNPDL	MYC
ANP32E	DDX1	HMGCS1	MRPL42	RAB1A	SRSF6	HNRNPF	PPP1CB
APP	DDX3X	HMGN1	MRPL48	RAB21	SSR3	HSPA9	PPP1CC
ARHGAP42	DEK	HNRNPA0	MRPS28	RAC1	SSU72	PAK2	TFDP2
ARHGEF12	DHX9	HNRNPA1	MRPS30	RACGAP1	STAMBP	SOD1	TGFBR1
ARID1A	DKC1	HNRNPA2B1	MTPN	RAD23B	STMN1		
ARL6IP1	DNAJC8	HNRNPA3	MXI1	RAN	STRAP		
ARPP19	DPYSL2	HNRNPD	MYC	RBBP4	SUMO2		
ATP6V1G1	DR1	HNRNPDL	NARS	RBFOX2	SYCP2L		
AUTS2	DRAP1	HNRNPF	NCBP2	RBM14	SYNCRIP		
AZIN1	DYNLL2	HNRNPH1	NCKAP1	RBMX	SYNE2		
BACH1	EID1	HNRNPH2	NDFIP1	RERE	TBL1XR1		
BASP1	EIF1AX	HNRNPH3	NELL2	RLIM	TCF12		
BAZ2A	EIF2A	HNRNPR	NEUROD1	RPL15	TFAM		
BCAT1	EIF2S1	HNRNPU	NGRN	RPL22L1	TFDP2		
BRD3	EIF2S3	HPRT1	NOLC1	RPL23A	TFRC		
BTBD1	EIF3E	HSBP1	NONO	RPL27A	TGFBR1		
BUB3	EIF4A2	HSP90AA1	NPM1	RPL37	TIMM13		
C12ORF57	EIF4B	HSPA4	NUCKS1	RPN2	TMBIM6		
CADM1	EIF4EBP2	HSPA9	NUDT21	RPS12	TMCO1		
CALD1	EIF4G2	IGF2BP1	NUFIP2	RPS23	TMEM33		
CALM1	ELAVL1	IGF2BP3	NUP155	RRM2	TMEM97		
CALM2	ELOVL5	INA	NUSAP1	RSL24D1	TMPO		
CAND1	ELOVL6	INCENP	ORC6	RTN4	TNPO1		
CANX	EMC4	INSIG1	OTX2	SCARB2	TOMM20		
CAPRIN1	ENAH	IPO7	PABPC1	SCD	TP53BP2		
CBFB	EPB41	JAM3	PAICS	SEC11A	TPM4		
CBX3	ERCC1	KHDRBS1	PAK2	SERP1	TRIM9		
CBX5	EXOC5	KIF5B	PAPOLA	SESN3	TUBA1B		
CCNB1IP1	FAM172A	KPNA4	PARP1	SF3B1	TUBB		
CCNG1	FHL1	KRAS	PAWR	SFPQ	TXNRD1		
CCNT1	FKBP1A	LAMTOR3	PCBP1	SKA2	U2AF1		
CCT4	FSCN1	LAMTOR5	PCBP2	SLBP	UBA52		
CCT5	FUBP1	LAPTM4B	PDLIM5	SLC25A36	UBE2J1		
CDC42	FYTTD1	LARP1	PFN2	SLC39A10	UNC119		
CDK6	G3BP1	LARP4	PGAM1	SLC39A14	UPF3B		
CDKN2A	GABARAPL2	LBR	PGK1	SLC39A6	URI1		
CENPF	GDI2	LDHA	PHF10	SMAD5	USP14		
CHERP	GLS	LIN28B	PNN	SMARCA4	USP33		
CKAP5	GNG12	LIN7C	POLR2D	SMARCC1	WAC		
CKB	GSPT1	LNPEP	PPA1	SMARCD1	WDR77		
CLIC4	GTF2F2	LONP2	PPP1CB	SMARCE1	WDR82		
CNBP	H1F0	LRPPRC	PPP1CC	SMC1A	XPOT		
CNOT6	H2AFV	MAD2L1	PPP1R15B	SMG1	XRCC5		
COX7C	H2AFY	MAGEF1	PPP2CA	SNX1	YWHAE		
CPSF6	H3F3B	MAP1B	PPP3R1	SOD1	YWHAG		
CSDE1	HDGF	MAP1LC3B	PPP4R2	SON	YWHAQ		
CTNNB1	HELLS	MAPRE1	PRKAR1A	SOX11	YWHAZ		
CUL3	HIPK1	MARCH6	PRKDC	SPCS1	ZBTB18		
CYCS	HIPK2	MARCKS	PRNP	SPCS3	ZNF638		
DAAM1	HK2	MCL1	PTBP2				

Supplementary Table 4.5B: Pathway analysis NSC201cb MSI1-bound genes associated with Chromatin, Transcription and Translation Pathway

Chromatin, Transcription, Translation
ADAM9
ANXA2
AZIN1
BCAP31
CSNK1D
GFAP
GNA12
HIPK2
HSP90AA1
HSP90AB1
LAPTM4B
MAP1A
MSN
OAZ2
PTTG1IP
RAB7A
RDX
SGTA
TIMP2
TMTC3
TRIB2

Supplementary Table 4.6: Top 50 transcript differentially expressed (FDR<0.05) in control vs sh*MSI1*-inhibited samples identified and quantified

Gene ID	logFC	*p*-value	adj. *p*-value
SAMD11	6.011459384	3.97E-15	2.26E-11
ID1	5.963583538	4.16E-15	2.26E-11
SMAD9	3.935721549	3.71E-13	1.34E-09
GNGT1	-3.292823676	2.19E-12	5.95E-09
FEZF1-AS1	-2.567671031	3.84E-12	8.33E-09
TNC	-2.310780904	1.24E-11	2.24E-08
DIAPH2-AS1	-2.26936437	1.96E-11	3.04E-08
VXN	-2.983732204	2.64E-11	3.21E-08
CCDC141	2.805275472	2.66E-11	3.21E-08
RXRG	2.692853253	4.27E-11	4.63E-08
KIRREL1	2.323398387	6.74E-11	6.64E-08
NID1	-3.25026725	7.77E-11	7.02E-08
GABRR3	-3.127691513	1.45E-10	1.21E-07
GPRC5B	3.721258921	4.78E-10	3.71E-07
ID3	4.378383972	5.76E-10	4.16E-07
C3orf70	-3.045293789	1.03E-09	6.97E-07
SMAD6	3.938520964	1.33E-09	8.46E-07
ST8SIA3	-2.876526013	1.52E-09	9.18E-07
EEF1A2	3.227050274	1.75E-09	1.00E-06
NEUROG1	2.342033954	1.94E-09	1.06E-06
RAB31	3.357439607	3.04E-09	1.57E-06
LIN28B	1.753633752	3.61E-09	1.68E-06
PRKCB	-3.989528579	3.67E-09	1.68E-06
TSPAN18	2.305888732	3.76E-09	1.68E-06
UCHL1	-3.37932154	3.87E-09	1.68E-06
CELF2	-3.212076214	4.84E-09	2.02E-06
ARSI	3.043440189	5.04E-09	2.03E-06
SOAT1	2.035855402	5.39E-09	2.09E-06
CRABP2	2.110811056	6.09E-09	2.28E-06
KIAA2012	2.702318067	6.69E-09	2.42E-06
SYNPR	2.937504115	8.58E-09	3.00E-06
DOCK8	2.180837079	9.22E-09	3.13E-06
TRIM55	-1.934371715	9.79E-09	3.22E-06
MCTP1	-2.790540676	1.07E-08	3.41E-06
SPAG1	-2.016176018	1.17E-08	3.64E-06
LIMCH1	2.327055316	1.76E-08	5.29E-06
REST	1.929339389	3.65E-08	1.07E-05
ETV3L	3.403079006	6.23E-08	1.78E-05
NRL	2.973061278	9.05E-08	2.52E-05
RALGAPA1	-1.936989934	1.11E-07	3.00E-05
PTPN13	1.917199568	1.14E-07	3.00E-05
B4GALT1	2.399457238	1.29E-07	3.32E-05
CNTN1	-1.942528026	1.40E-07	3.50E-05
NOX5	-1.78678262	1.42E-07	3.50E-05
BASP1	-1.574192654	1.84E-07	4.43E-05
SMYD1	-2.462259577	2.35E-07	5.55E-05
MIOS	1.508871056	3.09E-07	7.12E-05
PCDH10	-2.571616122	3.93E-07	8.71E-05
CHST2	2.418298282	3.93E-07	8.71E-05
IFI44	-2.264241404	4.26E-07	9.08E-05

Supplementary Table 4.7: Top 50 proteins differentially abundant (FDR<0.01) in control vs sh*MSI1* inhibited samples identified and quantified

Uniprot ID	GENE ID	Fold change	Significant B *p*-value (FDR<0.01)
Q5BKX6	**SLC45A4**	4.846941677	5.80E-92
P78324	**SIRPA**	3.347755699	2.54E-28
P30613	**PKLR**	4.671123379	2.77E-25
Q07617	**SPAG1**	2.329673661	8.02E-22
P68431	**HIST1H3A**	2.179437941	1.69E-17
Q13886	**KLF9**	2.635971199	1.96E-17
O75494	**SRSF10**	2.906584799	4.22E-17
P62805	**HIST1H4A**	2.131392532	1.73E-15
P80294	**MT1H**	2.796504093	2.48E-15
P04004	**VTN**	2.341765012	1.50E-14
Q96HD1	**CRELD1**	2.438947711	2.44E-14
P13640	**MT1G**	3.513815422	1.66E-12
Q86VI3	**IQGAP3**	2.535168787	2.06E-12
P20336	**RAB3A**	2.103624306	1.09E-11
Q9UL25	**RAB21**	2.006665815	1.91E-11
P01308	**INS**	1.898978317	8.80E-11
Q96ME7	**ZNF512**	2.208340497	5.00E-10
Q96AQ6	**PBXIP1**	1.928185476	6.71E-10
P46013	**MKI67**	1.684183221	6.94E-10
P04733	**MT1F**	2.174388867	1.47E-09
Q6ZN17	**LIN28B**	1.88662415	3.93E-09
Q96FQ6	**S100A16**	2.340069495	4.23E-09
Q96BY9	**SARAF**	2.118770666	1.03E-08
Q9Y5S1	**TRPV2 VRL**	2.061482458	1.21E-08
Q58FF6	**HSP90AB4P**	2.0013563	1.39E-08
P27824	**CANX**	1.626007741	1.62E-08
Q5SRI9	**MANEA**	2.232466419	2.06E-08
Q8NBN3	**TMEM87A**	1.832099759	3.55E-08
P09917	**ALOX5**	2.933063832	7.04E-08
Q9P2F8	**SIPA1L2**	2.001874046	8.18E-08
P16401	**HIST1H1B**	1.742238787	8.64E-08
Q9UIG0	**BAZ1B**	1.752624689	9.15E-08
Q5SSJ5	**HP1BP3**	1.854996629	1.30E-07
P48443	**RXRG**	2.030938152	1.48E-07
Q9BZB8	**CPEB**	2.871469982	1.86E-07
P02787	**TF PRO1400**	1.575325761	2.04E-07
Q9BXD5	**NPL**	2.179095834	2.53E-07
Q15714	**TSC22D1**	2.006106404	3.02E-07
Q9BVS4	**RIOK2**	1.820320742	5.43E-07
Q8IXH6	**TP53INP2**	2.093728856	7.24E-07
P11021	**HSPA5**	1.546775346	7.77E-07
Q14119	**VEZF1**	1.874691939	8.41E-07
P11387	**TOP1**	1.682218394	1.19E-06
Q8WWK9	**CKAP2**	1.689100083	1.39E-06
Q9BT88	**SYT11**	2.03813364	1.90E-06
Q9UN86	**G3BP2**	2.08688852	2.10E-06
Q16666	**IFI16**	1.661222959	2.46E-06
Q2TB10	**ZNF800**	2.00857166	3.80E-06
Q12767	**TMEM94**	0.286376723	6.28E-06
P35243	**RCVRN**	1.62841598	1.09681E-05

CHAPTER 5

DISCUSSION

5.1 Thesis overview

The characterization of the role of an RBP like MSI through an unbiased multi-omic systems biology approach facilitated the acquisition of a number of large datasets for integrative analysis. While in theory more data has the potential to further our understanding of the biology of genes and protein products, the tools to perturb genes of RBPs and the assays available are victim to their caveats and biases. The dynamic and multiple layers of cancer is described in Chapter 2 with a more detailed discussion of MSI1 and the technology employed to characterize its functional role in G3 MB in Chapter 3. Chapter 4 outlined the bulk of the observations MSI1 and G3 MB with detailed optimizations of the multi-omic analyses completed.

5.2 Future directions

Despite the vast amount of resources applied to further the understanding a dysregulated RNA-binding protein that is highly abundant in the developing brain and in G3 MB CSCs, extensive efforts are still required to translate these findings from benchtop to bedside.

The major findings in Chapter 4 imply MSI1 plays a substantial role in the propagation of cancer stem cells in G3 MB. Given the large emphasis on transcriptomic data to date with the small body of proteomic evidence suggesting aberrant post-transcriptional gene regulation processes in non-Wnt/Shh MB, this work has made some headway in the explanation for the imbalance in mRNA and

protein in these tumours. Limitations of this work are those inherent when employing cell lines cultured in an artificial environment (i.e., culture dish, brain of a immunodeficient mouse) to infer similar biological mechanisms are present as the developing human posterior fossa. Another limitation as alluded to in Chapter 3 is the use of shRNAi to study the effects of MSI1. While multiple hairpins were tested with similar resultant phenotype, the potential for off target effects cannot be underestimated.

Future plans for the data generated from this body of work will be to 1) further investigate the role of MSI2 in G3 MB as well as 2) employ CRISPR-Cas9 knock out constructs to test the essentiality of MSI1 as well as HIPK1 in NSC and G3 MB. Validating a downstream target of MSI1 (i.e., HIPK1), with a similar expression profile in primary patient samples, and is potentially targetable with novel kinase inhibitors, it has the potential to provide the MB community with an option for future therapeutic drug discovery. Further experiments to answer the question of whether there is a change in the rate of nascent mRNA production or in the post-translational landscape would be of interest, as would determining subcellular localization of MSI1 to provide further insight into its mechanism of action. Furthermore, to bring the findings back to the patients, further comparative analysis with the primary patient proteomic and phosphoproteomic data will be of value.

Early efforts to identify novel therapeutic options by employing the data generated from the MSI1 project was completed with Connectivity Mapping analysis. Repurposable drugs that were identified through this analysis included

gossypol (i.e., a known MSI1 inhibitor (Lan et al., 2010)) ciclosporin, etoposide (i.e., currently utilized for treatment of MB), puromycin, cycloheximide (Figure 5.1). The mechanism of action of the 2 latter compounds largely acting to inhibit translation provides further credence to the hypothesis that MSI1 aberrantly modulates the translational landscape of SU_MB002 (i.e., G3 MB). As a potential novel therapeutic adjunct to current standard of care, ciclosporin was initially tested *in vitro* and *in vivo*. *In vitro*, there was a significant reduction in functional stem cell properties but unfortunately during the *in vivo*, a number of mice succumb to the complications of the therapy rather than endpoint secondary to tumour burden making interpretation of the data difficult (Figure 5.2). While this study was worthwhile pursuing at the time, the validation of a downstream target of MSI1 holds greater promise for targeted drug discovery.

With the collaborations fostered, understanding of the principles of biochemical experimental techniques, and proposals for follow up and future experiments conceptualized, I hope to continue work in the field of RBP and neural cancers to be part of the international effort to find a cure for some of the most aggressive subtypes of pediatric brain tumours.

5.3 Closing remarks

This thesis represents the efforts to characterize the role of the RNA-binding protein, MSI1 in G3 MB. Using the cancer stem cell theory as our model system, I employed multiple G3 MB primary cell cultures, lentiviral gene manipulation, xenotransplant models and the unbiased study of G3 MB employing a multi-

omics systems biological approach, to explore the multiple facets of MSI1 function in G3MB. The exploration of post-transcriptional control of gene regulation by RNA binding proteins in cancer is still considered to be in its infancy and the goal of this work was to introduce a novel therapeutic paradigm to consider in our quest for targeted therapy in pediatric medulloblastoma.

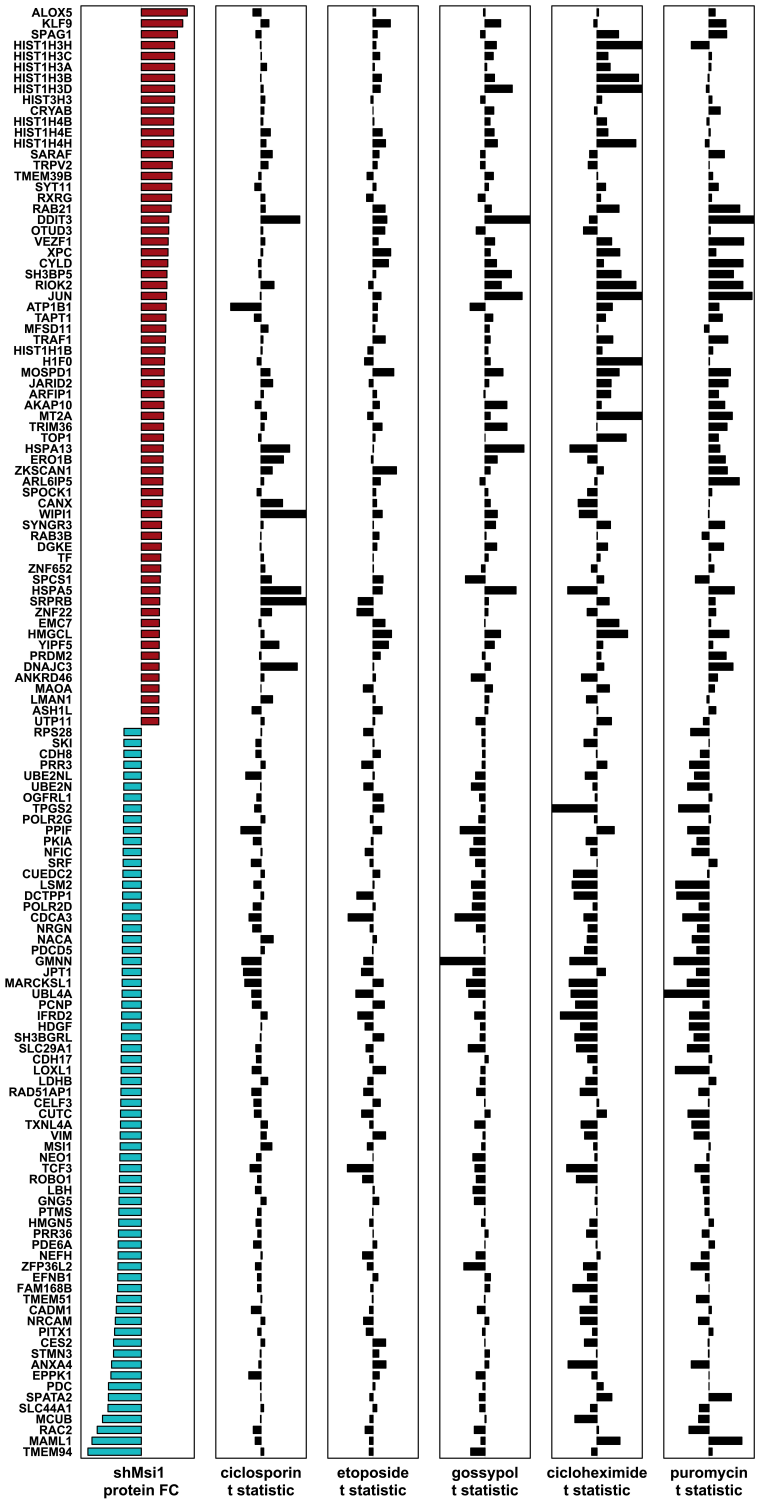


Figure 5.1: CMAP analysis and protein expression profile for Ciclosporin, Etoposide, Gossypol, Cicloheximide and Puromycin.

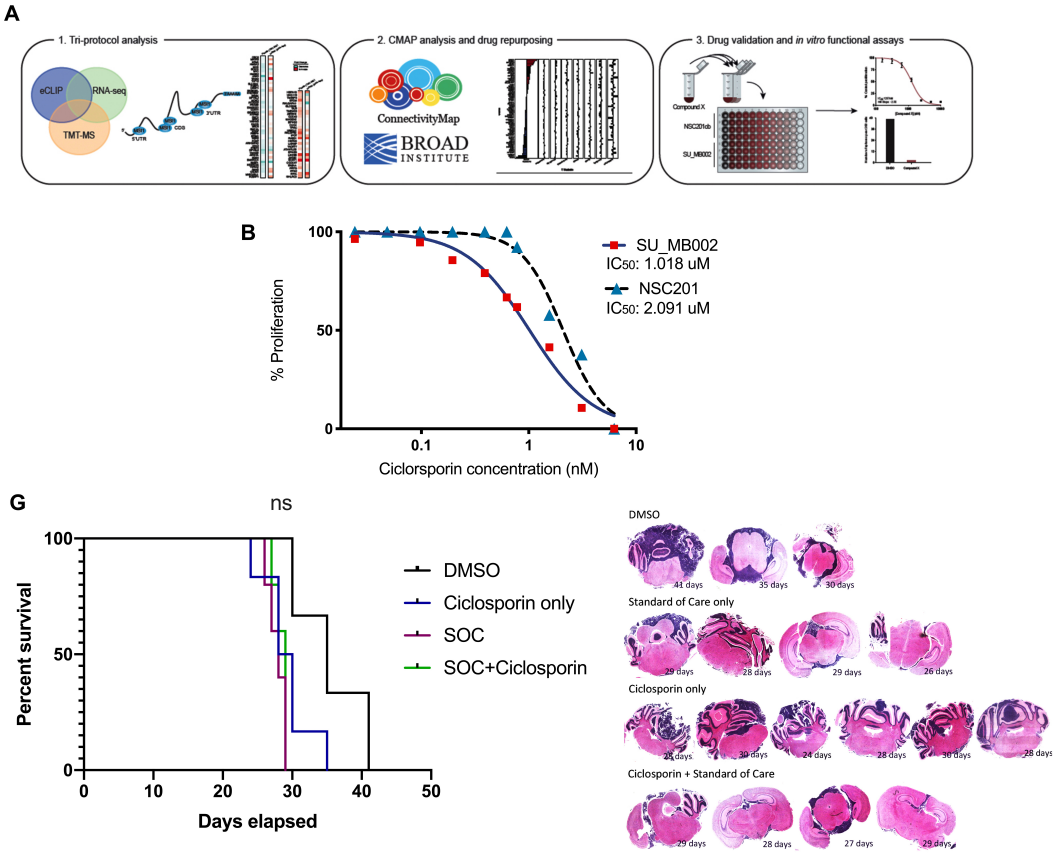


Figure 5.2: CMAP hit Ciclosporin *in vitro* and *in vivo* experiment. A. Schematic CMAP analysis and subsequent in vitro and in vivo experiments, **B.** survival curve showing no significant survival benefit of treatment with ciclosporin nor SOC (Cisplatin, Vincristine and Etoposide), **C.** Tumour burden studies showing limited reduction in tumour burden in all 4 treatment arms.

Bibliography

Abdouh, M., Facchino, S., Chatoo, W., Balasingam, V., Ferreira, J., and Bernier, G. (2009). BMI1 sustains human glioblastoma multiforme stem cell renewal. J Neurosci *29*, 8884-8896.

Adlakha, Y.K., and Seth, P. (2016). The expanding horizon of MicroRNAs in cellular reprogramming. Prog Neurobiol.

Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., and Clarke, M.F. (2003). Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A *100*, 3983-3988.

Aldaz, B., Sagardoy, A., Nogueira, L., Guruceaga, E., Grande, L., Huse, J.T., Aznar, M.A., Diez-Valle, R., Tejada-Solis, S., Alonso, M.M*., et al.* (2013). Involvement of miRNAs in the differentiation of human glioblastoma multiforme stem-like cells. PLoS One *8*, e77098.

Archer, T.C., Ehrenberger, T., Mundt, F., Gold, M.P., Krug, K., Mah, C.K., Mahoney, E.L., Daniel, C.J., LeNail, A., Ramamoorthy, D*., et al.* (2018). Proteomics, Post-translational Modifications, and Integrative Analyses Reveal Molecular Heterogeneity within Medulloblastoma Subgroups. Cancer Cell *34*, 396-410 e398.

Archer, T.C., Fertig, E.J., Gosline, S.J., Hafner, M., Hughes, S.K., Joughin, B.A., Meyer, A.S., Piccolo, S.R., and Shajahan-Haq, A.N. (2016). Systems Approaches to Cancer Biology. Cancer Res *76*, 6774-6777.

Asadi, M.H., Khalifeh, K., and Mowla, S.J. (2016). OCT4 spliced variants are highly expressed in brain cancer tissues and inhibition of OCT4B1 causes G2/M arrest in brain cancer cells. J Neurooncol *130*, 455-463.

Austin, S., Ziese, M., and Sternberg, N. (1981). A novel role for site-specific recombination in maintenance of bacterial replicons. Cell *25*, 729-736.

Baek, D., Villen, J., Shin, C., Camargo, F.D., Gygi, S.P., and Bartel, D.P. (2008). The impact of microRNAs on protein output. Nature *455*, 64-71.

Bailey, P.a.C.H. (1925). Medulloblastoma cerebelli: A common type of midcerebellar glioma of childhood. Archives of Neurology and Psychiatry *14*, 192-224.

Baltz, A.G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M*., et al.* (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol Cell *46*, 674-690.

Bandopadhayay, P., Bergthold, G., Nguyen, B., Schubert, S., Gholamin, S., Tang, Y., Bolin, S., Schumacher, S.E., Zeid, R., Masoud, S.*, et al.* (2014). BET bromodomain inhibition of MYC-amplified medulloblastoma. Clin Cancer Res *20*, 912-925.

Bao, S., Wu, Q., McLendon, R.E., Hao, Y., Shi, Q., Hjelmeland, A.B., Dewhirst, M.W., Bigner, D.D., and Rich, J.N. (2006). Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. Nature *444*, 756-760.

Barciszewska, A.M. (2016). MicroRNAs as efficient biomarkers in high-grade gliomas. Folia Neuropathol *54*, 369-374.

Becker, A.J., Mc, C.E., and Till, J.E. (1963). Cytological demonstration of the clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature *197*, 452-454.

Bernstein, E., Caudy, A.A., Hammond, S.M., and Hannon, G.J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. Nature *409*, 363-366.

Bignell, G.R., Greenman, C.D., Davies, H., Butler, A.P., Edkins, S., Andrews, J.M., Buck, G., Chen, L., Beare, D., Latimer, C.*, et al.* (2010). Signatures of mutation and selection in the cancer genome. Nature *463*, 893-898.

Blough, R.I., Petrij, F., Dauwerse, J.G., Milatovich-Cherry, A., Weiss, L., Saal, H.M., and Rubinstein, J.H. (2000). Variation in microdeletions of the cyclic AMP-responsive element-binding protein gene at chromosome band 16p13.3 in the Rubinstein-Taybi syndrome. Am J Med Genet *90*, 29-34.

Bourdeaut, F., Miquel, C., Richer, W., Grill, J., Zerah, M., Grison, C., Pierron, G., Amiel, J., Krucker, C., Radvanyi, F.*, et al.* (2014). Rubinstein-Taybi syndrome predisposing to non-WNT, non-SHH, group 3 medulloblastoma. Pediatr Blood Cancer *61*, 383-386.

Brandman, O., and Hegde, R.S. (2016). Ribosome-associated protein quality control. Nat Struct Mol Biol *23*, 7-15.

Bull, K.S., Liossi, C., Culliford, D., Peacock, J.L., Kennedy, C.R., Children's, C., and Leukaemia, G. (2014). Child-related characteristics predicting subsequent health-related quality of life in 8- to 14-year-old children with and without cerebellar tumors: a prospective longitudinal study. Neurooncol Pract *1*, 114-122.

Byrum, S.D., Raman, A., Taverna, S.D., and Tackett, A.J. (2012). ChAP-MS: a method for identification of proteins and histone posttranslational modifications at a single genomic locus. Cell Rep *2*, 198-205.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M.*, et al.* (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell *149*, 1393-1406.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2013). System-wide identification of RNA-binding proteins by interactome capture. Nat Protoc *8*, 491-500.

Cavalli, F.M.G., Remke, M., Rampasek, L., Peacock, J., Shih, D.J.H., Luu, B., Garzia, L., Torchia, J., Nor, C., Morrissy, A.S.*, et al.* (2017). Intertumoral Heterogeneity within Medulloblastoma Subgroups. Cancer Cell *31*, 737-754 e736.

Chakrabarti, A.M., Haberman, N., Praznik, A., Luscombe, N.M., and Ule, J. (2018). Data Sciene Issues in Studying Protein-RNA Interactions with CLIP technologies. Annual Review of Biocmedical Data Science *1*, 235-261.

Chandler, H., and Peters, G. (2013). Stressing the cell cycle in senescence and aging. Curr Opin Cell Biol *25*, 765-771.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

Chen, H.Y., Lin, L.T., Wang, M.L., Lee, S.H., Tsai, M.L., Tsai, C.C., Liu, W.H., Chen, T.C., Yang, Y.P., Lee, Y.Y.*, et al.* (2016a). Musashi-1 regulates AKT-derived IL-6 autocrinal/paracrinal malignancy and chemoresistance in glioblastoma. Oncotarget.

Chen, H.Y., Lin, L.T., Wang, M.L., Lee, S.H., Tsai, M.L., Tsai, C.C., Liu, W.H., Chen, T.C., Yang, Y.P., Lee, Y.Y.*, et al.* (2016b). Musashi-1 regulates AKT-derived IL-6 autocrinal/paracrinal malignancy and chemoresistance in glioblastoma. Oncotarget *7*, 42485-42501.

Chen, J., Li, Y., Yu, T.S., McKay, R.M., Burns, D.K., Kernie, S.G., and Parada, L.F. (2012). A restricted cell population propagates glioblastoma growth after chemotherapy. Nature *488*, 522-526.

Chen, L., and Kang, C. (2015). miRNA interventions serve as 'magic bullets' in the reversal of glioblastoma hallmarks. Oncotarget *6*, 38628-38642.

Chen, P.X., Li, Q.Y., and Yang, Z. (2015). Musashi-1 Expression is a Prognostic Factor in Ovarian Adenocarcinoma and Correlates with ALDH-1 Expression. Pathol Oncol Res *21*, 1133-1140.

Chhangawala, S., Rudy, G., Mason, C.E., and Rosenfeld, J.A. (2015). The impact of read length on quantification of differentially expressed genes and splice junction detection. Genome Biol *16*, 131.

Cho, Y.J., Tsherniak, A., Tamayo, P., Santagata, S., Ligon, A., Greulich, H., Berhoukim, R., Amani, V., Goumnerova, L., Eberhart, C.G.*, et al.* (2011). Integrative genomic analysis of medulloblastoma identifies a molecular subgroup that drives poor clinical outcome. J Clin Oncol *29*, 1424-1430.

Clamer, M., Tebaldi, T., Lauria, F., Bernabo, P., Gomez-Biagi, R.F., Marchioretto, M., Kandala, D.T., Minati, L., Perenthaler, E., Gubert, D.*, et al.* (2018). Active Ribosome Profiling with RiboLace. Cell Rep *25*, 1097-1108 e1095.

Clarke, M.F., Dick, J.E., Dirks, P.B., Eaves, C.J., Jamieson, C.H.M., Jones, D.L., Visvader, J., Weissman, I.L., and Wahl, G.M. (2006). Cancer Stem Cells--Perspectives on Current Status and Future Directions: AACR Workshop on Cancer Stem Cells. Cancer Research *66*, 9339-9344.

Clifford, S.C., Lusher, M.E., Lindsey, J.C., Langdon, J.A., Gilbertson, R.J., Straughton, D., and Ellison, D.W. (2006). Wnt/Wingless pathway activation and chromosome 6 loss characterize a distinct molecular sub-group of medulloblastomas associated with a favorable prognosis. Cell Cycle *5*, 2666-2670.

ClinicalTrials.gov (2018a). A Dose Exploration Study With MK-8628 in Participants With Selected Advanced Solid Tumors (MK-8628-006) (Bethesda (MD): National Library of Medicine (US).).

ClinicalTrials.gov (2018b). A Dose-Finding Study of MK-8628, a Small Molecule Inhibitor of the Bromodomain and Extra-Terminal (BET) Proteins, in Adults With Selected Advanced Solid Tumors (MK-8628-003) (Bethesda (MD) National Library of Medicine (US).).

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A.*, et al.* (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819-823.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol *26*, 1367-1372.

Cox, J.L., Wilder, P.J., Gilmore, J.M., Wuebben, E.L., Washburn, M.P., and Rizzino, A. (2013). The SOX2-interactome in brain cancer cells identifies the requirement of MSI2 and USP9X for the growth of brain tumor cells. PLoS One *8*, e62857.

Crick, F. (1970). Molecular biology in the year 2000. Nature *228*, 613-615.

Czermin, B., Melfi, R., McCabe, D., Seitz, V., Imhof, A., and Pirrotta, V. (2002). Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal Polycomb sites. Cell *111*, 185-196.

D'Urso, P.I., D'Urso, O.F., Storelli, C., Mallardo, M., Gianfreda, C.D., Montinaro, A., Cimmino, A., Pietro, C., and Marsigliante, S. (2012). miR-155 is up-regulated in primary and secondary glioblastoma and promotes tumour growth by inhibiting GABA receptors. Int J Oncol *41*, 228-234.

Dahlrot, R.H. (2014). The prognostic value of clinical factors and cancer stem cell-related markers in gliomas. Dan Med J *61*, B4944.

Dahlrot, R.H., Hansen, S., Herrstedt, J., Schroder, H.D., Hjelmborg, J., and Kristensen, B.W. (2013a). Prognostic value of Musashi-1 in gliomas. J Neurooncol *115*, 453-461.

Dahlrot, R.H., Hermansen, S.K., Hansen, S., and Kristensen, B.W. (2013b). What is the clinical value of cancer stem cell markers in gliomas? Int J Clin Exp Pathol *6*, 334-348.

de Araujo, P.R., Gorthi, A., da Silva, A.E., Tonapi, S.S., Vo, D.T., Burns, S.C., Qiao, M., Uren, P.J., Yuan, Z.M., Bishop, A.J.*, et al.* (2016). Musashi1 Impacts Radio-Resistance in Glioblastoma by Controlling DNA-Protein Kinase Catalytic Subunit. Am J Pathol *186*, 2271-2278.

Dick, J.E. (2003). Breast cancer stem cells revealed. Proc Natl Acad Sci U S A *100*, 3547-3549.

Driessens, G., Beck, B., Caauwe, A., Simons, B.D., and Blanpain, C. (2012). Defining the mode of tumour growth by clonal analysis. Nature *488*, 527-530.

Ecker, J., Oehme, I., Mazitschek, R., Korshunov, A., Kool, M., Hielscher, T., Kiss, J., Selt, F., Konrad, C., Lodrini, M.*, et al.* (2015). Targeting class I histone deacetylase 2 in MYC amplified group 3 medulloblastoma. Acta Neuropathol Commun *3*, 22.

Ellison, D. (2002). Classifying the medulloblastoma: insights from morphology and molecular genetics. Neuropathol Appl Neurobiol *28*, 257-282.

Ellison, D.W. (2010). Childhood medulloblastoma: novel approaches to the classification of a heterogeneous disease. Acta Neuropathol *120*, 305-316.

Ellison, D.W., Dalton, J., Kocak, M., Nicholson, S.L., Fraga, C., Neale, G., Kenney, A.M., Brat, D.J., Perry, A., Yong, W.H.*, et al.* (2011). Medulloblastoma: clinicopathological correlates of SHH, WNT, and non-SHH/WNT molecular subgroups. Acta Neuropathol *121*, 381-396.

Ellison, D.W., Onilude, O.E., Lindsey, J.C., Lusher, M.E., Weston, C.L., Taylor, R.E., Pearson, A.D., Clifford, S.C., and Committee, U.K.C.a.s.C.S.G.B.T. (2005). beta-Catenin status predicts a favorable outcome in childhood medulloblastoma: the United Kingdom Children's Cancer Study Group Brain Tumour Committee. J Clin Oncol *23*, 7951-7957.

Esteller, M. (2008). Epigenetics in cancer. N Engl J Med *358*, 1148-1159.

Esteller, M., Garcia-Foncillas, J., Andion, E., Goodman, S.N., Hidalgo, O.F., Vanaclocha, V., Baylin, S.B., and Herman, J.G. (2000). Inactivation of the DNA-repair gene MGMT and the clinical response of gliomas to alkylating agents. N Engl J Med *343*, 1350-1354.

Evans, A.E., Jenkin, R.D., Sposto, R., Ortega, J.A., Wilson, C.B., Wara, W., Ertel, I.J., Kramer, S., Chang, C.H., Leikin, S.L.*, et al.* (1990). The treatment of medulloblastoma. Results of a prospective randomized trial of radiation therapy with and without CCNU, vincristine, and prednisone. J Neurosurg *72*, 572-582.

Facchino, S., Abdouh, M., Chatoo, W., and Bernier, G. (2010). BMI1 confers radioresistance to normal and cancerous neural stem cells through recruitment of the DNA damage response machinery. J Neurosci *30*, 10096-10111.

Favaro, R., Appolloni, I., Pellegatta, S., Sanga, A.B., Pagella, P., Gambini, E., Pisati, F., Ottolenghi, S., Foti, M., Finocchiaro, G.*, et al.* (2014). Sox2 is required to maintain cancer stem cells in a mouse model of high-grade oligodendroglioma. Cancer Res *74*, 1833-1844.

Filippov, V., Solovyev, V., Filippova, M., and Gill, S.S. (2000). A novel type of RNase III family proteins in eukaryotes. Gene *245*, 213-221.

Fire, A., Albertson, D., Harrison, S.W., and Moerman, D.G. (1991). Production of antisense RNA leads to effective and specific inhibition of gene expression in C. elegans muscle. Development *113*, 503-514.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature *391*, 806-811.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev *17*, 1870-1881.

Forget, A., Martignetti, L., Puget, S., Calzone, L., Brabetz, S., Picard, D., Montagud, A., Liva, S., Sta, A., Dingli, F.*, et al.* (2018). Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. Cancer Cell *34*, 379-395 e377.

Fox, R.G., Lytle, N.K., Jaquish, D.V., Park, F.D., Ito, T., Bajaj, J., Koechlein, C.S., Zimdahl, B., Yano, M., Kopp, J.*, et al.* (2016). Image-based detection and targeting of therapy resistance in pancreatic adenocarcinoma. Nature *534*, 407-411.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L., and Paul, C.L. (1992). A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A *89*, 1827-1831.

Gajjar, A., Chintagumpala, M., Ashley, D., Kellie, S., Kun, L.E., Merchant, T.E., Woo, S., Wheeler, G., Ahern, V., Krasin, M.J.*, et al.* (2006). Risk-adapted craniospinal radiotherapy followed by high-dose chemotherapy and stem-cell rescue in children with newly diagnosed medulloblastoma (St Jude Medulloblastoma-96): long-term results from a prospective, multicentre trial. Lancet Oncol *7*, 813-820.

Gajjar, A., and Finlay, J.L. (2015). The management of children and adolescents with medulloblastoma in low and middle income countries. Pediatr Blood Cancer *62*, 549-550.

Gangemi, R.M., Griffero, F., Marubbi, D., Perera, M., Capra, M.C., Malatesta, P., Ravetti, G.L., Zona, G.L., Daga, A., and Corte, G. (2009). SOX2 silencing in glioblastoma tumor-initiating cells causes stop of proliferation and loss of tumorigenicity. Stem Cells *27*, 40-48.

Gibson, P., Tong, Y., Robinson, G., Thompson, M.C., Currle, D.S., Eden, C., Kranenburg, T.A., Hogg, T., Poppleton, H., Martin, J.*, et al.* (2010). Subtypes of medulloblastoma have distinct developmental origins. Nature *468*, 1095-1099.

Glas, M., Happold, C., Rieger, J., Wiewrodt, D., Bahr, O., Steinbach, J.P., Wick, W., Kortmann, R.D., Reifenberger, G., Weller, M.*, et al.* (2009). Long-term survival of patients with glioblastoma treated with radiotherapy and lomustine plus temozolomide. J Clin Oncol *27*, 1257-1261.

Glinsky, G.V. (2007). Stem cell origin of death-from-cancer phenotypes of human prostate and breast cancers. Stem Cell Rev *3*, 79-93.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A.*, et al.* (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science *286*, 531-537.

Goncalves da Silva, P.B., Teixeira Dos Santos, M.C., Rodini, C.O., Kaid, C., Leite Pereira, M.C., Furukawa, G., Gimenes da Cruz, D.S., Goldfeder, M.B., Reily Rocha, C.R., Rosenberg, C.*, et al.* (2017). High OCT4A levels drive tumorigenicity and metastatic potential of medulloblastoma cells. Oncotarget.

Good, P., Yoda, A., Sakakibara, S., Yamamoto, A., Imai, T., Sawa, H., Ikeuchi, T., Tsuji, S., Satoh, H., and Okano, H. (1998). The human Musashi homolog 1 (MSI1) gene encoding the homologue of Musashi/Nrp-1, a neural RNA-binding protein putatively expressed in CNS stem cells and neural progenitor cells. Genomics *52*, 382-384.

Gottardo, N.G., Hansford, J.R., McGlade, J.P., Alvaro, F., Ashley, D.M., Bailey, S., Baker, D.L., Bourdeaut, F., Cho, Y.J., Clay, M.*, et al.* (2014). Medulloblastoma Down Under 2013: a report from the third annual meeting of the International Medulloblastoma Working Group. Acta Neuropathol *127*, 189-201.

Graff, J.N., Higano, C.S., Hahn, N.M., Taylor, M.H., Zhang, B., Zhou, X., Venkatakrishnan, K., Leonard, E.J., and Sarantopoulos, J. (2016). Open-label, multicenter, phase 1 study of alisertib (MLN8237), an aurora A kinase inhibitor, with docetaxel in patients with solid tumors. Cancer *122*, 2524-2533.

Grichnik, J.M. (2006). Genomic instability and tumor stem cells. J Invest Dermatol *126*, 1214-1216.

Guo, B.H., Feng, Y., Zhang, R., Xu, L.H., Li, M.Z., Kung, H.F., Song, L.B., and Zeng, M.S. (2011a). Bmi-1 promotes invasion and metastasis, and its elevated expression is correlated with an advanced stage of breast cancer. Mol Cancer *10*, 10.

Guo, Y., Liu, S., Wang, P., Zhao, S., Wang, F., Bing, L., Zhang, Y., Ling, E.A., Gao, J., and Hao, A. (2011b). Expression profile of embryonic stem cell-associated genes Oct4, Sox2 and Nanog in human gliomas. Histopathology *59*, 763-775.

Gygi, S.P., Rochon, Y., Franza, B.R., and Aebersold, R. (1999). Correlation between protein and mRNA abundance in yeast. Mol Cell Biol *19*, 1720-1730.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.C., Munschauer, M.*, et al.* (2010). PAR-CliP--a method to identify transcriptome-wide the binding sites of RNA binding proteins. J Vis Exp.

Han, J., Zhang, F., Yu, M., Zhao, P., Ji, W., Zhang, H., Wu, B., Wang, Y., and Niu, R. (2012). RNA interference-mediated silencing of NANOG reduces cell proliferation and induces G0/G1 cell cycle arrest in breast cancer cells. Cancer Lett *321*, 80-88.

Hannon, G.J. (2002). RNA interference. Nature *418*, 244-251.

Hashimoto, K., and Tsuji, Y. (2017). Arsenic-Induced Activation of the Homeodomain-Interacting Protein Kinase 2 (HIPK2) to cAMP-Response Element Binding Protein (CREB) Axis. J Mol Biol *429*, 64-78.

Hayes, J., Peruzzi, P.P., and Lawler, S. (2014). MicroRNAs in cancer: biomarkers, functions and therapy. Trends Mol Med *20*, 460-469.

Hegi, M.E., Diserens, A.C., Godard, S., Dietrich, P.Y., Regli, L., Ostermann, S., Otten, P., Van Melle, G., de Tribolet, N., and Stupp, R. (2004). Clinical trial substantiates the predictive value of O-6-methylguanine-DNA methyltransferase promoter methylation in glioblastoma patients treated with temozolomide. Clin Cancer Res *10*, 1871-1874.

Hegi, M.E., Diserens, A.C., Gorlia, T., Hamou, M.F., de Tribolet, N., Weller, M., Kros, J.M., Hainfellner, J.A., Mason, W., Mariani, L.*, et al.* (2005). MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med *352*, 997-1003.

Hemmati, H.D., Nakano, I., Lazareff, J.A., Masterman-Smith, M., Geschwind, D.H., Bronner-Fraser, M., and Kornblum, H.I. (2003). Cancerous stem cells can arise from pediatric brain tumors. Proc Natl Acad Sci U S A *100*, 15178-15183.

Hendrickson, D.G., Hogan, D.J., McCullough, H.L., Myers, J.W., Herschlag, D., Ferrell, J.E., and Brown, P.O. (2009). Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. PLoS Biol *7*, e1000238.

Herrlinger, U., Rieger, J., Koch, D., Loeser, S., Blaschke, B., Kortmann, R.D., Steinbach, J.P., Hundsberger, T., Wick, W., Meyermann, R.*, et al.* (2006). Phase II trial of lomustine plus temozolomide chemotherapy in addition to radiotherapy in newly diagnosed glioblastoma: UKT-03. J Clin Oncol *24*, 4412-4417.

Hill, R.M., Kuijper, S., Lindsey, J.C., Petrie, K., Schwalbe, E.C., Barker, K., Boult, J.K., Williamson, D., Ahmad, Z., Hallsworth, A.*, et al.* (2015). Combined MYC and P53 defects emerge at medulloblastoma relapse and define rapidly progressive, therapeutically targetable disease. Cancer Cell *27*, 72-84.

Hitoshi, S., Alexson, T., Tropepe, V., Donoviel, D., Elia, A.J., Nye, J.S., Conlon, R.A., Mak, T.W., Bernstein, A., and van der Kooy, D. (2002). Notch pathway molecules are essential for the maintenance, but not the generation, of mammalian neural stem cells. Genes Dev *16*, 846-858.

Holliday, R. (1987). The inheritance of epigenetic defects. Science *238*, 163-170.

Honig, A., Weidler, C., Hausler, S., Krockenberger, M., Buchholz, S., Koster, F., Segerer, S.E., Dietl, J., and Engel, J.B. (2010). Overexpression of polycomb protein BMI-1 in human specimens of breast, ovarian, endometrial and cervical cancer. Anticancer Res *30*, 1559-1564.

Hooper, C.M., Hawes, S.M., Kees, U.R., Gottardo, N.G., and Dallas, P.B. (2014). Gene expression analyses of the spatio-temporal relationships of human

medulloblastoma subgroups during early human neurogenesis. PLoS One *9*, e112909.

Hosen, N., Yamane, T., Muijtjens, M., Pham, K., Clarke, M.F., and Weissman, I.L. (2007). Bmi-1-green fluorescent protein-knock-in mice reveal the dynamic regulation of bmi-1 expression in normal and leukemic hematopoietic cells. Stem Cells *25*, 1635-1644.

Huntzinger, E., and Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet *12*, 99-110.

Imai, T., Tokunaga, A., Yoshida, T., Hashimoto, M., Mikoshiba, K., Weinmaster, G., Nakafuku, M., and Okano, H. (2001). The neural RNA-binding protein Musashi1 translationally regulates mammalian numb gene expression by interacting with its mRNA. Mol Cell Biol *21*, 3888-3900.

Ingolia, N.T., Lareau, L.F., and Weissman, J.S. (2011). Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell *147*, 789-802.

Ishimura, R., Nagy, G., Dotu, I., Chuang, J.H., and Ackerman, S.L. (2016). Activation of GCN2 kinase by ribosome stalling links translation elongation with translation initiation. Elife *5*.

Izant, J.G., and Weintraub, H. (1984). Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. Cell *36*, 1007-1015.

Jackson, M., Hassiotou, F., and Nowak, A. (2015). Glioblastoma stem-like cells: at the root of tumor recurrence and a therapeutic target. Carcinogenesis *36*, 177-185.

Jarvelin, A.I., Noerenberg, M., Davis, I., and Castello, A. (2016). The new (dis)order in RNA regulation. Cell Commun Signal *14*, 9.

Ji, X. (2008). The mechanism of RNase III action: how dicer dices. Curr Top Microbiol Immunol *320*, 99-116.

Johannessen, T.C., Wang, J., Skaftnesmo, K.O., Sakariassen, P.O., Enger, P.O., Petersen, K., Oyan, A.M., Kalland, K.H., Bjerkvig, R., and Tysnes, B.B. (2009). Highly infiltrative brain tumours show reduced chemosensitivity associated with a stem cell-like phenotype. Neuropathol Appl Neurobiol *35*, 380-393.

Johnston, D.L., Keene, D., Kostova, M., Strother, D., Lafay-Cousin, L., Fryer, C., Scheinemann, K., Carret, A.S., Fleming, A., Percy, V.*, et al.* (2014). Incidence of medulloblastoma in Canadian children. J Neurooncol *120*, 575-579.

Jordan, C.T. (2004). Cancer stem cell biology: from leukemia to solid tumors. Curr Opin Cell Biol *16*, 708-712.

Kanemura, Y., Mori, K., Sakakibara, S., Fujikawa, H., Hayashi, H., Nakano, A., Matsumoto, T., Tamura, K., Imai, T., Ohnishi, T.*, et al.* (2001). Musashi1, an evolutionarily conserved neural RNA-binding protein, is a versatile marker of human glioma cells in determining their cellular origin, malignancy, and proliferative activity. Differentiation *68*, 141-152.

Kawauchi, D., Robinson, G., Uziel, T., Gibson, P., Rehg, J., Gao, C., Finkelstein, D., Qu, C., Pounds, S., Ellison, D.W.*, et al.* (2012). A mouse model of the most aggressive subgroup of human medulloblastoma. Cancer Cell *21*, 168-180.

Kechavarzi, B., and Janga, S.C. (2014). Dissecting the expression landscape of RNA-binding proteins in human cancers. Genome Biol *15*, R14.

Keene, J.D. (2007). RNA regulons: coordination of post-transcriptional events. Nat Rev Genet *8*, 533-543.

Kenney, A.M., Widlund, H.R., and Rowitch, D.H. (2004). Hedgehog and PI-3 kinase signaling converge on Nmyc1 to promote cell cycle progression in cerebellar neuronal precursors. Development *131*, 217-228.

Kleihues, P.C., WK (2000). Tumours of the Central Nervous System. World Health Organization Classification of Tumours (Lyon: IARC).

Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol *17*, 909-915.

Kool, M., Korshunov, A., Remke, M., Jones, D.T., Schlanstein, M., Northcott, P.A., Cho, Y.J., Koster, J., Schouten-van Meeteren, A., van Vuurden, D.*, et al.* (2012). 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.

Kool, M., Koster, J., Bunt, J., Hasselt, N.E., Lakeman, A., van Sluis, P., Troost, D., Meeteren, N.S., Caron, H.N., Cloos, J.*, et al.* (2008). Integrated genomics identifies five medulloblastoma subtypes with distinct genetic profiles, pathway signatures and clinicopathological features. PLoS One *3*, e3088.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res *19*, 1639-1645.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A.*, et al.* (2016).

Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res *44*, W90-97.

Kumar, M., and Carmichael, G.G. (1998). Antisense RNA: function and fate of duplex RNA in cells of higher eukaryotes. Microbiol Mol Biol Rev *62*, 1415-1434.

Kurdyukov, S., and Bullock, M. (2016). DNA Methylation Analysis: Choosing the Right Method. Biology (Basel) *5*.

Lagadec, C., Vlashi, E., Frohnen, P., Alhiyari, Y., Chan, M., and Pajonk, F. (2014). The RNA-binding protein Musashi-1 regulates proteasome subunit expression in breast cancer- and glioma-initiating cells. Stem Cells *32*, 135-144.

Lampe, I., and Macintyre, R.S. (1954). Experiences in the radiation therapy of medulloblastoma of the cerebellum. Am J Roentgenol Radium Ther Nucl Med *71*, 659-668.

Lan, S.Y., Yu, T., Xia, Z.S., Yuan, Y.H., Shi, L., Lin, Y., Huang, K.H., and Chen, Q.K. (2010). Musashi 1-positive cells derived from mouse embryonic stem cells can differentiate into neural and intestinal epithelial-like cells in vivo. Cell Biol Int *34*, 1171-1180.

Lapointe, C.P., Wilinski, D., Saunders, H.A., and Wickens, M. (2015). Protein-RNA networks revealed through covalent RNA marks. Nat Methods *12*, 1163-1170.

Lavallee-Adam, M., Rauniyar, N., McClatchy, D.B., and Yates, J.R., 3rd (2014). PSEA-Quant: a protein set enrichment analysis on label-free and label-based protein quantification data. J Proteome Res *13*, 5496-5509.

Leedham, S.J., Schier, S., Thliveris, A.T., Halberg, R.B., Newton, M.A., and Wright, N.A. (2005). From gene mutations to tumours--stem cells in gastrointestinal carcinogenesis. Cell Prolif *38*, 387-405.

Leung, C., Lingbeek, M., Shakhova, O., Liu, J., Tanger, E., Saremaslani, P., Van Lohuizen, M., and Marino, S. (2004). Bmi1 is essential for cerebellar development and is overexpressed in human medulloblastomas. Nature *428*, 337-341.

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. Nature *424*, 147-151.

Li, C., Lee, C.J., and Simeone, D.M. (2009). Identification of human pancreatic cancer stem cells. Methods Mol Biol *568*, 161-173.

Li, D.W., Tang, H.M., Fan, J.W., Yan, D.W., Zhou, C.Z., Li, S.X., Wang, X.L., and Peng, Z.H. (2010). Expression level of Bmi-1 oncoprotein is associated with

progression and prognosis in colon cancer. J Cancer Res Clin Oncol *136*, 997-1006.

Li, N., Yousefi, M., Nakauka-Ddamba, A., Li, F., Vandivier, L., Parada, K., Woo, D.H., Wang, S., Naqvi, A.S., Rao, S.*, et al.* (2015). The Msi Family of RNA-Binding Proteins Function Redundantly as Intestinal Oncoproteins. Cell Rep *13*, 2440-2455.

Li, Q., Lee, J.A., and Black, D.L. (2007). Neuronal regulation of alternative pre-mRNA splicing. Nat Rev Neurosci *8*, 819-831.

Li, Y., Choi, P.S., Casey, S.C., and Felsher, D.W. (2014). Activation of Cre recombinase alone can induce complete tumor regression. PLoS One *9*, e107589.

Liau, B.B., Sievers, C., Donohue, L.K., Gillespie, S.M., Flavahan, W.A., Miller, T.E., Venteicher, A.S., Hebert, C.H., Carey, C.D., Rodig, S.J.*, et al.* (2017). Adaptive Chromatin Remodeling Drives Glioblastoma Stem Cell Plasticity and Drug Tolerance. Cell Stem Cell *20*, 233-246 e237.

Lipardi, C., Wei, Q., and Paterson, B.M. (2001). RNAi as random degradative PCR: siRNA primers convert mRNA into dsRNAs that are degraded to generate new siRNAs. Cell *107*, 297-307.

Liu, Y., Zhou, J., and White, K.P. (2014). RNA-seq differential expression studies: more sequence or more replication? Bioinformatics *30*, 301-304.

Louis, D.N., Ohgaki, H., Wiestler, O.D., Cavenee, W.K., Burger, P.C., Jouvet, A., Scheithauer, B.W., and Kleihues, P. (2007). The 2007 WHO classification of tumours of the central nervous system. Acta Neuropathol *114*, 97-109.

Louis, D.N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W.K., Ohgaki, H., Wiestler, O.D., Kleihues, P., and Ellison, D.W. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol *131*, 803-820.

Lovci, M.T., Ghanem, D., Marr, H., Arnold, J., Gee, S., Parra, M., Liang, T.Y., Stark, T.J., Gehman, L.T., Hoon, S.*, et al.* (2013). Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. Nat Struct Mol Biol *20*, 1434-1442.

Lowry, N.A., and Temple, S. (2009). Identifying the perpetrator in medulloblastoma: Dorian Gray versus Benjamin Button. Cancer Cell *15*, 83-85.

Ma, L., Xu, Y.L., Ding, W.J., Shao, H.F., and Teng, Y.C. (2015). Prognostic value of Musashi-1 in endometrioid adenocarcinoma. Int J Clin Exp Pathol *8*, 4564-4572.

MacNicol, A.M., Hardy, L.L., Spencer, H.J., and MacNicol, M.C. (2015). Neural stem and progenitor cell fate transition requires regulation of Musashi1 function. BMC Dev Biol *15*, 15.

Manoranjan, B., Venugopal, C., McFarlane, N., Doble, B.W., Dunn, S.E., Scheinemann, K., and Singh, S.K. (2012). Medulloblastoma stem cells: where development and cancer cross pathways. Pediatr Res *71*, 516-522.

Manoranjan, B., Wang, X., Hallett, R.M., Venugopal, C., Mack, S.C., McFarlane, N., Nolte, S.M., Scheinemann, K., Gunnarsson, T., Hassell, J.A.*, et al.* (2013). FoxG1 interacts with Bmi1 to regulate self-renewal and tumorigenicity of medulloblastoma stem cells. Stem Cells *31*, 1266-1277.

Maury, E., and Hashizume, R. (2017). Epigenetic modification in chromatin machinery and its deregulation in pediatric brain tumors: Insight into epigenetic therapies. Epigenetics, 0.

McCulloch, E.A., and Till, J.E. (1960). The radiation sensitivity of normal mouse bone marrow cells, determined by quantitative marrow transplantation into irradiated mice. Radiat Res *13*, 115-125.

McFarland, J.M., Ho, Z.V., Kugener, G., Dempster, J.M., Montgomery, P.G., Bryan, J.G., Krill-Burger, J.M., Green, T.M., Vazquez, F., Boehm, J.S.*, et al.* (2018). Improved estimation of cancer dependencies from large-scale RNAi screens using model-based normalization and data integration. Nat Commun *9*, 4610.

McMahon, A.C., Rahman, R., Jin, H., Shen, J.L., Fieldsend, A., Luo, W., and Rosbash, M. (2016). TRIBE: Hijacking an RNA-Editing Enzyme to Identify Cell-Specific Targets of RNA-Binding Proteins. Cell *165*, 742-753.

Merkerova, M., Bruchova, H., Kracmarova, A., Klamova, H., and Brdicka, R. (2007). Bmi-1 over-expression plays a secondary role in chronic myeloid leukemia transformation. Leuk Lymphoma *48*, 793-801.

Michiels, E.M., Schouten-Van Meeteren, A.Y., Doz, F., Janssens, G.O., and van Dalen, E.C. (2015). Chemotherapy for children with medulloblastoma. Cochrane Database Syst Rev *1*, CD006678.

Milde, T., Lodrini, M., Savelyeva, L., Korshunov, A., Kool, M., Brueckner, L.M., Antunes, A.S., Oehme, I., Pekrun, A., Pfister, S.M.*, et al.* (2012). HD-MB03 is a novel Group 3 medulloblastoma model demonstrating sensitivity to histone deacetylase inhibitor treatment. J Neurooncol *110*, 335-348.

Min, J., Zhang, Y., and Xu, R.M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev *17*, 1823-1828.

Mittal, N., Roy, N., Babu, M.M., and Janga, S.C. (2009). Dissecting the expression dynamics of RNA-binding proteins in posttranscriptional regulatory networks. Proc Natl Acad Sci U S A *106*, 20300-20305.

Mulhern, R.K., Merchant, T.E., Gajjar, A., Reddick, W.E., and Kun, L.E. (2004). Late neurocognitive sequelae in survivors of brain tumours in childhood. Lancet Oncol *5*, 399-408.

Muller, J., Hart, C.M., Francis, N.J., Vargas, M.L., Sengupta, A., Wild, B., Miller, E.L., O'Connor, M.B., Kingston, R.E., and Simon, J.A. (2002). Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell *111*, 197-208.

Musiyenko, A., Bitko, V., and Barik, S. (2007). RNAi-dependent and -independent antiviral phenotypes of chromosomally integrated shRNA clones: role of VASP in respiratory syncytial virus growth. J Mol Med (Berl) *85*, 745-752.

Muto, J., Imai, T., Ogawa, D., Nishimoto, Y., Okada, Y., Mabuchi, Y., Kawase, T., Iwanami, A., Mischel, P.S., Saya, H.*, et al.* (2012). RNA-binding protein Musashi1 modulates glioma cell growth through the post-transcriptional regulation of Notch and PI3 kinase/Akt signaling pathways. PLoS One *7*, e33431.

Nagy, A. (2000). Cre recombinase: the universal reagent for genome tailoring. Genesis *26*, 99-109.

Nakamura, M., Okano, H., Blendy, J.A., and Montell, C. (1994). Musashi, a neural RNA-binding protein required for Drosophila adult external sensory organ development. Neuron *13*, 67-81.

Nakano, A., Kanemura, Y., Mori, K., Kodama, E., Yamamoto, A., Sakamoto, H., Nakamura, Y., Okano, H., Yamasaki, M., and Arita, N. (2007). Expression of the Neural RNA-binding protein Musashi1 in pediatric brain tumors. Pediatr Neurosurg *43*, 279-284.

Nishimoto, Y., and Okano, H. (2010). New insight into cancer therapeutics: induction of differentiation by regulating the Musashi/Numb/Notch pathway. Cell Res *20*, 1083-1085.

Northcott, P.A., Buchhalter, I., Morrissy, A.S., Hovestadt, V., Weischenfeldt, J., Ehrenberger, T., Grobner, S., Segura-Wang, M., Zichner, T., Rudneva, V.A.*, et al.* (2017). The whole-genome landscape of medulloblastoma subtypes. Nature *547*, 311-317.

Northcott, P.A., Korshunov, A., Pfister, S.M., and Taylor, M.D. (2012a). The clinical implications of medulloblastoma subgroups. Nat Rev Neurol *8*, 340-351.

Northcott, P.A., Korshunov, A., Witt, H., Hielscher, T., Eberhart, C.G., Mack, S., Bouffet, E., Clifford, S.C., Hawkins, C.E., French, P.*, et al.* (2011). Medulloblastoma comprises four distinct molecular variants. J Clin Oncol *29*, 1408-1414.

Northcott, P.A., Shih, D.J., Remke, M., Cho, Y.J., Kool, M., Hawkins, C., Eberhart, C.G., Dubuc, A., Guettouche, T., Cardentey, Y.*, et al.* (2012b). Rapid, reliable, and reproducible molecular sub-grouping of clinical medulloblastoma samples. Acta Neuropathol *123*, 615-626.

Oh, S., Kim, HS (2016). Emerging power of proteomics for delineation of intrinsic tumor subtypes and resistance mechanisms to anti-cancer therapies. Expert Review of Proteomics *12*.

Ohyama, T., Nagata, T., Tsuda, K., Kobayashi, N., Imai, T., Okano, H., Yamazaki, T., and Katahira, M. (2012). Structure of Musashi1 in a complex with target RNA: the role of aromatic stacking interactions. Nucleic Acids Res *40*, 3218-3231.

Okano, H., Kawahara, H., Toriya, M., Nakao, K., Shibata, S., and Imai, T. (2005). Function of RNA-binding protein Musashi-1 in stem cells. Exp Cell Res *306*, 349-356.

Orzan, F., Pellegatta, S., Poliani, P.L., Pisati, F., Caldera, V., Menghi, F., Kapetis, D., Marras, C., Schiffer, D., and Finocchiaro, G. (2011). Enhancer of Zeste 2 (EZH2) is up-regulated in malignant gliomas and in glioma stem-like cells. Neuropathol Appl Neurobiol *37*, 381-394.

Otte, A.P., and Kwaks, T.H. (2003). Gene repression by Polycomb group protein complexes: a distinct complex for every occasion? Current opinion in genetics & development *13*, 448-454.

Ozsolak, F., and Milos, P.M. (2011). RNA sequencing: advances, challenges and opportunities. Nat Rev Genet *12*, 87-98.

P'ng, C., Green, J., Chong, L.C., Waggott, D., Prokopec, S.D., Shamsi, M., Nguyen, F., Mak, D.Y.F., Lam, F., Albuquerque, M.A.*, et al.* (2019). BPG: Seamless, automated and interactive visualization of scientific data. BMC Bioinformatics *20*, 42.

Packer, R.J. (2007). Craniospinal radiation therapy followed by adjuvant chemotherapy for newly diagnosed average-risk medulloblastoma. Curr Neurol Neurosci Rep *7*, 130-132.

Packer, R.J., Sutton, L.N., Atkins, T.E., Radcliffe, J., Bunin, G.R., D'Angio, G., Siegel, K.R., and Schut, L. (1989). A prospective study of cognitive function in children receiving whole-brain radiotherapy and chemotherapy: 2-year results. J Neurosurg *70*, 707-713.

Paczkowska, M., Barenboim, J., Sintupisut, N., Fox, N., Zhu, H., Abd-Rabbo, D., Group, P.N.a.P.A., Boutros, P., and Reimand, J. (2018). Integrative pathway enrichment analysis of multivariate omics data. BbioRxiv.

Paddison, P.J., Caudy, A.A., Bernstein, E., Hannon, G.J., and Conklin, D.S. (2002). Short hairpin RNAs (shRNAs) induce sequence-specific silencing in mammalian cells. Genes Dev *16*, 948-958.

Paduch, R. (2015). Theories of cancer origin. Eur J Cancer Prev *24*, 57-67.

Palmer, S.L., Gajjar, A., Reddick, W.E., Glass, J.O., Kun, L.E., Wu, S., Xiong, X., and Mulhern, R.K. (2003). Predicting intellectual outcome among children treated with 35-40 Gy craniospinal irradiation for medulloblastoma. Neuropsychology *17*, 548-555.

Palmer, S.L., Goloubeva, O., Reddick, W.E., Glass, J.O., Gajjar, A., Kun, L., Merchant, T.E., and Mulhern, R.K. (2001). Patterns of intellectual development among survivors of pediatric medulloblastoma: a longitudinal analysis. J Clin Oncol *19*, 2302-2308.

Panosyan, E.H., Laks, D.R., Masterman-Smith, M., Mottahedeh, J., Yong, W.H., Cloughesy, T.F., Lazareff, J.A., Mischel, P.S., Moore, T.B., and Kornblum, H.I. (2010). Clinical outcome in pediatric glial and embryonal brain tumors correlates with in vitro multi-passageable neurosphere formation. Pediatric Blood & Cancer *55*, 644-651.

Partap, S., Curran, E.K., Propp, J.M., Le, G.M., Sainani, K.L., and Fisher, P.G. (2009). Medulloblastoma incidence has not changed over time: a CBTRUS study. J Pediatr Hematol Oncol *31*, 970-971.

Patel, A.P., Tirosh, I., Trombetta, J.J., Shalek, A.K., Gillespie, S.M., Wakimoto, H., Cahill, D.P., Nahed, B.V., Curry, W.T., Martuza, R.L.*, et al.* (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science *344*, 1396-1401.

Pei, Y., Moore, C.E., Wang, J., Tewari, A.K., Eroshkin, A., Cho, Y.J., Witt, H., Korshunov, A., Read, T.A., Sun, J.L.*, et al.* (2012). An animal model of MYC-driven medulloblastoma. Cancer Cell *21*, 155-167.

Peng, S., Maihle, N.J., and Huang, Y. (2010). Pluripotency factors Lin28 and Oct4 identify a sub-population of stem cell-like cells in ovarian cancer. Oncogene *29*, 2153-2159.

Polkinghorn, W.R., and Tarbell, N.J. (2007). Medulloblastoma: tumorigenesis, current clinical paradigm, and efforts to improve risk stratification. Nat Clin Pract Oncol *4*, 295-304.

Pomeroy, S.L., Tamayo, P., Gaasenbeek, M., Sturla, L.M., Angelo, M., McLaughlin, M.E., Kim, J.Y., Goumnerova, L.C., Black, P.M., Lau, C.*, et al.* (2002). Prediction of central nervous system embryonal tumour outcome based on gene expression. Nature *415*, 436-442.

Pritchard, C.C., Cheng, H.H., and Tewari, M. (2012). MicroRNA profiling: approaches and considerations. Nat Rev Genet *13*, 358-369.

Rahman, N., and Scott, R.H. (2007). Cancer genes associated with phenotypes in monoallelic and biallelic mutation carriers: new lessons from old players. Hum Mol Genet *16 Spec No 1*, R60-66.

Ramaswamy, V., Northcott, P.A., and Taylor, M.D. (2011). FISH and chips: the recipe for improved prognostication and outcomes for children with medulloblastoma. Cancer Genet *204*, 577-588.

Ramaswamy, V., Remke, M., Bouffet, E., Bailey, S., Clifford, S.C., Doz, F., Kool, M., Dufour, C., Vassal, G., Milde, T.*, et al.* (2016). Risk stratification of childhood medulloblastoma in the molecular era: the current consensus. Acta Neuropathol.

Ramskold, D., Luo, S., Wang, Y.C., Li, R., Deng, Q., Faridani, O.R., Daniels, G.A., Khrebtukova, I., Loring, J.F., Laurent, L.C.*, et al.* (2012). Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol *30*, 777-782.

Read, T.A., Fogarty, M.P., Markant, S.L., McLendon, R.E., Wei, Z., Ellison, D.W., Febbo, P.G., and Wechsler-Reya, R.J. (2009). Identification of CD15 as a marker for tumor-propagating cells in a mouse model of medulloblastoma. Cancer Cell *15*, 135-147.

Reimand, J., Kull, M., Peterson, H., Hansen, J., and Vilo, J. (2007). g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Res *35*, W193-200.

Rentas, S., Holzapfel, N.T., Belew, M.S., Pratt, G.A., Voisin, V., Wilhelm, B.T., Bader, G.D., Yeo, G.W., and Hope, K.J. (2016). Musashi-2 attenuates AHR signalling to expand human haematopoietic stem cells. Nature *532*, 508-511.

Reya, T., Morrison, S.J., Clarke, M.F., and Weissman, I.L. (2001). Stem cells, cancer, and cancer stem cells. Nature *414*, 105-111.

Ris, M.D., Packer, R., Goldwein, J., Jones-Wallace, D., and Boyett, J.M. (2001). Intellectual outcome after reduced-dose radiation therapy plus adjuvant chemotherapy for medulloblastoma: a Children's Cancer Group study. J Clin Oncol *19*, 3470-3476.

Robertson, D., Savage, K., Reis-Filho, J.S., and Isacke, C.M. (2008). Multiple immunofluorescence labelling of formalin-fixed paraffin-embedded (FFPE) tissue. BMC Cell Biol *9*, 13.

Robinson, G., Parker, M., Kranenburg, T.A., Lu, C., Chen, X., Ding, L., Phoenix, T.N., Hedlund, E., Wei, L., Zhu, X.*, et al.* (2012). Novel mutations target distinct subgroups of medulloblastoma. Nature *488*, 43-48.

Robinson, G.W., Rudneva, V.A., Buchhalter, I., Billups, C.A., Waszak, S.M., Smith, K.S., Bowers, D.C., Bendel, A., Fisher, P.G., Partap, S.*, et al.* (2018). Risk-adapted therapy for young children with medulloblastoma (SJYC07): therapeutic and molecular outcomes from a multicentre, phase 2 trial. Lancet Oncol *19*, 768-784.

Rogers, H.A., Sousa, S., Salto, C., Arenas, E., Coyle, B., and Grundy, R.G. (2012). WNT/beta-catenin pathway activation in Myc immortalised cerebellar progenitor cells inhibits neuronal differentiation and generates tumours resembling medulloblastoma. Br J Cancer *107*, 1144-1152.

Rutkowski, S., Bode, U., Deinlein, F., Ottensmeier, H., Warmuth-Metz, M., Soerensen, N., Graf, N., Emser, A., Pietsch, T., Wolff, J.E.*, et al.* (2005). Treatment of early childhood medulloblastoma by postoperative chemotherapy alone. N Engl J Med *352*, 978-986.

Sakakibara, S., Imai, T., Hamaguchi, K., Okabe, M., Aruga, J., Nakajima, K., Yasutomi, D., Nagata, T., Kurihara, Y., Uesugi, S.*, et al.* (1996). Mouse-Musashi-1, a neural RNA-binding protein highly enriched in the mammalian CNS stem cell. Dev Biol *176*, 230-242.

Sakakibara, S., Nakamura, Y., Satoh, H., and Okano, H. (2001). Rna-binding protein Musashi2: developmentally regulated expression in neural precursor cells and subpopulations of neurons in mammalian CNS. J Neurosci *21*, 8091-8107.

Sakakibara, S., Nakamura, Y., Yoshida, T., Shibata, S., Koike, M., Takano, H., Ueda, S., Uchiyama, Y., Noda, T., and Okano, H. (2002). RNA-binding protein Musashi family: roles for CNS stem cells and a subpopulation of ependymal cells revealed by targeted disruption and antisense ablation. Proc Natl Acad Sci U S A *99*, 15194-15199.

Sakakibara, S., and Okano, H. (1997). Expression of neural RNA-binding proteins in the postnatal CNS: implications of their roles in neuronal and glial cell development. J Neurosci *17*, 8300-8312.

Sanchez-Diaz, P.C., Burton, T.L., Burns, S.C., Hung, J.Y., and Penalva, L.O. (2008). Musashi1 modulates cell proliferation genes in the medulloblastoma cell line Daoy. BMC Cancer *8*, 280.

Sauer, B. (1998). Inducible gene targeting in mice using the Cre/lox system. Methods *14*, 381-392.

Saxena, S., Jonsson, Z.O., and Dutta, A. (2003). Small RNAs with imperfect match to endogenous mRNA repress translation. Implications for off-target activity of small inhibitory RNA in mammalian cells. J Biol Chem *278*, 44312-44319.

Schepers, A.G., Snippert, H.J., Stange, D.E., van den Born, M., van Es, J.H., van de Wetering, M., and Clevers, H. (2012). Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas. Science *337*, 730-735.

Schoffski, P., Jones, S.F., Dumez, H., Infante, J.R., Van Mieghem, E., Fowst, C., Gerletti, P., Xu, H., Jakubczak, J.L., English, P.A*., et al.* (2011). Phase I, open-label, multicentre, dose-escalation, pharmacokinetic and pharmacodynamic trial of the oral aurora kinase inhibitor PF-03814735 in advanced solid tumours. Eur J Cancer *47*, 2256-2264.

Schuller, U., Heine, V.M., Mao, J., Kho, A.T., Dillon, A.K., Han, Y.G., Huillard, E., Sun, T., Ligon, A.H., Qian, Y*., et al.* (2008). Acquisition of granule neuron precursor identity is a critical determinant of progenitor cell competence to form Shh-induced medulloblastoma. Cancer Cell *14*, 123-134.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.

Sciuscio, D., Diserens, A.C., van Dommelen, K., Martinet, D., Jones, G., Janzer, R.C., Pollo, C., Hamou, M.F., Kaina, B., Stupp, R*., et al.* (2011). Extent and patterns of MGMT promoter methylation in glioblastoma- and respective glioblastoma-derived spheres. Clin Cancer Res *17*, 255-266.

Selbach, M., Schwanhausser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature *455*, 58-63.

Seymour, J.F., Kim, D.W., Rubin, E., Haregewoin, A., Clark, J., Watson, P., Hughes, T., Dufva, I., Jimenez, J.L., Mahon, F.X*., et al.* (2014). A phase 2 study of MK-0457 in patients with BCR-ABL T315I mutant chronic myelogenous leukemia and philadelphia chromosome-positive acute lymphoblastic leukemia. Blood Cancer J *4*, e238.

Seymour, T., Nowak, A., and Kakulas, F. (2015a). Targeting Aggressive Cancer Stem Cells in Glioblastoma. Front Oncol *5*, 159.

Seymour, T., Twigger, A.J., and Kakulas, F. (2015b). Pluripotency Genes and Their Functions in the Normal and Aberrant Breast and Brain. Int J Mol Sci *16*, 27288-27301.

Sharpless, N.E., and Sherr, C.J. (2015). Forging a signature of in vivo senescence. Nat Rev Cancer *15*, 397-408.

Shi, Q., Qin, L., Wei, W., Geng, F., Fan, R., Shin, Y.S., Guo, D., Hood, L., Mischel, P.S., and Heath, J.R. (2012). Single-cell proteomic chip for profiling intracellular signaling pathways in single tumor cells. Proc Natl Acad Sci U S A *109*, 419-424.

Sijen, T., Fleenor, J., Simmer, F., Thijssen, K.L., Parrish, S., Timmons, L., Plasterk, R.H., and Fire, A. (2001). On the role of RNA amplification in dsRNA-triggered gene silencing. Cell *107*, 465-476.

Singer, O., and Verma, I.M. (2008). Applications of lentiviral vectors for shRNA delivery and transgenesis. Curr Gene Ther *8*, 483-488.

Singh, S., Clarke, I., Terasaki, M., and Bonn, V. (2003). Identification of a cancer stem cell in human brain tumors. Cancer Research.

Singh, S.K., Hawkins, C., Clarke, I.D., Squire, J.A., Bayani, J., Hide, T., Henkelman, R.M., Cusimano, M.D., and Dirks, P.B. (2004). Identification of human brain tumour initiating cells. Nature *432*, 396-401.

Singh, S.K., Manoranjan B, Venugopal C (2013). Evolution of brain tumor-initiating cell research: in pursuit of a moving target. Future Neurology *8*, 1-3.

Smith, C.E., Long, D.M., Jones, T.K., Jr., and Levitt, S.H. (1973). Experiences in treating medulloblastoma at the University of Minnesota hospitals. Radiology *109*, 179-182.

Smith, N.A., Singh, S.P., Wang, M.B., Stoutjesdijk, P.A., Green, A.G., and Waterhouse, P.M. (2000). Total silencing by intron-spliced hairpin RNAs. Nature *407*, 319-320.

Srinivasan, S., Patric, I.R., and Somasundaram, K. (2011). A ten-microRNA expression signature predicts survival in glioblastoma. PLoS One *6*, e17438.

Stolzenburg, S., Rots, M.G., Beltran, A.S., Rivenbark, A.G., Yuan, X., Qian, H., Strahl, B.D., and Blancafort, P. (2012). Targeted silencing of the oncogenic transcription factor SOX2 in breast cancer. Nucleic Acids Res *40*, 6725-6740.

Stupp, R., Hegi, M.E., Mason, W.P., van den Bent, M.J., Taphoorn, M.J., Janzer, R.C., Ludwin, S.K., Allgeier, A., Fisher, B., Belanger, K.*, et al.* (2009). Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. Lancet Oncol *10*, 459-466.

Sturm, D., Bender, S., Jones, D.T., Lichter, P., Grill, J., Becher, O., Hawkins, C., Majewski, J., Jones, C., Costello, J.F.*, et al.* (2014). Paediatric and adult

glioblastoma: multiform (epi)genomic culprits emerge. Nat Rev Cancer *14*, 92-107.

Subapanditha, M.K., Adile, A.A., Venugopal, C., and Singh, S.K. (2019). Flow Cytometric Analysis of Brain Tumor Stem Cells. Methods Mol Biol *1869*, 69-77.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S.*, et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Suva, M.L., Rheinbay, E., Gillespie, S.M., Patel, A.P., Wakimoto, H., Rabkin, S.D., Riggi, N., Chi, A.S., Cahill, D.P., Nahed, B.V.*, et al.* (2014). Reconstructing and reprogramming the tumor-propagating potential of glioblastoma stem-like cells. Cell *157*, 580-594.

Suva, M.L., Riggi, N., Janiszewska, M., Radovanovic, I., Provero, P., Stehle, J.C., Baumer, K., Le Bitoux, M.A., Marino, D., Cironi, L.*, et al.* (2009). EZH2 is essential for glioblastoma cancer stem cell maintenance. Cancer Res *69*, 9211-9218.

Swartling, F.J., Grimmer, M.R., Hackett, C.S., Northcott, P.A., Fan, Q.W., Goldenberg, D.D., Lau, J., Masic, S., Nguyen, K., Yakovenko, S.*, et al.* (2010). Pleiotropic role for MYCN in medulloblastoma. Genes Dev *24*, 1059-1072.

Tabor, M.H., Clay, M.R., Owen, J.H., Bradford, C.R., Carey, T.E., Wolf, G.T., and Prince, M.E. (2011). Head and neck cancer stem cells: the side population. Laryngoscope *121*, 527-533.

Tait, D.M., Thornton-Jones, H., Bloom, H.J., Lemerle, J., and Morris-Jones, P. (1990). Adjuvant chemotherapy for medulloblastoma: the first multi-centre control trial of the International Society of Paediatric Oncology (SIOP I). Eur J Cancer *26*, 464-469.

Tavernarakis, N., Wang, S.L., Dorovkov, M., Ryazanov, A., and Driscoll, M. (2000). Heritable and inducible genetic interference by double-stranded RNA encoded by transgenes. Nat Genet *24*, 180-183.

Taxman, D.J., Livingstone, L.R., Zhang, J., Conti, B.J., Iocca, H.A., Williams, K.L., Lich, J.D., Ting, J.P., and Reed, W. (2006). Criteria for effective design, construction, and gene knockdown by shRNA vectors. BMC Biotechnol *6*, 7.

Taylor, M.D., Northcott, P.A., Korshunov, A., Remke, M., Cho, Y.J., Clifford, S.C., Eberhart, C.G., Parsons, D.W., Rutkowski, S., Gajjar, A.*, et al.* (2012). Molecular subgroups of medulloblastoma: the current consensus. Acta Neuropathol *123*, 465-472.

Taylor, R.E., Bailey, C.C., Robinson, K., Weston, C.L., Ellison, D., Ironside, J., Lucraft, H., Gilbertson, R., Tait, D.M., Walker, D.A.*, et al.* (2003). Results of a randomized study of preradiation chemotherapy versus radiotherapy alone for nonmetastatic medulloblastoma: The International Society of Paediatric Oncology/United Kingdom Children's Cancer Study Group PNET-3 Study. J Clin Oncol *21*, 1581-1591.

Thompson, M.C., Fuller, C., Hogg, T.L., Dalton, J., Finkelstein, D., Lau, C.C., Chintagumpala, M., Adesina, A., Ashley, D.M., Kellie, S.J.*, et al.* (2006). Genomics identifies medulloblastoma subgroups that are enriched for specific genetic alterations. J Clin Oncol *24*, 1924-1931.

Tian, Q., Sangar, V., and Price, N.D. (2016). Emerging Proteomic Technologies Provide Enormous and Underutilized Potential for Brain Cancer Research. Mol Cell Proteomics *15*, 362-367.

Till, J.E., and McCulloch, E.A. (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat Res *14*, 213-222.

Tirosh, I., Venteicher, A.S., Hebert, C., Escalante, L.E., Patel, A.P., Yizhak, K., Fisher, J.M., Rodman, C., Mount, C., Filbin, M.G.*, et al.* (2016). Single-cell RNA-seq supports a developmental hierarchy in human oligodendroglioma. Nature *539*, 309-313.

Tiscornia, G., Tergaonkar, V., Galimi, F., and Verma, I.M. (2004). CRE recombinase-inducible RNA interference mediated by lentiviral vectors. Proc Natl Acad Sci U S A *101*, 7347-7351.

Toda, M., Iizuka, Y., Yu, W., Imai, T., Ikeda, E., Yoshida, K., Kawase, T., Kawakami, Y., Okano, H., and Uyemura, K. (2001). Expression of the neural RNA-binding protein Musashi1 in human gliomas. Glia *34*, 1-7.

Toledo, C.M., Ding, Y., Hoellerbauer, P., Davis, R.J., Basom, R., Girard, E.J., Lee, E., Corrin, P., Hart, T., Bolouri, H.*, et al.* (2015). Genome-wide CRISPR-Cas9 Screens Reveal Loss of Redundancy between PKMYT1 and WEE1 in Glioblastoma Stem-like Cells. Cell Rep *13*, 2425-2439.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nat Methods *13*, 731-740.

Ui-Tei, K., Naito, Y., Nishi, K., Juni, A., and Saigo, K. (2008a). Thermodynamic stability and Watson-Crick base pairing in the seed duplex are major determinants of the efficiency of the siRNA-based off-target effect. Nucleic Acids Res *36*, 7100-7109.

Ui-Tei, K., Naito, Y., Zenno, S., Nishi, K., Yamato, K., Takahashi, F., Juni, A., and Saigo, K. (2008b). Functional dissection of siRNA sequence by systematic

DNA substitution: modified siRNA with a DNA seed arm is a powerful tool for mammalian gene silencing with significantly reduced off-target effect. Nucleic Acids Res *36*, 2136-2151.

Ule, J., Jensen, K., Mele, A., and Darnell, R.B. (2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. Methods *37*, 376-386.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. Science *302*, 1212-1215.

Ullal, A.V., Peterson, V., Agasti, S.S., Tuang, S., Juric, D., Castro, C.M., and Weissleder, R. (2014). Cancer cell profiling by barcoding allows multiplexed protein analysis in fine-needle aspirates. Sci Transl Med *6*, 219ra219.

UniProtKB (2017). MSI1H_HUMAN (uniprot.org).

Uren, P.J., Vo, D.T., de Araujo, P.R., Potschke, R., Burns, S.C., Bahrami-Samani, E., Qiao, M., de Sousa Abreu, R., Nakaya, H.I., Correa, B.R.*, et al.* (2015). RNA-Binding Protein Musashi1 Is a Central Regulator of Adhesion Pathways in Glioblastoma. Mol Cell Biol *35*, 2965-2978.

van Nostrand, E.L., Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen J-Y, Cody NAL, Dominguez D, Olsen S, Sundararaman B, Zhan L, Bazile C, Bouvrette LPB, Bergalet J, Duff MO,Garcia KE, Gelboin-Burkhart C, Hochman M, Lambert NJ, Li H, Nguyen TB, Palden T, Rabano I, Sathe S, Stanton R, Su A, Wang R, Yee BA, Zhou B, Louie AL, Aigner S, Fu X-D, Lecuyer E, Burge CB, Graveley BR, Yeo GW. (2018). A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. bioRxiv.

Van Nostrand, E.L., Nguyen, T.B., Gelboin-Burkhart, C., Wang, R., Blue, S.M., Pratt, G.A., Louie, A.L., and Yeo, G.W. (2017). Robust, Cost-Effective Profiling of RNA Binding Protein Targets with Single-end Enhanced Crosslinking and Immunoprecipitation (seCLIP). Methods Mol Biol *1648*, 177-200.

Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K.*, et al.* (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat Methods *13*, 508-514.

van Steensel, B., and Henikoff, S. (2000). Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. Nat Biotechnol *18*, 424-428.

Vanner, R.J., Remke, M., Gallo, M., Selvadurai, H.J., Coutinho, F., Lee, L., Kushida, M., Head, R., Morrissy, S., Zhu, X.*, et al.* (2014). Quiescent sox2(+) cells drive hierarchical growth and relapse in sonic hedgehog subgroup medulloblastoma. Cancer Cell *26*, 33-47.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A.*, et al.* (2001). The sequence of the human genome. Science *291*, 1304-1351.

Verhaak, R.G., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P.*, et al.* (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. Cancer Cell *17*, 98-110.

Vizcaino, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T.*, et al.* (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res *44*, 11033.

Vo, D.T., Qiao, M., Smith, A.D., Burns, S.C., Brenner, A.J., and Penalva, L.O. (2011). The oncogenic RNA-binding protein Musashi1 is regulated by tumor suppressor miRNAs. RNA Biol *8*, 817-828.

Vo, D.T., Subramaniam, D., Remke, M., Burton, T.L., Uren, P.J., Gelfond, J.A., de Sousa Abreu, R., Burns, S.C., Qiao, M., Suresh, U.*, et al.* (2012). The RNA-binding protein Musashi1 affects medulloblastoma growth via a network of cancer-related genes and is an indicator of poor prognosis. Am J Pathol *181*, 1762-1772.

Vogel, C., Abreu Rde, S., Ko, D., Le, S.Y., Shapiro, B.A., Burns, S.C., Sandhu, D., Boutz, D.R., Marcotte, E.M., and Penalva, L.O. (2010). Sequence signatures and mRNA concentration can explain two-thirds of protein abundance variation in a human cell line. Mol Syst Biol *6*, 400.

Vrzalikova, K., Skarda, J., Ehrmann, J., Murray, P.G., Fridman, E., Kopolovic, J., Knizetova, P., Hajduch, M., Klein, J., Kolek, V.*, et al.* (2008). Prognostic value of Bmi-1 oncoprotein expression in NSCLC patients: a tissue microarray study. J Cancer Res Clin Oncol *134*, 1037-1042.

Wang, E.T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008a). Alternative isoform regulation in human tissue transcriptomes. Nature *456*, 470-476.

Wang, H., Pan, K., Zhang, H.K., Weng, D.S., Zhou, J., Li, J.J., Huang, W., Song, H.F., Chen, M.S., and Xia, J.C. (2008b). Increased polycomb-group oncogene Bmi-1 expression correlates with poor prognosis in hepatocellular carcinoma. J Cancer Res Clin Oncol *134*, 535-541.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature *431*, 873-878.

Wang, S., Garcia, A.J., Wu, M., Lawson, D.A., Witte, O.N., and Wu, H. (2006). Pten deletion leads to the expansion of a prostatic stem/progenitor cell subpopulation and tumor initiation. Proc Natl Acad Sci U S A *103*, 1480-1485.

Ward, R.J., Lee, L., Graham, K., Satkunendran, T., Yoshikawa, K., Ling, E., Harper, L., Austin, R., Nieuwenhuis, E., Clarke, I.D.*, et al.* (2009). Multipotent CD15+ cancer stem cells in patched-1-deficient mouse medulloblastoma. Cancer Res *69*, 4682-4690.

Washburn, M.P., Koller, A., Oshiro, G., Ulaszek, R.R., Plouffe, D., Deciu, C., Winzeler, E., and Yates, J.R., 3rd (2003). Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A *100*, 3107-3112.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis, 2nd edn (Springer International Piublishing).

Wu, G., Feng, X., and Stein, L. (2010). A human functional protein interaction network and its application to cancer data analysis. Genome Biol *11*, R53.

Wu, X., Northcott, P.A., Dubuc, A., Dupuy, A.J., Shih, D.J., Witt, H., Croul, S., Bouffet, E., Fults, D.W., Eberhart, C.G.*, et al.* (2012). Clonal selection drives genetic divergence of metastatic medulloblastoma. Nature *482*, 529-533.

Wu, Z., Wang, Q., Wang, L., Li, G., Liu, H., Fan, F., Li, Z., Li, Y., and Tu, Y. (2013). Combined aberrant expression of Bmi1 and EZH2 is predictive of poor prognosis in glioma patients. J Neurol Sci *335*, 191-196.

Xia, D., Reardon, D.A., Bruce, J.L., and Lindeman, N.I. (2016). The Clinical Implications of Inconsistently Methylated Results from Glioblastoma MGMT Testing by Replicate Methylation-Specific PCR. J Mol Diagn *18*, 864-871.

Xu, Q., Modrek, B., and Lee, C. (2002). Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. Nucleic Acids Res *30*, 3754-3766.

Yang, Z.J., Ellis, T., Markant, S.L., Read, T.A., Kessler, J.D., Bourboulas, M., Schuller, U., Machold, R., Fishell, G., Rowitch, D.H.*, et al.* (2008). Medulloblastoma can be initiated by deletion of Patched in lineage-restricted progenitors or stem cells. Cancer Cell *14*, 135-145.

Yeang, C.H., McCormick, F., and Levine, A. (2008). Combinatorial patterns of somatic gene mutations in cancer. FASEB J *22*, 2605-2622.

Yoo, H.J., Kim, H., Park, H.J., Kim, D.S., Ra, Y.S., and Shin, H.Y. (2016). Neurocognitive Function and Health-Related Quality of Life in Pediatric Korean Survivors of Medulloblastoma. J Korean Med Sci *31*, 1726-1734.

Young, R.A. (2011). Control of the embryonic stem cell state. Cell *144*, 940-954.

Zomerman, W.W., Plasschaert, S.L.A., Conroy, S., Scherpen, F.J., Meeuwsen-de Boer, T.G.J., Lourens, H.J., Guerrero Llobet, S., Smit, M.J., Slagter-Menkema, L., Seitz, A.*, et al.* (2018). Identification of Two Protein-Signaling States Delineating Transcriptionally Heterogeneous Human Medulloblastoma. Cell Rep *22*, 3206-3216.