Functional characterization of recurrent *HNRNPU* mutations in *MYC*-driven aggressive lymphomas

by

Qurat Ul Ain Qureshi

BSc (Hons., Molecular Biology and Biochemistry), Simon Fraser University, 2020

Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science

in the
Department of Molecular Biology and Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY
Fall 2023

Declaration of Committee

Name:	**Qurat Ul Ain Qureshi**
Degree:	**Master of Science**
Title:	**Functional characterization of recurrent *HNRNPU* mutations in *MYC*-driven aggressive lymphomas**
Committee:	**Chair: Valentin Jaumouillé** Assistant Professor, Molecular Biology and Biochemistry
	Timothy E. Audas Co-Supervisor Associate Professor, Molecular Biology and Biochemistry
	Ryan D. Morin Co-supervisor Professor, Molecular Biology and Biochemistry
	Esther Verheyen Committee Member Professor, Molecular Biology and Biochemistry
	Sharon Gorski Examiner Professor, Molecular Biology and Biochemistry

Abstract

hnRNPU is a ubiquitously expressed, pleiotropic DNA and RNA binding protein involved in RNA metabolism. It is recurrently mutated in Burkitt and high-grade B-cell lymphomas, but the functional consequence of this recurrence is unknown. Here, I show that heterozygous *HNRNPU* nonsense mutations are enriched in lymphomas with *MYC* translocations. The results imply that *HNRNPU* is haploinsufficient and mutations of a single allele promote cell cycle entry by altering the gene expression and splicing landscape of the B cells that harbor them. This reduced hnRNPU expression lowers MYC levels, possibly to buffer MYC-induced proteotoxic stress and shifts dependence from MYC to E2Fs to maintain cell cycle promotion. Finally, I show that, owing to this dependence on E2Fs for cell cycle progression, *HNRNPU* mutated lymphomas are more sensitive to E2F inhibitors. These results highlight hnRNPU-mediated regulation of MYC and its downstream effects as possible new avenues for therapeutic intervention in MYC-driven lymphomas.

Keywords: Aggressive lymphoma; Cancer genomics; RNA binding proteins; MYC; hnRNPU

Dedication

To all patients who participated in this project, whose samples will continue to improve cancer research and patient care.

And to my family, for their unwavering love and support.

Acknowledgements

The work outlined in this thesis would not have been achievable without the exceptional work environments I had the privilege to be a part of. I am grateful for the opportunity to collaborate with extraordinary scientists in the department of Molecular Biology and Biochemistry at Simon Fraser University and the Centre for Lymphoid Cancer at the BC Cancer Research Centre. The collaborative culture has not only positively impacted this project but has also significantly contributed to my development as a researcher.

I would like to express my deepest gratitude to my supervisors, Dr. Tim Audas and Dr. Ryan Morin, whose guidance, encouragement, and support was instrumental in this research. Your insightful feedback and invaluable expertise have shaped not only this thesis but also my academic growth.

My fellow lab members, I am grateful for your encouragement, discussions, and occasional moments of respite, which have provided much-needed balance and motivation. I want to extend my gratitude to Haya for being a constant motivator, Chris for his invaluable awk and R expertise, to Lionel and Eva – our conversations while splitting cells have been the best part of my lab days and to Krysta and Kostia for always being there whenever I hit a roadblock or sought guidance.

I extend my deepest gratitude to my friends and family for their love, encouragement, and understanding throughout this challenging period. A special thanks goes to my parents for their sacrifices and unwavering support, doing everything in their power to make sure I had the opportunity to pursue my dreams. To my husband, Mohammad, thank you for patiently listening to my frustrations about failed experiments, helping rehearse my presentations and celebrating every little win. Your endless support and belief in my abilities have been my pillar of strength.

Lastly, I extend my appreciation to all those whose names may not appear here but who, in various ways, have supported and encouraged me on this academic journey. This thesis stands as a culmination of collective support and collaboration, and for that, I am profoundly grateful.

Table of Contents

List of Figures

List of Acronyms

ABC DLBCL	Activated B Cell-like Diffuse Large B-cell Lymphoma
ADAR	Adenosine Deaminase Acting on RNA
AID	Activation Induced (cytidine) Deaminase
aSHM	Aberrant Somatic Hypermutation
BL	Burkitt Lymphoma
CDK	Cyclin Dependent Kinase
CSR	Class Switch Recombination
ddPCR	Digital Droplet Polymerase Chain Reaction
DLBCL	Diffuse Large B-cell Lymphoma
EBV	Epstein Barr Virus
eCLIP	enhanced Crosslinking Immunoprecipitation
eGFP	enhanced Green Fluorescent Protein
FACS	Fluorescence Activated Cell Sorting
FL	Follicular Lymphoma
GAMBL	Genomic Analysis of Mature B-cell Lymphomas
GC B cells	Germinal Centre B cells
HGBL-DH/TH	Double-hit or Triple-hit High-Grade B-cell Lymphomas
hnRNP	Heterogenous Nuclear Ribonuclear Proteins
IgV	Immunoglobulin Variable Domain
MCL	Mantle Cell Lymphoma
NHL	Non-Hodgkin Lymphoma
NMD	Nonsense Mediated Decay
PBL	Plasmablastic Lymphoma
qPCR	quantitative real-time PCR
RBP	RNA Binding Protein
SHM	Somatic Hypermutation
UTR	Untranslated Region
UPR	Unfolded Protein Response
WT	Wild Type

Chapter 1. Introduction

1.1. Cancer

Cancer refers to a collection of more than 200 diseases that are caused by the accumulation of somatic mutations and are unified by features such as abnormal cell proliferation and the potential of local invasion or metastasis to distant tissues. The earliest documented reference to human cancer dates to 3000 BC, in ancient Egypt. The text (Edwin Smith Papyrus) describes cancer as an incurable disease[1]. In our contemporary landscape, although disease outcomes have improved, cancer continues to impose a significant global health burden. It is estimated that 1 in 2 Canadians are expected to receive a cancer diagnosis in their lifetime and despite the many advances in our understanding and treatment of the disease, cancer is the leading cause of death in Canada[2].

1.1.1. Hallmarks of cancer

One of the major challenges in characterizing and treating cancer is that different cancer types exhibit a remarkable diversity in their cellular origins, genetic backgrounds, and mutations that drive them. The vast complexity of cancer phenotypes and genotypes can be distilled into a small number of underlying principles shared by almost all cancers known as “the hallmarks of cancer”[3]. Since its conception, the hallmarks of cancer have expanded to detail ten hallmark capabilities and four enabling characteristics outlined in **Figure 1-1**[4].

The hallmark features define the functional capabilities required by human cells for them to transform from a normal to neoplastic state. These need not be acquired in every type of cancer, and indeed some of these are features of certain healthy cell types. Nonetheless, these collectively represent a comprehensive view of the core mechanisms that underlie the ability of cancer cells to proliferate, evade normal regulatory mechanisms, sustain their metabolic and nutrient requirements, and spread throughout the body. Beyond the core hallmarks, enabling characteristics such as the acquisition of mutations and microbial infection can provide opportunity for certain cells to achieve the hallmarks of cancer. Understanding the hallmarks and the enabling

characteristics have been crucial for the development of targeted therapies and is a main objective of cancer research.

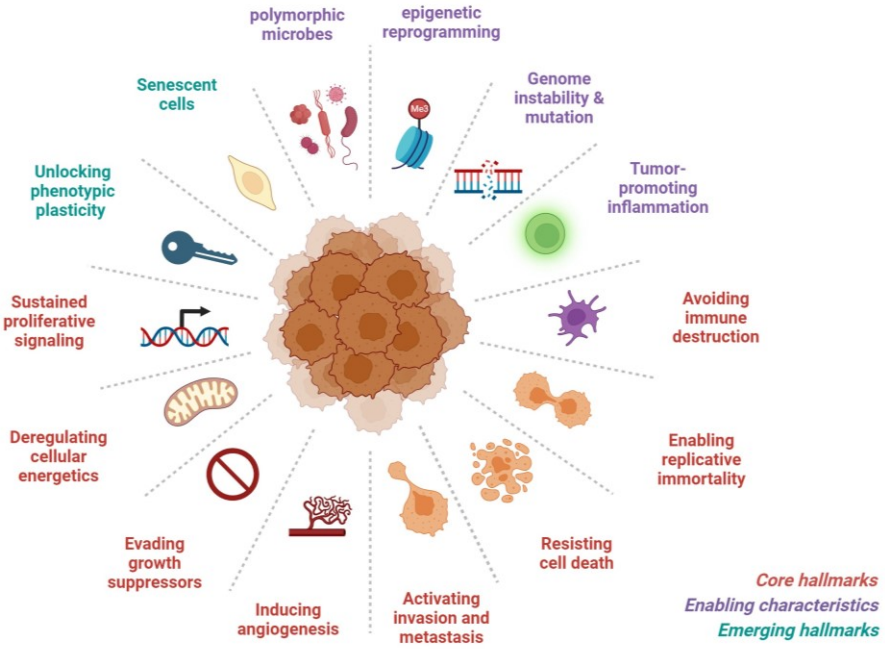


Figure 1-1. Hallmarks of cancer.
Adapted from Hanahan and Weinberg (2021). Created with Biorender.com

1.1.2. Driver and passenger mutations in cancer evolution

Cancer development is an evolutionary process among individual somatic cells where genomic alterations are acquired and undergo selection based on their influence on the fitness of cells in their microenvironment[5,6]. The eventual accumulation of a sufficient combination of such alterations drives the transformation of normal human cells into clonal populations of neoplastic (cancer) cells. In recent years, high throughput sequencing technologies have enabled the detection of thousands of mutations in single samples and in large cohorts, revolutionizing the ability of researchers to comprehensively identify somatic mutations in a tumour biopsy[7]. The resulting sequencing datasets however are large and complex, obscuring the clinically important

mutations in a background of errors, noise, and random mutations that did not contribute to oncogenesis[8].

In general, a small minority of the clonal mutations in cancer have conferred a selective growth advantage to, and thus experienced positively selection, during evolution of the cancer[9]. By definition, these are known as driver mutations[10]. These mutations exist in a sea of biologically inert somatic mutations (passengers) that coincidentally occur during the mutational processes and are retained in the population only because they cannot be lost during clonal expansion of cells in which they arise. Passenger mutations may affect cellular functions and processes beyond those relevant to cancer development[11]. The identification of driver mutations amidst passenger mutations is crucial for comprehending the molecular mechanisms of carcinogenesis.

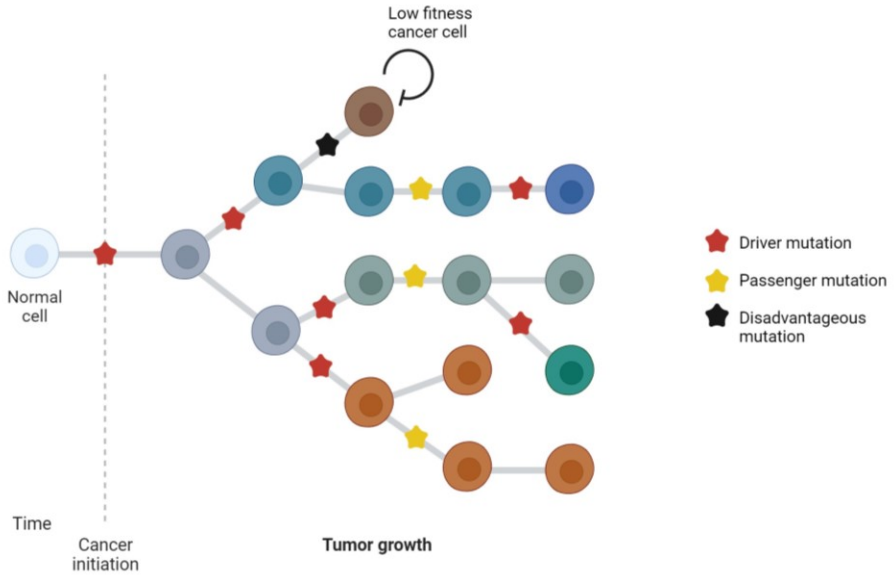


Figure 1-2. Somatic evolution and cancer development with driver and passenger events.
Ongoing mutation and selection cause sub clonal divergence within the tumour clone. Some of these events are neutral (passenger mutations), some are disadvantageous, some are inert, and others, called driver mutations, provide cancer cells with a selective advantage and lead to clonal expansion. Created with Biorender.com

1.1.3. Oncogenes and tumour suppressors

Cancer-associated genes are commonly categorized as either tumour suppressor genes or oncogenes based on the role of their gene product in cancer cells. An oncogene is a cellular gene that is constitutively activated by mutation, structural rearrangement or gene amplifications[12,13]. An activating somatic mutation in one allele of an oncogene is generally sufficient to confer a selective growth advantage to the cell[14]. Often, oncogenes encode proteins that function to stimulate cell division, inhibit cell differentiation, and restrict apoptosis[15]. As an example, the Ras GTPase family proteins are recurrently mutated in cancer where mutations lead to their constitutive activation[16]. These proteins have a role in a number of processes including cell cycle progression, cell growth and migration. In contrast, a tumour suppressor gene encodes a protein that acts to regulate cell division. When a tumour suppressor gene is inactivated, it contributes to tumour growth by inactivating proteins that normally act to limit cell proliferation[17]. As an example, the tumor suppressor p53 plays a role in induction of cell cycle arrest and/or apoptosis upon detecting DNA damage[18]. Loss of function mutations in the p53 gene is observed in nearly 50% of all cancers[19]. Tumour-suppressor genes are targeted in the opposite way by genetic alterations: mutations reduce the activity of the gene product. Such inactivation may arise from missense mutations at residues that are essential for its activity, from mutations that result in a premature termination codon and degradation of the RNA by the non-sense mediated decay pathway, from deletions or insertions of various sizes, or from epigenetic silencing[20–23]. Some recently described tumour-suppressor genes have been hypothesized to exert a selective advantage on a cell when only one allele is inactivated and the other remains functional (a genetic scenario known as haploinsufficiency)[24].

1.1.4. Cancer classification

Internationally accepted classifications of malignant tumours are currently based on the histotype, site of origin, morphologic grade, and spread of cancer throughout the body[25]. Histologically, cancers can be broadly grouped into five major categories, which include carcinoma, sarcoma, myeloma, leukemia, and lymphoma. Carcinoma specifically pertains to cancers that originate from epithelial tissues, which form the internal or external linings of the body[26]. These are the most prevalent cancer types, constituting approximately 80-90 percent of all cancer cases. Sarcomas, on the other hand, originate

from the supportive or connective tissues, such as bone or muscle[27]. Other major types are hematological malignancies. Leukemias are a group of hematological malignancies that arise from the dysregulated proliferation of developing leukocytes. It's noteworthy that unlike solid tumours, which often necessitate genetic alterations and complex cellular reprogramming for effective metastatic spreading, leukemic cells are known as “liquid tumours” and possess a unique inherent ability for migration and invasion[28]. Furthermore, lymphomas are another type of hematological malignancy that develop as solid tumours in the glands or nodes of the lymphatic system, a network of vessels, nodes, and organs that develop lymphocytes[29]. Due to presence of these tumors in the lymphatic system, increased lymphatic dissemination via functionally contiguous lymph nodes is often observed, limiting requirements for metastatic spread[30]. Lastly, myeloma originates in plasma cells, responsible for antibody production.

While conventional tumour categorizations primarily centre on the source tissue or organ, as well as histopathological, clinical, and epidemiological information, WHO classifications have been incorporating molecular-genetic attributes of tumours since the third edition in 2000[31]. The integration of molecular pathology features is somewhat restrained due to the scarcity of diagnostic capabilities in laboratories worldwide. Nevertheless, these advancements are significantly influencing the diagnosis and treatment of cancer.

1.2. Lymphoma

Lymphomas are a heterogeneous group of malignancies derived from the basic cells of the lymphoid tissue, the lymphocytes, at different stages of their maturation[32]. They are the most common hematological malignancy worldwide and are the 5th most common form of cancer in Canada, with 11,400 new diagnosis in 2022 alone[33]. Due to the diverse biological and clinical behavior of lymphoma subtypes, response to treatment is generally variable and despite advances in treatment, many subtypes of lymphoma remain incurable with current management strategies[34]. This underscores the importance of understanding the pathobiology of these lymphomas to develop novel diagnostic, prognostic and/or therapeutic strategies for these diseases.

1.2.1. Lymphoma subtypes and their cell of origin

Lymphomas are broadly classified into two groups, Hodgkin lymphoma and Non-Hodgkin lymphoma (NHL) accounting for roughly 10% and 90% of all lymphoma diagnoses respectively[35]. Hodgkin lymphomas arise from B-cell precursors and are typically characterized by the presence of large, multinucleated Reed/Sternberg cells[36]. Most cases of Hodgkin lymphomas are chemo sensitive and the overall survival is about 86%[35]. NHLs are stratified broadly by cell origin, namely B-cell, T-cell, or natural killer cell origin[37]. Each of these are further divided into different subtypes by their distinct morphology, immunophenotype, genetic, molecular, and clinical features[38].

Approximately 85% of NHLs are of B-cell origin, coopting the regulatory biologic features of their normal B cell counterparts and shape it to execute their oncogenic purposes[39]. They can be broadly divided into low-grade/indolent and high grade/aggressive B-cell lymphomas. Indolent lymphomas are slow growing and usually responsive to many treatment modalities, however their protracted nature results in a long clinical course and are usually incurable[40–42]. Follicular lymphoma (FL), is the most common indolent B-cell lymphoma, accounting for 22.1% of all NHL cases worldwide[43]. FL tumourigenesis starts in precursor B cells and becomes full-blown tumour when the cells reach the germinal centre maturation step[44].

In contrast to the slow growth and long outcomes of indolent lymphomas, high grade/aggressive lymphomas are characterized by rapidly growing and generally aggressive tumours with inferior patient outcomes if left untreated but are generally responsive to therapy[45,46]. The most common aggressive lymphoma, Diffuse large B-cell lymphoma (DLBCL), accounts for nearly 40% of all NHL diagnoses and has a heterogeneous clinical presentation and response to chemotherapy[46]. The two dominant subtypes of DLBCL are classified based on their gene expression profile. Germinal centre B cell–like (GCB) DLBCL, harbour the gene expression program of normal germinal centre (GC) B cells, and activated B cell–like (ABC) DLBCL, express genes that are characteristically induced following B cell receptor engagement and activation of normal B cells[47]. Burkitt lymphoma (BL) is another highly aggressive B-cell NHL and the most common pediatric NHL[48]. BL is the fastest growing human tumour and also originates from germinal centre B cells[49]. Other aggressive B cell NHLs are rare and

include Mantle Cell lymphoma (MCL) originating from the mantle zone of the germinal centre, and plasmablastic lymphoma (PBL)[50,51].

1.2.2. Role of the germinal centre reaction in the development of B-cell lymphomas

Most of the aforementioned B cell neoplasms originate from GC B cells as depicted in **Figure 1-3**. The germinal centre is a transient microanatomical structure within the follicles of secondary lymphoid tissues[52]. Here, mature B cells undergo repeated rounds of clonal expansion, genetic diversification, and selection of their immunoglobulin genes. This ultimately results in the generation of high-affinity, clonally expanded B cells destined to become either memory B cells or plasma cells[53]. These processes occur in distinct regions of the germinal centre known as the dark zone and the light zone. After proliferative expansion and genetic diversification in the dark zone, B cells move to the light zone where those that have high affinity for the antigen are selected for survival. One of the unusual features of GC B cells is that they manifest phenotypic features that mimic many of the hallmarks of cancer. Some of these include:

Massive proliferation and clonal expansion: GC B cells are highly proliferative, with cell cycles that can be as short as 5–6 h[54]. Cell proliferation is initiated by transient induction of the transcription factor MYC early in the GC reaction[55]. MYC expression is then reactivated in a very small subset of light zone B cells undergoing positive selection, primed for re-entry into the dark zone to undergo further cycles of proliferation[56]. Transcription factor 3 (TCF3) activates cell proliferation via the cell cycle regulator, E2F, when MYC expression is transiently lowered[52]. The proliferation is also maintained through the actions of the transcriptional repressor EZH2, which represses proliferation checkpoint genes[57,58].

Resisting cell death: Checkpoint genes, specifically those involved in sensing and responding to DNA damage (such as TP53) are directly repressed in GC-B cells through another transcriptional repressor, BCL6[59,60].

Genome instability: In the germinal centre, antigen stimulated B cells undergo a DNA mutational process in their immunoglobulin variable (IgV) region to generate higher affinity B-cell receptors to the antigen of interest. This process, called somatic hypermutation (SHM), is initiated by cytosine deamination catalyzed by the activation

induced (cytidine) deaminase enzyme (AID) in the dark zone of the germinal centre. SHM causes point mutations at very high rates (10^{-3} bp per cell generation) which is 10^6 fold higher than the spontaneous mutation rate in somatic cells[61,62]. Through error prone repair pathways this process results in a characteristic pattern of mutations in the IgV regions where some of these increase the affinity of the antibody. SHM is also known to physiologically target non-Ig genes such as BCL6, albeit at lower rates[63,64]. In addition to point mutations in the variable region, procession of AID-induced lesions can produce double-stranded DNA breaks in switch regions of Ig genes. This triggers class-switch recombination from IgM to downstream IgG, IgE, or IgA in the light zone of the germinal centre[65,66].

One key aspect of AID biology is the balance between mutagenic diversity and genomic integrity. Most human B-cell NHLs originate from GC or post-GC B cells, as demonstrated by their somatically mutated immunoglobulin genes[67]. Aberrant regulation of somatic hypermutation (aSHM), in part through dysregulated AID activity, has been reported to cause widespread alterations of genes that are not physiological SHM targets[68–70]. AID-related mutations occur at a frequency of ~8% in hematological malignancies, many of which have putative driver roles in lymphoma[71]. Some aSHM targets are proto-oncogenes, such as MYC and BCL6, which become susceptible to chromosomal translocations in B-cell NHLs which is thought to be a byproduct of SHM induced double-strand breaks[68]. Aberrant regulation of class switch recombination has also been described to predispose B cells to chromosomal translocations[72]. Translocations involving the immunoglobulin loci (IGH, IGL, IGK) juxtaposed with oncogenes have been associated with most types of malignant B-cell lymphomas with some regarded as hallmark characteristics of the lymphoma subtype[73].

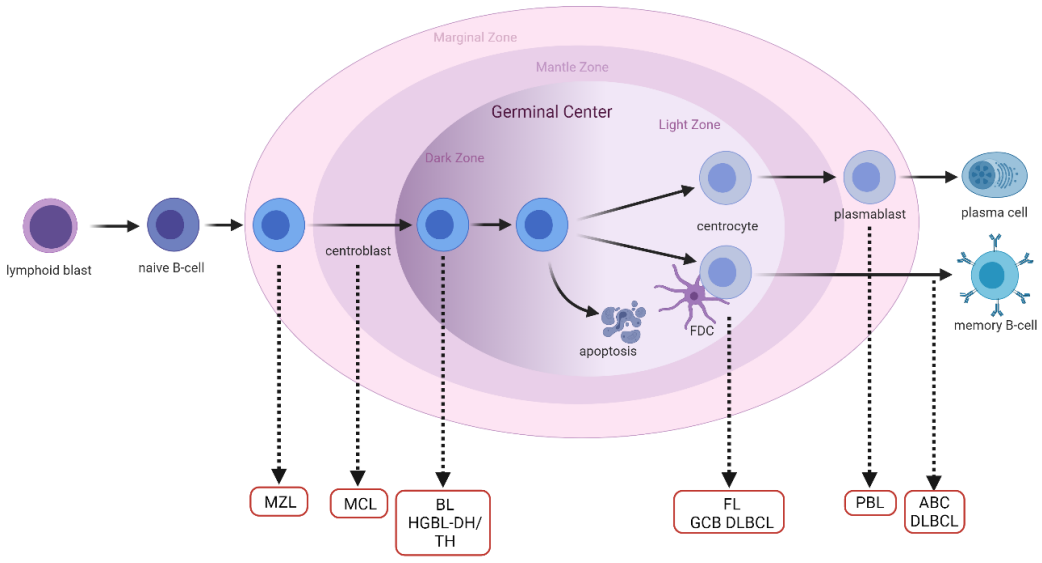


Figure 1-3. Major B-cell non-Hodgkin lymphoma subtypes and their cell of origin. Illustrated here are the fundamental stages in the later development of B-cell differentiation. This process commences with a lymphoid blast exiting the bone marrow and proceeding through the germinal centre (GC), where it undergoes either positive or negative selection. If successful, it ultimately matures into a plasma cell or memory B-cell. Each of the several prevalent types of mature B-cell cancers is highlighted within a red box, originating from the lymphoid cell from which they are most likely derived. Adapted from Morin et al. (2021). Created with biorender.com.

1.3. MYC-driven aggressive lymphomas

The proto-oncogene MYC, located on chromosome 8q24, is one of the most frequently dysregulated genes in B-cell NHLs[74]. MYC is a transcription factor of the helix-loop-helix-leucine zipper family and is described to regulate the expression of 15% of all human genes including some intergenic sites[75]. Regulation of transcription by MYC involves the formation of heterodimeric complexes with the MAX protein. The MYC-MAX heterodimer is the active form that binds to specific DNA sequences called E-boxes (canonical sequence CACGTG) in the regulatory regions of target genes[76]. MYC regulates a diverse array of cellular functions, including proliferation, growth, and apoptosis. Due to its central role in human cells, MYC is tightly regulated at both the transcriptional and translational levels and its dysregulation contributes to oncogenesis. MYC-driven tumours display “oncogene addiction” to MYC where its continuing activity is necessary for cancer maintenance and its genetic inhibition relieves the tumour burden[77].

Unlike some common oncogenes, *MYC* dysregulation does not result from mutations that change its function. Instead, *MYC* expression is enhanced through amplification or, more commonly in lymphomas, chromosomal translocations[78]. In contrast to the GC B-cells in which *MYC* expression is transient, B-cell lymphomas with *MYC* translocations have constitutive *MYC* expression. These changes cause the cells to proliferate uncontrollably and promotes angiogenesis, genomic instability, and transformation. In BL, MYC plays a central role in initiating and maintaining tumor growth. *MYC* chromosomal translocations are considered a genetic hallmark of BL but also appear less commonly in other B-cell lymphomas. The most common form of these translocations is a juxtaposition of potent enhancers belonging to the immunoglobulin heavy or light chain loci, with the *MYC* locus[79,80]. *MYC* translocations are also found in about 10% of tumors with DLBCL morphology and are a feature of more aggressive disease[81]. *MYC* rearrangements in tumors with DLBCL morphology are often present in the context of a complex karyotype and the adverse prognosis associated with *MYC* rearrangement is largely derived from concurrent *BCL2* and/or *BCL6* rearrangements[82]. Tumors with *MYC* translocation accompanied by *BCL2* and/or *BCL6* are termed as double-hit or triple-hit high-grade B-cell lymphomas (HGBCL-DH/TH) and have dismal clinical outcomes. These patients are refractory to most chemotherapy regimens and die within the first year of diagnosis[80].

1.3.1. Cell cycle dysregulation in *MYC* driven lymphomas

Promotion of the cell cycle is a major oncogenic feature provided by MYC. Ectopic *MYC* expression in quiescent cells is sufficient to mediate cell cycle entry, whereas inhibition of *MYC* expression causes cells to withdraw from the cell cycle[83]. Progression through the cell cycle phases (G1, S, G2 and M) is under the control of a family of serine/threonine protein kinases called cyclin-dependent kinases (CDKs). The CDKs form complexes with cyclins to activate them and phosphorylate downstream effectors in specific cell cycle phases[84]. D-type cyclins preferentially bind and activate CDK4 and CDK6 at early G1-phase of the cell cycle, leading to the phosphorylation of the retinoblastoma protein (pRB) and the release of the E2Fs which are a large family of transcription factors that regulate cell cycle progression[85,86]. Cyclin E1/2-CDK2 complexes in the late G1-phase further phosphorylate pRB, allowing the expression of E2F target genes required for the transition to S-phase.

MYC promotes cell cycle entry by not only activating or inducing cyclins and CDKs but also through the downregulation of a set of proteins that act as cell-cycle brakes[87]. MYC also promotes the cell cycle entry by directly inducing E2F activities[85,88] (**Figure 1-4**). Induction of E2Fs is considered an essential component of MYC pathways that control cell proliferation and cell fate decisions[89]. Numerous studies have shown extensive E2F crosstalk with MYC signal transduction pathways[90–92]. E2F1, a member of the E2F family, was determined to be transcriptionally activated by MYC. In lymphoma models, Eμ-MYC transgenic mice lacking one or both E2f1 alleles exhibited a slower onset of lymphoma development[90]. Elevated expression of E2F1, is also frequently observed in BL and is thought to collaborate with MYC in BL formation[93]. Reduction of E2F1 inhibits BL tumour formation and decreases their proliferation rate[93]. This demonstrates an indispensable role of E2Fs in MYC driven lymphomagenesis and shows that this interaction is vital to dysregulated cell cycle progression in lymphoma.

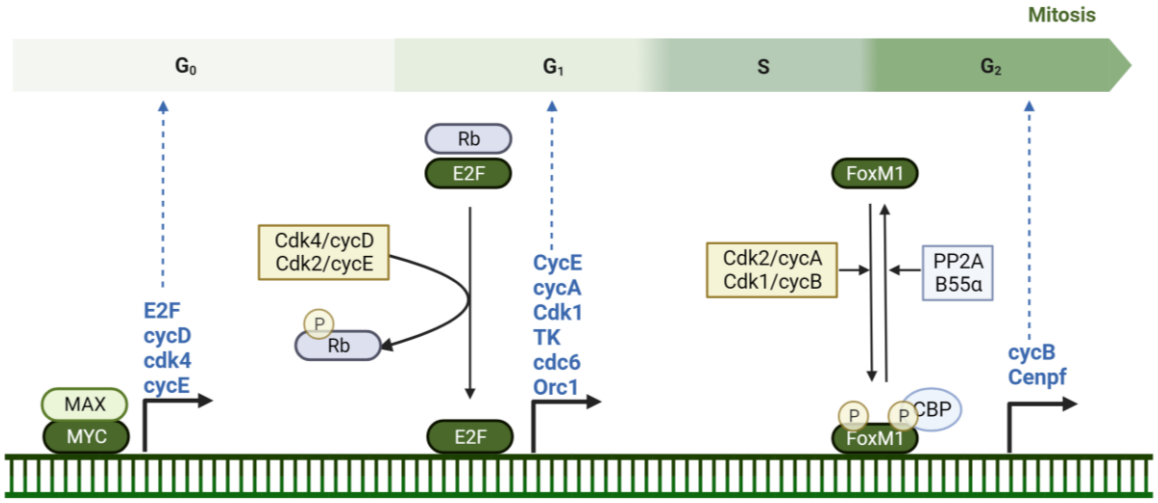


Figure 1-4. MYC stimulates cell cycle entry.
MYC induces expression of genes such as D-types cycles and E2Fs to allow cells to enter the cell cycle[87]. During the G1 phase of the cell cycle, Cdk4/cyclin D (cycD) and Cdk2/cyclin E (cycE) complexes phosphorylate (P) pRb, leading to the activation of E2F proteins and the expression of E2F responsive genes. This cluster of genes encode cell cycle regulators required for G1/S transition [cyclin E, cyclin A (cycA) and Cdk1], enzymes involved in nucleotide biosynthesis [thymidine kinase (TK)] and components of the DNA replication machinery [Cdc6 and origin recognition complex subunit 1 (Orc1)]. During the G2 phase of the cell cycle, Cdk2/cyclin A and Cdk1/cyclin B (cycB) complexes sequentially phosphorylate FoxM1, leading to the relief of its self-inhibition and the recruitment of a histone deacetylase CREB binding protein (CBP) that activates the expression of FoxM1 target genes. This cluster of genes encodes cell cycle regulators required for the execution of mitosis (cyclin B) and interactors of the

kinetochore complex crucial for proper chromosome segregation [centromere protein F (Cenpf)]. Adapted from Lim *et al.* (2013). Created with biorender.com.

1.3.2. Evading cell death in *MYC* driven lymphoma

In recent years, it has become clear that oncogenes that promote cell proliferation often possess intrinsic pro-apoptotic activities[94–96]. For example, the transforming activity of MYC is tempered by its capacity to induce apoptosis under physiological conditions[97,98]. MYC triggers apoptosis by either directly engaging tumour suppressor genes or indirectly by inducing replicative, transcriptional and proteotoxic stress as a consequence of rapid cell proliferation[99,100]. *MYC* dysregulation alone is therefore considered to be insufficient to drive tumourigenesis. The regulation of MYC-induced cell stress in tumour initiation and lymphoma progression, however, is poorly understood. Studies have shown that heightened sensitivity to apoptosis by ectopic MYC expression is observed only in premalignant cells but not after malignant transformation[101]. This demonstrates that tumour cells acquire specific mechanisms for blunting the apoptotic effects that MYC exerts in their normal counterparts.

These mechanisms are interrelated and are thought to involve some aspects of the proapoptotic p53 pathway, the pro-survival BCL2 pathway, or both. Mutations disabling p53, a tumour suppressor that activates pro-apoptotic machinery are frequent in *MYC*-driven lymphomas and are thought to be synergistic with *MYC*, resulting in more aggressive lymphomas[102]. In contrast, BCL2 is a pro-survival factor that blocks apoptotic cell death in lymphocytes and is a frequent target of chromosomal translocations in NHLs[103]. HGBCL-DH-BCL2 have concomitant *MYC* and *BCL-2* translocations with the *BCL2* translocation thought to represent the initiating event. With *MYC* translocations as secondary event, it is likely that BCL2 expression allows these cells to selectively overcome MYC-induced apoptosis[82]. Because BCL2 is not expressed in BL, the collateral damage imposed by MYC in BL must be overcome by other means.

MYC is also able to trigger apoptosis indirectly by inducing proteotoxic stress. Oncogenic activation of *MYC* results in abnormally high protein synthesis rates resulting in proteotoxic stress which if unresolved, leads to apoptosis[104]. In a recent study, heightened proteotoxic stress and active unfolded protein response was observed when *MYC* and *BCL2* were simultaneously overexpressed in primary germinal centre B

cells[105]. To evade this, loss of function mutations in DDX3X, a regulator of ribosome biogenesis and global protein synthesis, have been observed in some MYC-driven lymphomas with the highest abundance in BL[105]. These mutations moderate MYC-driven global protein synthesis, thereby buffering MYC-induced proteotoxic stress during early lymphomagenesis.

It is interesting to consider that MYC-induced stress remains present despite the presence of active pro-survival pathways. Given that *DDX3X* mutations are not universally found in all tumours, nor are *P53* mutations, the question arises: how do cancer cells without these mutations cope with the stress triggered by MYC activation? Interestingly it has been found that one mechanism of evading MYC-induced cell death is moderating *MYC* expression itself. It was recently shown through *in vivo* studies that modest elevation of *MYC* enhances transformation as expected, however, robust overexpression leads to a dramatic increase in apoptosis[106]. The requirement to keep activated oncogenes at low levels to avoid engaging tumour suppression is an important selective pressure governing the early stages of tumour microevolution. Identifying factors that moderate MYC expression are thus important for both understanding the pathobiology of these tumours and to identify potential therapeutic vulnerabilities.

1.3.3. Current treatment strategies for *MYC* driven lymphomas and precision medicine approaches

The current standard of care for most aggressive lymphomas is a chemo-immunotherapy regimen called R-CHOP. It consists of a combination of drugs, incorporating rituximab (a monoclonal anti-CD-20 antibody), along with the chemotherapeutic agents cyclophosphamide, doxorubicin, vincristine, and the steroid prednisone[107]. The response to treatment in *MYC*-driven lymphomas with chemoimmunotherapy is variable. BL remains the most curable with an overall response rate of 89-90% if tumours are detected early, although with significant toxicity[108]. In patients with relapsing disease, the prognosis is typically poor. Patients with *MYC* driven DLBCL or HGBCL-DH/TH have poor overall and progression free survival with many showing primary treatment failure, partial response or relapse after initial treatment with R-CHOP[109]. In light of poor outcomes with *MYC*-related lymphomas, both intensified chemoimmunotherapy regimens and consolidative stem cell transplantation have been evaluated for management.

Efforts to improve upon the R-CHOP backbone have included dose intensification as well as the addition of new agents; the dose-adjusted rituximab, etoposide (a topoisomerase II inhibitor), prednisone, vincristine, cyclophosphamide, and doxorubicin (DA-R-EPOCH) regimen has been identified as a potential replacement for R-CHOP in high-risk cases[107]. Currently, combination chemoimmunotherapy with dose adjusted DA-R-EPOCH is the only treatment regimen that has the potential to induce complete response in a proportion of HGBCL-DH/TH patients however no overall survival advantage in MYC-rearranged lymphomas[110]. Additionally, EPOCH-R is more intensive therefore has a higher toxicity and is less tolerable in older and fragile patients[111]. In general, patients with relapsed/refractory BL or HGBCL have very poor outcomes with an overall response rate of 39% to salvage therapy and a median overall survival of only 2.8 months[112]. The lack of a clear benefit from intensification of therapy (with or without stem cell transplantation) suggests that incorporating novel and targeted agents should be pursued.

Given the lack of favorable chemotherapeutic regimens for *MYC*-driven cancers and increased risk of relapse when treated with standard chemotherapy, targeting the MYC oncoprotein directly has been pursued. There are two major challenges that have prevented successful development and application of MYC inhibitors. Firstly, given that MYC is a ubiquitously expressed general transcription factor, targeting it might not be well tolerated in normal tissues. Secondly, the MYC protein lacks a preferred binding pocket for traditional small molecule drugs[113]. Due to these reasons alternate approaches to target *MYC* dependent neoplasms need to be explored. Some of these include targeting *MYC* gene expression directly or indirectly (transcription, translation, stability), targeting the MYC:MAX interaction or targeting the accessibility of MYC to downstream genes. Since *MYC* dysregulation also incurs cellular stress, deregulated MYC would rewire signal pathways and adapt to such circumstances. Hence, cancer cells could also be selectively killed through targeting critical nodes in these rewired pathways. Finding and exploring such synthetic lethal pathways as well as factors that contribute to *MYC* dysregulation are critical in improving the therapeutic landscape of these lymphomas.

1.4. Gene expression modulation by hnRNPs in lymphoma

RNA processing is a fundamental aspect of gene expression and is critical for transferring genetic information into functional phenotypes. Nearly all RNA molecules are subjected to some form of processing on route to their mature/active form (**Figure 1-5**). Specifically, the immature transcripts of protein-coding genes are spliced, capped, and polyadenylated prior to export into the cytoplasm, where regulation of stability, decay, and translation contribute to the expression of the final gene product. These different aspects of RNA processing are mediated or regulated by many different proteins and protein complexes. There are approximately 1914 human RNA binding proteins (RBPs), accounting for 7.5% of protein-coding genes[114]. Some of these, such as the spliceosome, are well-defined and play consistent roles in varied cell types; however, others, such as the heterogeneous nuclear ribonucleoproteins (hnRNPs) assume a more diverse repertoire that is highly context specific. Mounting evidence have shown that RBPs are involved in various important cellular processes, for instance, cell transport, localization, development, differentiation, and metabolism. Additionally, RBPs engage in almost every step of post-transcriptional regulation, supervise the formation and function of transcripts, and maintain cell homeostasis.

Cancer cells exhibit widespread abnormalities in RNA processing, including driver alterations that functionally contribute to cancer development and progression. Variations in an RBP's expression or localization has the capacity to impact oncogene expression levels or those of tumour-suppressor genes. They can also influence genes important to genome stability. As a result, different transcriptomic and cellular phenotypes arise under the influence of RBP-centered gene regulation, such as differences in proliferation or apoptosis, as well as in other functions like angiogenesis or epithelial to mesenchymal transition. Eventually all these can, in turn, give rise to different profiles of cancer invasion and metastasis as well as different cancer prognoses. It is therefore becoming more and more apparent that RBPs can act as prospective targets for future cancer treatments.

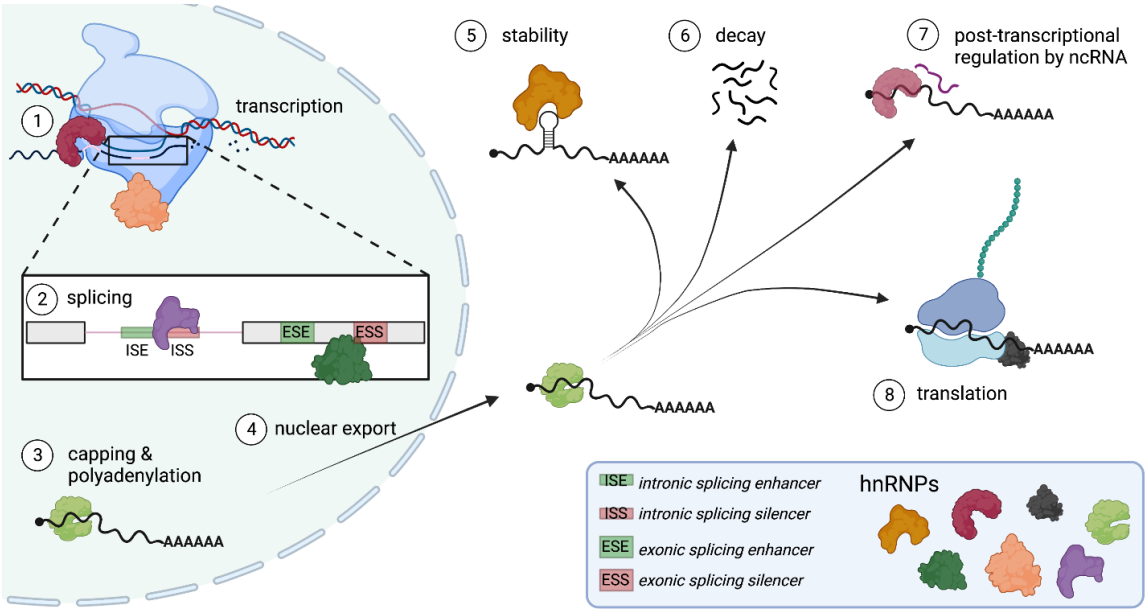


Figure 1-5. hnRNPs are involved throughout the RNA lifecycle.

Varied hnRNPs have been described to play a role in (1) transcriptional regulation and elongation; (2) binding to cis-regulatory elements to affect splicing, and (3) regulation of capping and polyadenylation. Some hnRNPs are also involved in (4) nuclear export and may remain associated with an exported mRNA in the cytoplasm. There, hnRNPs are involved in regulating (5) stability and (6) decay, (7) mediating post-transcriptional regulation with non-coding RNAs, and (8) regulation of translation. hnRNPs: heterogeneous nuclear ribonucleoproteins. Created with biorender.com.

1.4.1. The hnRNP protein family

hnRNPs are a large family of RNA binding proteins (RBPs) that are engaged throughout the RNA lifecycle. These are typically defined as proteins which bind heterogeneous nuclear RNAs that are not components of other ribonucleoprotein (RNP) complexes, such as small nucleolar RNPs.[115] hnRNPs are among the most abundant proteins in the nucleus, with levels similar to histone proteins.[115] Ranging from hnRNP A1 to hnRNP U, these proteins contain one or more RNA binding motifs and additional auxiliary domains, which may contribute to protein:protein interactions and affect subcellular localization[116–118]. The RNA binding motifs found in hnRNP proteins generally

belong to the highly conserved RNA recognition motif (RRM), K-homology (KH), the RGG box, and the poorly conserved quasi-RRM domains. These motifs broadly govern the sequence-specific interactions between hnRNPs and RNA molecules,[119–121] thus small variations in these domains confer diverse sequence specificity to each hnRNP. hnRNPs have the essential function of directing alternative splicing coupled to nonsense-mediated decay (NMD),[122–124] and some hnRNPs can affect the stability and decay of RNA species via additional pathways. hnRNP F, H1, K, L, and U have been shown to stabilize various transcripts, including p53[125] . The stability of the RNA naturally affects its availability for translation; however, hnRNPs can further impact protein levels through translational regulation, including cap-dependent translation,[126,127] elongation,[128] and translation from internal ribosomal entry sites.[129]

The complex and multifaceted roles of hnRNPs in regulating gene expression from transcription through translation makes them particularly important in tissues composed of continually developing and differentiating cells.

1.4.2. The role of hnRNPs in normal B-cell development and differentiation

Mammalian B-cell development is a tightly regulated process controlled by the carefully ordered expression of specific genes that play important roles in their differentiation and cellular functions. The cellular stages of B-cell differentiation have been described in detail however little is known about the molecular mechanisms controlling the various functions in these cells. RNA regulation is thought to play a key role during B cell development mediated in part by RBPs, including several essential hnRNPs.[130] Several studies illustrate that expression of various genes encoding RBPs varies between B-cell subsets, primarily examining naïve, memory, and plasmablasts.[131] Naive and memory B cells have a higher proportion of intron retention than pro- and pre-B cells as well as germinal centre and plasma cells.[132] These differences in gene expression and splicing collectively point to differing requirements for RNA processing in different subsets of B cells.

A key process in the maturation of B cells is diversification of the antibody repertoire. This is accomplished through three mechanisms: somatic V(D)J recombination, somatic hypermutation via AID, and class switch recombination. Several

hnRNPs are involved in these processes. PTBP1 is required for affinity maturation,[133] and both hnRNP K and L are required for the DNA cleavage and repair via end-joining in CSR and somatic hypermutation.[134,135] Other hnRNPs such as hnRNP C, I, and U have also been described as components of AID:RNA complexes.[136] In a recent study, hnRNP U was shown to facilitate CSR by forming and stabilizing the DNA double strand repair ribonucleoprotein complex and preventing excessive R-loop accumulation, which otherwise would cause persistent DNA breaks and aberrant DNA repair, leading to genomic instability[137].

hnRNPs have multi-faceted roles regulating the key genes orchestrating the germinal centre reaction. During this process, B cell survival is dependent on specific regulation of cell death and proliferation. These processes and the necessary homeostasis between B cell types is regulated in part by *MYC, BCL2, and BCL6*.[138–140] Several hnRNPs are known to influence expression of these three genes[141–143]. Cumulatively, hnRNPs appear to play essential roles in regulating key regulators of normal B cell development and differentiation, with specific impacts on genetic diversity, proliferation, and apoptosis. Deregulation of these cellular processes through regulatory mutations, changes in gene dosage, or protein-coding alterations are fundamental contributors to the development of lymphoid cancers. It is thus likely that dysregulation of hnRNPs serves an important role in lymphomagenesis.

1.4.3. hnRNPs in malignant B cells

Perturbations in the RBP-RNA network have been causally associated with many B-cell malignancies. It was recently reported that RBPs play important roles in DLBCL biology and are significantly associated with overall survival.[144] As hnRNPs are thought to regulate key lymphoma drivers such as MYC, BCL2 and BCL6, dysregulation in these interactions is frequently observed in lymphoma. hnRNP A1 was recently reported as an important regulator of MYC expression and tumour cell expansion in multiple myeloma by upregulating MYC through stimulation of its cap-independent translation[145]. hnRNP K is overexpressed in patients lacking MYC genomic alterations, thereby representing an alternate mechanism of MYC pathway activation in B-cell malignancies[146]. *PCBP1* is recurrently mutated in BL where mutations predominantly affect the third KH domain and/or NLS, pointing to a reduced function or complete loss. [147–150] In MCL, mutations in *HNRNPH1* aberrantly upregulate the amount of hnRNPH1 protein, dysregulating the

alternative splicing landscape and resulting in inferior progression free survival[123]. Overexpression of hnRNPL and binding of it to the 3' UTR of BCL2 allows the aberrant *BCL2-IGH* fusion mRNAs to evade NMD and results in BCL2 overexpression[151]. Various other hnRNPs are recurrently mutated in lymphoma but their functional implications in lymphomagenesis remain largely unknown or incomplete underscoring the requirement of mechanistic studies to understand their role in lymphomagenesis. In addition, since their expression and functions are tissue specific, their exact roles in the context of B cells need to be identified.

1.4.4. Discovery of novel hnRNPU mutations in Burkitt lymphoma

hnRNPU is the largest protein in the hnRNP complex and is one of the most abundant and ubiquitously expressed hnRNP proteins[152]. It is a pleiotropic DNA and RNA binding protein that plays various important roles in RNA metabolism and nuclear organisation[153,154]. The domain structure of hnRNPU reflects these different roles, with an N-terminus DNA binding domain, a central actin binding domain, and a C-terminus, intrinsically disordered, RNA binding domain[155]. Our group has recently identified hnRNPU as a significantly mutated gene in Burkitt lymphoma[156]. These mutations occur in a heterozygous manner and are concentrated at the N-terminus of the gene. Moreover, there is an enrichment of mutations that lead to a premature termination codon. This distinctive mutational pattern seems to suggest selective pressure favouring a truncated hnRNPU isoform. However, this does not exclude the possibility that these transcripts are untranslated and instead undergo degradation through nonsense-mediated decay. The discovery of these mutations was novel because *HNRNPU* mutations had never been reported in Burkitt lymphoma. Similar patterns of nonsense mutations or deletions involving the hnRNPU locus are implicated in a rare form of neurodevelopmental disorder known collectively as "*HNRNPU* related neurodevelopmental disorders". hnRNPU loss in these disorders result in mild to severe intellectual disability and developmental delay[157]. The mutations in such disorders are *de novo* and heterozygous consistent with this gene exhibiting hsploinsufficiency[158]. In cancers other than BL, mutations potentially resulting in a loss or reduced function of hnRNPU have not been observed. Conversely, hnRNPU amplification is observed in various cancers, including hepatocellular carcinoma and triple-negative breast cancer, where it is associated with a poor prognosis[159,160]. Due to its seemingly contradictory

effects in different cancers, we aimed to mechanistically understand the impact of hnRNPU nonsense mutations in Burkitt lymphoma.

1.5. Hypothesis

Recurrent *HNRNPU* mutations dysregulate its expression in cells that harbor them. Due to its role in regulating gene expression, these alterations may have a downstream effect on the transcriptome, notably in the expression of key lymphoma drivers. I will address this hypothesis through two overarching aims:

Aim 1: Identify alterations in the transcriptomic and splicing landscape due to *HNRNPU* mutations in B cells.

Aim 2: Mechanistically understand how these alterations contribute to lymphomagenesis.

Copyright statement

Portions of this chapter have been previously published in the following manuscript. The corresponding text has been edited for length, consistency and to include recent findings.

Qureshi, Q. U. A., Audas, T. E., Morin, R. D. & Coyle, K. M. Emerging roles for heterogeneous ribonuclear proteins in normal and malignant B cells. *Biochem. Cell Biol.* **101**, 160–171 (2023).[161]

Chapter 2. Methods

2.1. Sequencing data reanalysis

Data from primary (patient) samples are from an ongoing meta-analysis of mature B-cell neoplasms including unpublished data from the Lymphoma/Leukemia Molecular Profiling Project. My analysis includes whole genome and/or capture sequencing data from six B-cell NHL subtypes. Simple somatic mutations were identified using the SLMS-3 pipeline (relying on Sage, LoFreq, Mutect2, and Strelka2) as previously described[162]. *MYC* rearrangements were identified using the Genome Rearrangement Identification Software Suite (GRIDSS)[163] and Manta[164]. *HNRNPU* shRNA knockdown followed by sequencing and eCLIP-seq data in HEPG2 and K562 cells was obtained from the encode project consortium[165]. CLIP sites in these cell lines were resolved to single nucleotide resolution using PureClip[166].

2.2. Cell culture and reagents

HEK293, Ramos, Raji, JVM2, Mino, SP49, DOHH2 and SU-DHL-4 cell lines were a gift from Dr. Christian Steidl (Lymphoid Cancer, BC Cancer Research Centre). B-cell lymphoma lines were cultured in RPMI 1640 medium (Gibco) supplemented with 10% FBS. HEK293 cells were cultured in DMEM (Gibco) supplemented with 10% FBS. Cell lines were incubated at 37°C and 5% carbon dioxide and passaged every 2-3 days. Protein was extracted from $2x10^6$ cells with Pierce RIPA buffer (ThermoFisher, 78430) containing protease inhibitor (ThermoFisher, P8340) and quantified using the Pierce BCA Protein Assay Kit (ThermoFisher, 23227). 5 µg of protein lysate was resolved on 1.5 mm 12% Tris-Glycine polyacrylamide gels (BioRad, 1610175) and transferred to a nitrocellulose membrane by wet transfer using the Trans-Blot turbo transfer pack (Bio-Rad). Antibodies for hnRNPU (Abcam, ab180952), MYC (Abcam, ab32072), Histone H3 (Cell signaling, 9715S), E2F1 (Abcam, ab137415), E2F2 (Abcam, ab235837) , E2F6 (Abcam, ab155978) , E2F8 (Abcam, ab109596) were diluted according to manufacturer's recommendations. Anti-Rabbit HRP conjugate (Promega, W4018) was used to visualize the bands with Clarity Western ECL Substrate (Biorad, 1705061) on a ChemiDoc digital imager (BioRad). Image quantification was performed with ImageJ.

2.3. Transient transfections

Transient transfections were performed in B-cell lines using the Amaxa nucleofector with Kit V (VCA-1003, Lonza) according to manufacturer's recommendations. Briefly, 2 x 10^6 cells were resuspended in kit V nucleofector solution with either 10 µg of plasmid DNA or 30 pmol of siRNA and electroporated on the M-013 program. Transient transfections in HEK293 cells were performed with Lipofectamine 3000 (L3000001, Invitrogen). 1.5x10^5 cells were seeded on a 6 well plate the day prior to transfections. When cells reached 60—70% confluence, either 2 µg of plasmid DNA or 25 pmol of siRNA were diluted in Opti-MEM (31985062, Gibco) and then incubated with the transfection reagent for 20 min before applying dropwise onto the cells. All cells were incubated for 48 h prior to harvesting; half of the cells per replicate were used for protein analyses and the other half for RNA extractions.

2.4. CRISPR gene editing

Raji cells were first authenticated via STR profiling (C287, ABM) and tested negative for mycoplasma. Once confirmed they were transfected using the Amaxa nucleofector with IDT's Alt-R CRISPR/Cas9 system following manufacturer's recommendations. Briefly, crRNAs were designed using the IDT custom design tool to target exon 1 of the *HNRNPU* gene (CGGGCGACGAGAACGGGCAC). 2x10^6 cells per line were electroporated with the generated RNP complex in nucleofector solution V (program M-013). Cells were grown to confluence, then single cell expanded in 96-well plates in 100 µL MethoCult H4435 (StemCell). Genomic DNA was extracted using the DNeasy blood and tissue kit (Qiagen) according to the manufacturer's recommendations. *HNRNPU* was then PCR amplified from the genomic DNA with right primer GCCCTCACCATGAGTTCCT and left primer TCCGCCTTTCTGTTCTGTTT. The mutations in clones were verified by Sanger sequencing (GeneWiz) with primer: GAGGAAGGAATCTCCGCTCT. To obtain single-cell expanded WT clones, parental cells were single cell expanded in the same conditions as CRISPR cells and validated as WT by Sanger sequencing.

2.5. RNA-sequencing and analysis

RNA was extracted from $2x10^6$ cells per sample using the RNeasy mini kit (QIAGEN) according to manufacturer's recommendations with additional DNase I treatment. Qualities of total RNA samples and RNA integrity numbers were determined using an Agilent Bioanalyzer RNA Nanochip. Ribo-depleted RNA libraries were prepared and sequenced by the Genome Sciences Centre in Vancouver, BC. Sequencing was performed on the Illumina NovaSeq platform with 150 bp PE sequencing (100M individual reads). RNA-seq reads were aligned using STAR version 2.7.10a[167] to GRCH38. Feature counts was used to summarize gene counts[168]. The Bioconductor R package, DESeq2[169], was used to correct the read counts for library size and to obtain differentially expressed genes between conditions of interest employing a threshold of abs(log2FoldChange) > 0.585 and $p < 0.05$. DESeq2 results were then fed into the FGSEA[170] Bioconductor R package for pathway enrichment analyses, with a threshold of $p < 0.05$ set for enriched pathways. EnrichR[171] was also used to assess enriched pathways, particularly for reactome analysis. Differential splicing analysis was performed using rMATs 4.0.1.

2.6. Quantitative real-time PCR (qPCR)

Total RNA was extracted as described for RNA-seq. RNA concentration was determined by NanoDrop and 0.25 µg of RNA was used as input for cDNA synthesis using iScript cDNA synthesis kit (BioRAD). cDNA was diluted 1:10 and mixed with primers and SYBR Green Supermix (Quantabio). Quantification and analysis were performed with the QuantStudio 3 RT PCR machine (ThermoFisher). Relative expression of mRNA was determined after normalization to the geometric mean of human Actin and GAPDH levels using the ΔΔCt method. All plots represent at least three biological replicates of extracted RNA. Data was analyzed using GraphPad Prism 9 (GraphPad Software). qPCR primers used in this study are as follows:

Target	Forward	Reverse
HNRNPU	GGATAAGATGATGGTGGCAGG	TTCTTTCGGGCAGCAATCTC
MYC	TCTCTCCGTCCTCGGATTCT	TTCTTGTTCCTCCTCAGAGTCG
E2F1	CCGTGGACTCTTCGGAGAACT	GGCTGATCCCACCTACGGTC

E2F2	CATCCAGTGGGTAGGCAGGG	AAGGCCTGCTCCGTGTTCAT
E2F6	CGGCGAGGAAGTTACCCAGT	TAGAGCTTCAGCAGGCCCTC
E2F8	AGGCTCAAAGAGGGCAAGCA	TGGGAACAAGGTTGCGGAGA
ACTIN	CCACGGCTGCTTCCAGC	TAGTTTCGTGGATGCCACAG
GAPDH	GAAGGTGAAGGTCGGAGTCA	AATGAAGGGGTCATTGATGG

2.7. *XBP-1* alternative splicing detection with digital droplet PCR (ddPCR)

RNA extractions and cDNA synthesis performed as for qPCR. ddPCR was performed on the QX200 system (BioRad) using the primers and probes as outlined below. We employed a multiplex digital PCR approach to assess the levels of the two XBP-1 splicing isoforms within a single sample. Utilizing identical forward and reverse primers, both XBP-1 isoforms were amplified. Forward primer: GGGAATGAAGTGAGGCCAGT, reverse primer: CTCTGAATCTGAAGAGTCAATACCG. Full length XBP-1 was detected with the probe 5'FAM-AGCACTCAGACTACGTGCAC and spliced isoform was detected with the probe: 5'HEX-TCCGCAGCAGGTGCAGGC. Assay was validated by inducing ER stress with 2 μM Thapsigargin for 8 hours and visualizing splice isoforms on agarose gel. All ddPCR results are quantified as a ratio of the absolute copy numbers of splicing isoforms in each sample. All plots represent at three biological replicates of extracted RNA. Data was analyzed using GraphPad Prism 9 (GraphPad Software).

2.8. siRNA knockdown

siRNA knockdowns were performed with an siRNA specific for *HNRNPU* (Assay ID: 145413, ThermoFisher). Control transfections were performed with a negative control siRNA (4404021, ThermoFisher). Transient transfections were performed with the protocol mentioned above. Cells were incubated with the siRNA for 48 h prior to harvesting for RNA extractions and western blotting.

2.9. mRNA half-life analysis

$1x10^6$ cells were seeded for the *HNRNPU* mutant Raji lines or siRNA knockdowns for *HNRNPU* were performed in HEK293 cells as described above. 4 µM actinomycin D was added per well and incubated with cells at four time points, 0,30,60 or 90 minutes. Cells were harvested and RNA extractions and analysis were performed as described above. mRNA quantities were determined relative to the amount of mRNA at time point 0. Time points were plotted and fit to a linear regression using GraphPad.

2.10. Plasmid construction

2.10.1. *HNRNPU*-eGFP vector

The hnRNPU coding sequence was amplified from the pFRT/TO/HIS/FLAG/HA-HNRNPU vector (plasmid #38068, Addgene) with Q5 master mix (M0491L, NEB) using the following primers:

Forward primer: gatacaAAGCTTTGATGAGTTCCTCGCCTGTT
Reverse primer: gatacaGGTACCCTAATAATATCCTTGGTGATAATGCT

The PCR amplificon was run on an agarose gel and gel extracted (28704, Qiagen) according to the manufacturer's recommendations. The amplicon was digested with the restriction enzymes *Hind*III and *Kpn*I (NEB). The digested fragment was then cloned into the pEGFP-C1 vector, placing hnRNPU in frame and downstream of eGFP.

2.10.2. *MYC* minigene vectors

Two variants of the *MYC* minigene vector were cloned, one containing all introns and exons up until the stop codon (full-length minigene) and the other lacking intron 1 (Δintron-1). The full length *MYC* sequence containing the 5' UTR, all introns and exons up until the stop codon was PCR amplified from normal human donor DNA (G1471, Promega) in two PCR reactions with the following primers:

N-terminus

- Forward Primer: gatacaAAGCTTGGGATCGCGCTGAGTATAAA
- Reverse Primer: TGAAGGAGAAGGCGAGAGGC

C-terminus

- Forward Primer: TAGGGCGCGAGTGGGAAC
- Reverse Primer:

  gatacaTCTAGATTActtgtcatcgtcgtccttgtagtcCGCACAAGAGTTCCGTAGCTG

Both fragments were amplified with PCR. A C-terminal FLAG-tag was introduced into the sequence using the C-terminus reverse primers. The N-terminus fragment was digested with *Hind*III and *Apa*I and the C-terminus fragment was digested with *Apa*I and *Xba*I. The pRL-TK vector (Promega, E2241) was digested with *Hind*III and *Xba*I to remove the luciferase insert. A three-part ligation was then performed with the N-terminus, C-terminus *MYC* fragments and the digested vector. The complete sequence was validated with Sanger sequencing (GeneWiz) using tiled primers along the length of the gene.

For the Δintron-1 variant of the *MYC* minigene, the full-length minigene was used as input for the PCR reaction.

5’ UTR-Exon 1 was amplified with the following primers. The reverse primer contained an overhang complimentary for exon 2 that would be necessary for the next round of amplifications.

- Forward primer: gatacaAAGCTTGGGATCGCGCTGAGTATAAA
- Reverse primer:

  CGTCGCGGGAGGCTGCTGGTTTTCCACTACCCGAAAAAAATCCAGCGTCTAAGCAG

Exon 2-Stop codon sequence was amplified with the following primers. The forward primer contained an overhang complimentary to the end of exon 1:

- Forward primer: TTCGGGTAGTGGAAAACCAGCAGCCTCCCGCGACGATGCCCCTCAACGTTAGCTT
- Reverse Primer: gatacaTCTAGATTActtgtcatcgtcgtccttgtagtcCGCACAAGAGTTCCGTAGCTG

The amplicons from the above reactions were gel extracted, co-purified and used as substrate for the next reaction. An overlap PCR strategy allowed for by complementary overhangs introduced by the above reaction. This would generate the full MYC sequence lacking intron 1.

- Forward primer: gatacaAAGCTTGGGATCGCGCTGAGTATAAA
- Reverse primer: gatacaTCTAGATTActtgtcatcgtcgtccttgtagtcCGCACAAGAGTTCCGTAGCTG

This amplicon was gel extracted and digested with *Hind*III and *Xba*I (NEB) and cloned into the PRL-TK backbone. The entire sequence was subsequently validated with sanger sequencing (GeneWiz) using tiled primers along the length of the gene.

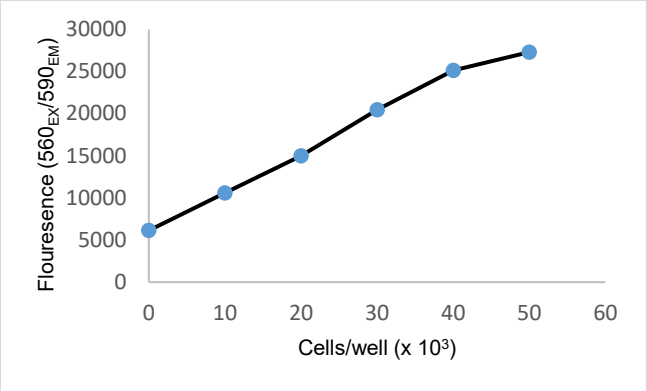
2.11. Fluorescence-activated cell sorting

To enrich for positively transfected cells with the *HNRNPU*-eGFP construct FACS was performed using the FACSAria Fusion Cell Sorter (BD) using a 100mm nozzle. Gates were applied to enrich for cells with medium-high GFP expression and were collected in 15 mL round-bottom polystyrene test tubes containing cell culture media. Immediately after sorting, the cells were spun down at 300xg for 5 minutes and lysed for western blot or RNA extractions.

2.12. Cell viability assay

Cell viability was measured with a fluorimetric test as previously described[172]. Raji cells were first seeded at varying densities to optimize the concentration within the readable, linear range of detection, as shown below. This range was determined to be 10,000-40,000 cells per well, and all future experiments were conducted within this range.

For cell viability assays, 100 uL of cells at a concentration of 200,000 cells/mL were seeded in a 96 well tissue culture plate and incubated for 24 hours. After the incubation period, 20 µL of cell titre blue (Promega, G8080) was added and incubated for 2 hours before determining fluorescence ($560_{Ex}/590_{Em}$) using a fluorometer (SpectraMax A5 PlateReader, Molecular Devices). Readings were taken at two time points, at seeding and after 24 h in 3 technical replicates per sample. Cell viability was normalized to that at seeding to account for differences in cell counts. The percentage of viable cells were calculated relative to the control cells. Each experiment was performed in at least three biological replicates.



2.13. Fluorescence imaging

Raji cells were first transfected with the eGFP-*HNRNU* construct and adhered in a monolayer to coverslips treated with concanavalian A. Coverslips were coated in 1 mg/ml concanavalin A for 2 h before washing twice with 1X phosphate buffered saline (PBS) (Gibco) and drying overnight. The coverslips were then incubated with 2 x 10^6 cells/mL cells and allowed to adhere for 2 h and immediately fixed with 1 mL of ice-cold methanol (-20°C) over a 5-min period, followed by a single rinse with room temperature 1X PBS. After discarding the PBS, the cells were briefly washed with a 0.2 ng/mL Hoechst 33342 solution (prepared in 1X PBS). This was followed by another wash with 1X PBS, and the coverslips were mounted on glass slides using Fluoromount G (ThermoFisher). The edges of the coverslips were sealed with nail polish. Fluorescence images were captured at 40X magnification using an EVOS FL Auto 2 microscope.

2.14. Inhibitor assays

50 µl of 400,000 cells/mL were seeded in a 96 well tissue culture plate. E2F inhibitor - HLM00647 (324461, Sigma) or MYC inhibitor - Mycro3 (HY-100669, MCE) were diluted in cell culture media to various concentrations. 50 µL of diluted inhibitor was added to the cells at a final concentration of 5 -100 µM. Cells were incubated in standard tissue culture conditions for 24 h. After 24 h, viability assay was performed as described above. All cell viabilities were normalized to a DMSO control. Drug concentrations were plotted after log-transformation and plotted as log[inhibitor] vs response – variable slope (four parameters) using GraphPad. One curve was plotted per biological replicate with two technical replicates each to determine the IC50.

2.15. Statistical analyses

All statistical analysis were performed with R version 1.4.1 or graphpad prism. For multiple comparisons a one-way ANOVA for multiple comparisons was used. Analyses involving two populations were analysed with a students T-test.

Chapter 3. Results

3.1. Splicing factors are recurrently mutated in mature B-cell lymphomas

To identify splicing factors recurrently mutated in B-cell NHLs, I investigated sequencing data from our ongoing meta-analysis of mature B-cell lymphoma (GAMBL). When examining the simple somatic mutations present in these lymphomas alongside a catalog of established splicing factors[173], it was observed that splicing factors are recurrently mutated in B-cell NHLs at varying frequencies as shown in **Figure 3-1**. This includes many hnRNPs, with *HNRNPU*, *HNRNPQ* (SYNCRIP), *HNRNPD*, and *HNRNPM* ranking among the top 10 most commonly mutated splicing factors.

Furthermore, a trend of mutual exclusivity in mutations amongst different splicing factors was observed as shown in **Figure 3-1** where we see that it is uncommon to find more than one splicing factor mutated within the same tumour. Given the crucial role and expression of distinct splicing factors at each stage of B-cell development, it is understandable that the prevalence of some splicing factor mutations may vary depending on the lymphoma subtype. For instance, *HNRNPD* mutations are most prevalent in DLBCL, while *PCBP1* mutations are predominantly found in BL. The exact mechanism or role they play in disease progression, however, remains largely unknown.

In the cohort evaluated, *HNRNPU* is the most recurrently mutated splicing factor across all lymphomas with mutations observed in BL, HGBCL-DH/TH, DLBCL AND FL. These mutations are interesting because nonsense mutations are common and they appear heterozygous, indicating that biallelic loss of this gene may be rare. These have not previously been reported in the context of lymphoma, so we sought to further evaluate the role of hnRNPU in B-cells and explore the functional relevance of these mutations.

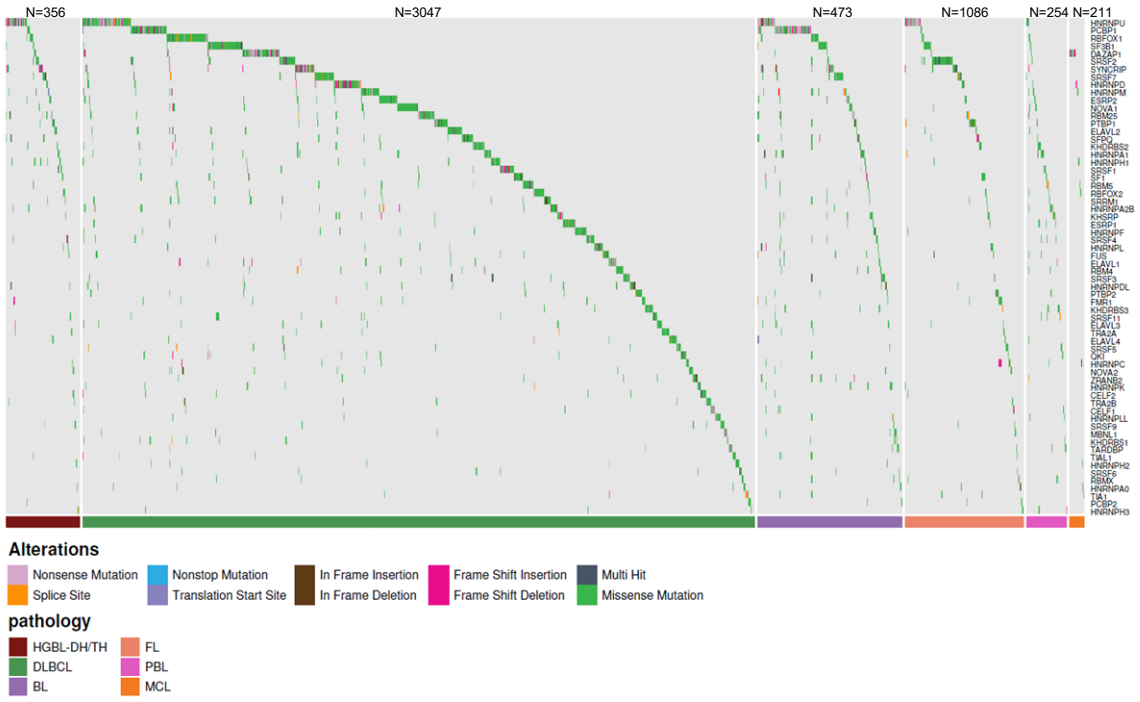


Figure 3-1. Recurrently mutated splicing factors in mature B-cell lymphomas

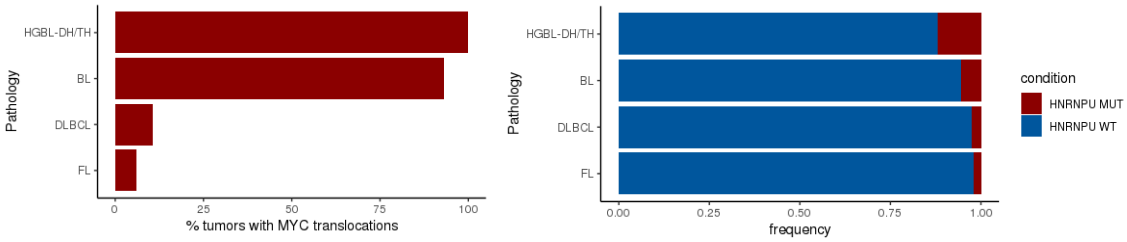
A total of 5427 tumour samples were evaluated for splicing factor mutations. Cases with at least one mutated splicing factor are shown in the oncoplot with total cases evaluated for each lymphoma subtype displayed above. We find that splicing factors are recurrently mutated in B-cell NHLs with *HNRNPU* being the most frequently mutated splicing factor.

HNRNPU mutations are most frequent in HGBL-DH/TH tumours (12%) followed by BL tumours (5.6%). Mutations are also found in DLBCL and FL but at lower frequencies, specifically 2.6% and 2%, respectively. HGBL-DH/TH and BL are similar in that both are derived from the dark zone of the germinal centre (centroblasts) and harbour *MYC* rearrangements. As shown in **Figure 3-2A**, *HNRNPU* mutations are enriched in tumors with *MYC* translocations perhaps linking a functional importance of hnRNPU in MYC-driven lymphomagenesis. We also found *HNRNPU* mutations to be enriched in Epstein-Barr virus (EBV) + tumors. EBV has a tropism for B-lymphocytes and is a primary risk factor in BL, specifically in the equatorial belt of Africa and other parts of the world where malaria is hyperendemic[174]. Latent EBV subsequently predisposes to malignant transformation, especially in the setting of impaired cell mediated immunity and chronic antigenic activation, where the virus replicates and expresses viral antigens that promote growth and survival of the cell[175]. *HNRNPU*

mutations preferentially occur in EBV+ BL (7.35% of cases) relative to EBV- (2.28% of cases) and this difference is statistically significant (**Figure 3-2B**). We also see *HNRNPU* mutations in EBV+ HGBL-DH/TH although with only 3 cases in this cohort being EBV+ it remains inconclusive.

Based on the findings presented above, we aimed to investigate the expression levels of hnRNPU in relation to MYC expression, translocation status, and EBV status in cell lines. For this study, we selected two cell lines from HGBL-DH/TH and BL, each being either EBV+ or EBV-. We then compared the hnRNPU and MYC protein expression patters in these cell lines to a non-lymphoma cell line, HEK293, and cell lines from MCL where *HNRNPU* mutations are rarely observed (**Figure 3-2C**). The unifying feature amongst all cell lines is that those with high MYC expression also express high expression of hnRNPU. We did not observe any pattern with regards to EBV status from the cell lines tested.

A



B

Pathology	EBV status	% HNRNPU mutant	P-value
BL	Positive	7.35	0.054
	Negative	2.28	
HGBL-DH/TH	Positive	33.33	0.365
	Negative	9.56	

C

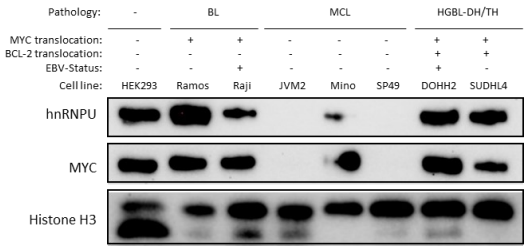


Figure 3-2. *HNRNPU* expression and mutations are enriched in lymphomas with high MYC expression

A) Lymphoma subtypes were evaluated for *MYC* rearrangement status with GRIDSS and Manta. Enrichment of *HNRNPU* mutations are seen in 12.07% cases mutated in HGBCL-DH/TH, 5.7% in BL, 2.6% in DLBCL and 2% in FL. B) *HNRNPU* mutations are also enriched in EBV+ tumors. P value calculated with fishers' exact test. C) lymphoma cell lines with high hnRNPU expression also have high expression of MYC protein. Western blot shown as representative image from three biological replicates.

3.2. CRISPR-induced *HNRNPU* mutations in BL cell line led to decreased protein levels

Due the recurrence of *HNRNPU* mutations in MYC-driven lymphomas, our objective was to establish a cell line model suitable to evaluate the function of these mutations in lymphoma cells. Given the pleiotropic role of hnRNPU, we sought to comprehensively determine the effects of these mutations on the transcriptome. We first investigated the distribution of mutations on the *HNRNPU* locus and found that there was a predominance of nonsense mutations clustered at the 5' end of the gene. Interestingly, most nonsense mutations were the result of a single nucleotide substitution converting a Glutamine residue to a stop codon as shown in **Figure 3-3A**. These mutations clustered at the 5' end and were C -> A mutations, which can be obtained through the enzymatic activity of AID. We were curious whether these mutation hotspots arose due to aberrant AID activity (i.e. aSHM). Indeed, most nonsense mutations existed in WRCY motifs (W = A/T, R = A/G, Y = C/T) which are known AID target hotspots[176]. This alludes to aSHM as a putative mechanism for acquisition of these mutations. Most mutations occur before the SPRY domain, which would effectively remove it and the C-terminus RNA binding domain but may leave N-terminus DNA binding domain intact. We sought to evaluate whether these mutations indeed caused complete loss of the protein or produced a truncated isoform that retained only certain activities relative to the full-length protein.

We therefore used a CRISPR-Cas9 approach whereby we designed a guide RNA targeting the region proximal to the mutation hotspot (aa152-158). We chose the BL cell line Raji as our model since it best reflected the patient cells harbouring these mutations. Specifically, Raji cells have a MYC-IgH translocation, are EBV+ and express high levels of hnRNPU and MYC as shown in our cell screen (**Figure 3-2C**). Following validation of cell lines positive for *HNRNPU* frameshifts, we chose two cell lines, CRISPR 36 and 38 for our downstream analysis. Interestingly, both cell lines are compound heterozygotes with a frame shift deletion in one allele and an in-frame deletion in the other (**Figure 3-3B**). Since amino acids 60-250 of hnRNPU are in an intrinsically disordered region, we hypothesise that the in-frame deletion removing 2-4 amino acids would not negatively impact the protein and thus consider this allele to produce a variant of hnRNPU that is functional. The deletions in CRISPR36 and 38 result in frame shifts at aa150 and aa155 respectively. These eventually result in a stop

codon at aa203 for CRISPR36 and aa199 for CRISPR38. We did not identify any compound heterozygotes representing two frameshift mutations. As seen in **Figure 3-3C**, we observed an additional band at the 100kDa mark for CRISPR38. Raji cells are diploid for *HNRNPU* and the DNA sequencing from this cell revealed only the 41bp and 6bp deletions consistent with this cell line being a compound heterozygote. We were curious about the mechanism of achieving this isoform given that it must contain the first 100 amino acids of hnRNPU as the antibody used is an N-terminus antibody that binds this region. This isoform is also too large to be a truncated variant. We then looked at the RNA-sequencing alignment for this cell line and we observed a 348bp skipping event that would skip the site of deletions in both alleles and result in an in-frame transcript. This was found in nearly 25% of all the RNA-sequencing reads but was absent from the DNA sequencing reads from *HNRNPU*, suggesting that this is an RNA level event. There are no canonical splice sites at the edges of this deletion event, so it is not likely a splicing event. There are, however, strings of Gs at either side of the deletion, perhaps contributing to G-quadruplex structure which are known to contribute to instability based deletions[177]. We find that with greater passages the cells evolve to favour this shorter isoform relative to the WT isoform.

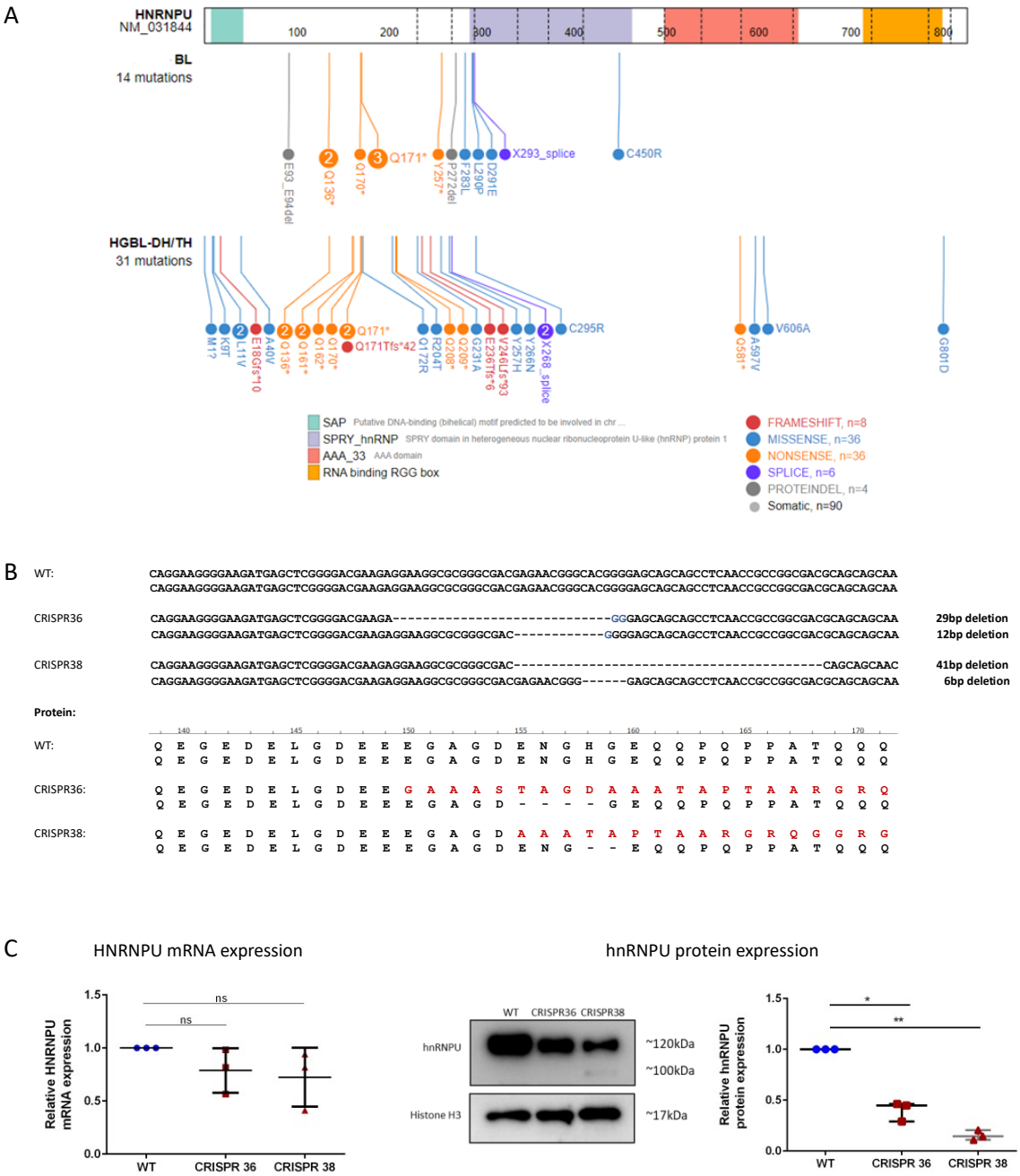


Figure 3-3. CRISPR induced HNRNPU mutations in Raji cell line
A) Recurrent *HNRNPU* nonsense mutations form hotspot at the N-terminus of the gene in both BL and HGBL-DH/TH. Mutations plotted with proteinpaint[178] B) CRISPR induced deletions in the *HNRNPU* locus of Raji cells results in the formation of compound heterozygotes with a frame shift deletion in one allele and a small in-frame deletion in the other. C) *HNRNPU* expression measured at mRNA level with qPCR and western blot for protein level. We see that while there is variable expression of *HNRNPU* at the mRNA level, there is a consistently low level of hnRNPU protein in these cell lines. qPCR performed in three biological replicates; western blots shown as representative image from three biological replicates. Significance was calculated by ANOVA ($p^* \leq 0.05$; $^{**}p \leq 0.01$)

We then analyzed the protein and mRNA levels of *HNRNPU* in the mutant cell lines relative to a WT cell line that had been single cell split, expanded in the same way and at the same time as the mutant lines and were validated to have wild type *HNRNPU*. We found that while there was variable expression of *HNRNPU* at the mRNA level, the protein expression was consistently lower in the mutants (**Figure 3-3C**). We also did not observe the presence of any truncated isoform even in the context of inhibited proteasomal activity, indicating that the transcripts bearing frameshift mutations are likely degraded by NMD and are not translated.

3.3. *HNRNPU* mutations alter the transcriptomic and splicing landscape in MYC driven B-cell lymphomas

The observation of decreased hnRNPU protein levels in the mutant cell lines support the notion that a single copy of *HNRNPU* that is close to the WT sequence, is insufficient to maintain normal levels of protein. We hypothesised that this decrease would have a downstream effect on the transcriptome, altering the gene expression and splicing landscape to promote lymphoma progression. To understand this further we generated RNA-seq libraries from the two *HNRNPU* mutant cell lines and compared the gene expression and splicing landscape to the wild-type cell line.

3.3.1. Impact of *HNRNPU* mutations on the gene expression landscape

We first performed transcriptomic analyses to gain insights into the molecular mechanisms underlying *HNRNPU* mutations. Through differential expression analysis we found that there were 909 significantly differentially expressed transcripts when comparing *HNRNPU* mutant to WT samples (P-value <0.05, Log2FC > 0.58) (**Figure 3-4A**). Within these, there were more transcripts significantly upregulated (563) than downregulated (346) in the mutants. As seen through hierarchical clustering based on gene expression, we see that both mutant cell lines cluster together and have similar gene expression patterns that are distinct from the wild type cells and the gene expression in both mutant cell lines (CIRPR36 and CRISPR38) is more consistent (**Figure 3-4B**).

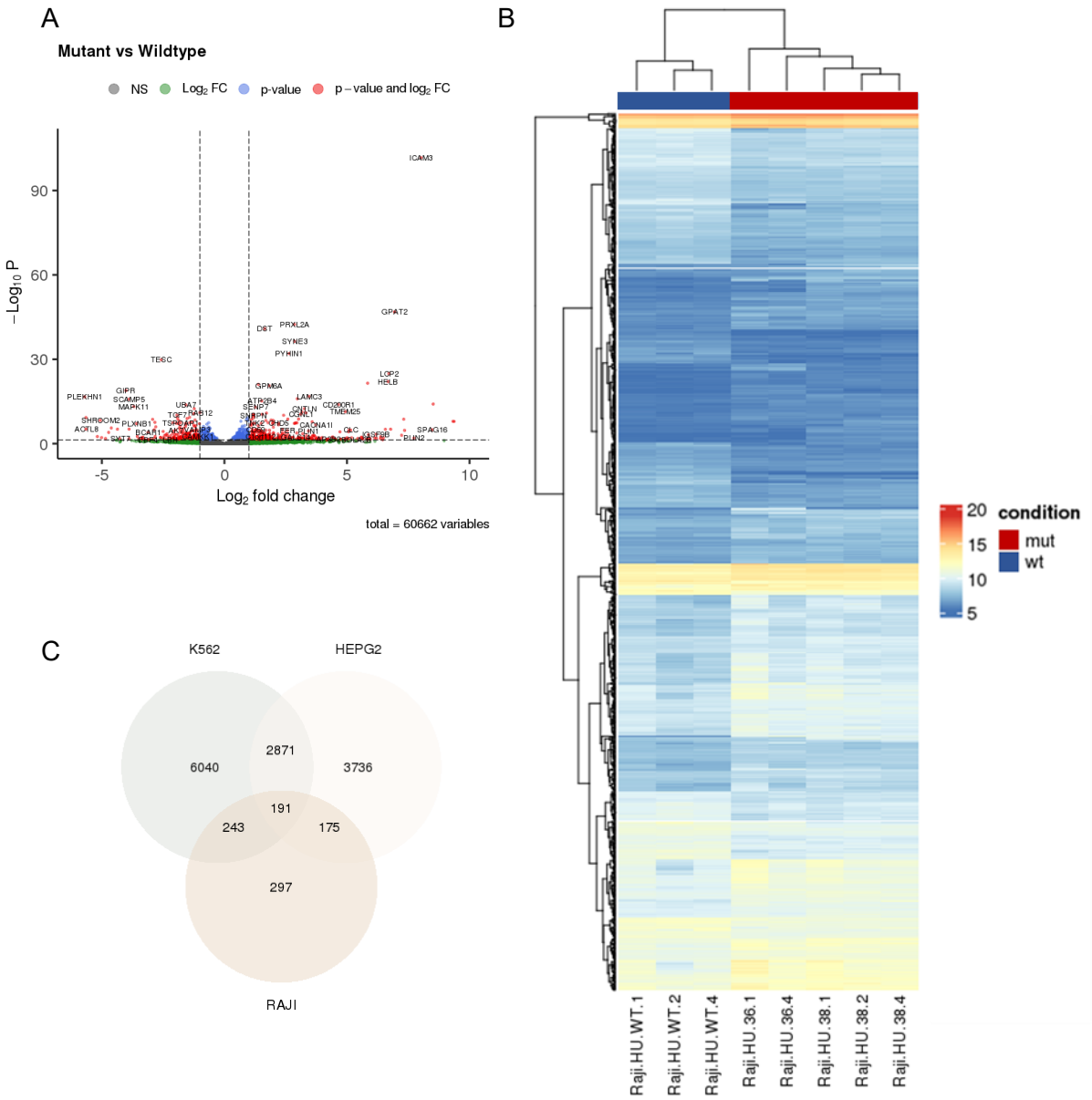


Figure 3-4. *HNRNPU* mutations alter the gene expression landcape of Raji cells
A) Volcano plot demonstrating 563 genes upregulated and 346 genes downregulated. B) Heatmap showing hierarchical clustering of the expression of 909 DE genes in WT (*HNRNPU*+/+) and Mutant (*HNRNPU*+/-) Raji lines. C) Comparisons of differentially expressed transcripts in *HNRNPU* mutant cell lines compared to hnRNPU shRNA knockdown in HEPG2 and K562 cells.

In addition, we reanalyzed *HNRNPU* knockdown followed by sequencing data in two cell lines, HEPG2 and K562 obtained from the encode project. Our objective was to identify the degree of overlap among genes that showed differential expression in these cell lines relative to those found in the *HNRNPU* mutant Raji cells. HepG2, a hepatocellular carcinoma cell line, stands out as quite distinct from B cells in terms of cell lineage. In contrast, K562 originates from lymphoblasts and shares a closer lineage with B cells. Few transcripts (191) were consistently differentially expressed between the mutant Raji cells, *HNRNPU* knockdown HEPG2 cells and *HNRNPU* knockdown K562 cells (**Figure 3-4C**). The mutant Raji cells contain most gene expression similarities with the *HNRNPU* knockdown K562 cells, consistent with the notion that K562 cells are closer in lineage to B cells it was. This emphasized the notion that hnRNPs exhibit distinct functions and cellular targets that vary according to the cell type.

Next, we wanted to understand what cellular functions are broadly impacted by *HNRNPU* mutations. Using differentially expressed genes from the *HNRNPU* mutant CRISPR lines, we performed gene set enrichment analysis. Using the hallmark pathways data set, we found that the most significantly upregulated pathways were related broadly to promotion of the cell cycle (**Figure3-5A**). These include enrichment in targets of the transcription factors MYC and E2F (**Figure 3-5B**). These results were interesting since MYC is one of the primary oncogenic drivers of BL and HGBL-DH/TH and are thought to carry out cell cycle dysregulation in cooperation with E2Fs[92]. Genes related to the G2M checkpoint, mitotic spindle and DNA repair were positively enriched, indicative of increased cell cycle entry and proliferation. Downregulated gene sets included those belonging to various metabolic pathways such as angiogenesis, cholesterol homeostasis and fatty acid metabolism.

Given that many of the significantly upregulated pathways were linked to the initiation and advancement of the cell cycle, along with the elevated expression of both MYC and E2F target genes (**Figure 3-5C**), we postulated that mutations in *HNRNPU* may be influencing MYC and E2F family proteins to carry out these roles. To investigate this further, we examined the gene expression of *MYC* and *E2F* family members in the differentially expressed transcripts from the RNA seq data. Notably, both *MYC* and *E2F* family members were significantly differentially expressed. The expression pattern, however, was intriguing, as *MYC* expression was significantly reduced in these cells. Within the E2F family members, *E2F1*, *E2F2*, and *E2F8* showed significant upregulation,

and *E2F6* exhibited a substantial downregulation (**Figure 3-5D**). The E2F family members, exist as canonical activators (E2F1,E2F2), canonical repressors (E2F6) or atypical repressors (E2F8)[179]. Since activating E2Fs (*E2F1* and *E2F2*) are upregulated and repressive *E2F6* is downregulated in the RNA-seq data, it is understandable that E2F targets are upregulated, and cell cycle entry is subsequently promoted. What is perplexing however, is that MYC targets are seemingly upregulated even when *MYC* itself is downregulated (**Figure 3-5C**). We propose that the higher representation of MYC targets in this dataset could be attributed to the shared pathway for cell cycle progression between E2F and MYC, along with the extensive crosstalk between these two transcription factors. Despite the reduced expression of MYC itself, the upregulation of E2F family members may consequently enhance the expression of MYC target genes.

3.3.2. Impact of *HNRNPU* mutations on the alternative splicing landscape

Given the roles of *HNRNPU* in RNA splicing, we analyzed the RNA-seq data to identify differences in isoforms detected in the *HNRNPU* mutant cell lines compared with the WT cell line using rMATS. We found a total of 908 alternative splice variants (FDR = 0.01, dPSI > 0.1) where the majority, 698 (77%), were skipped exon events, alternative 3’ and 5’ splice sites were observed in 41(4.5%) and 27(2.9%) events respectively, and intron retention was observed in 78(8.6%) events (**Figure 3-6A**). Due to the uncertainty in resolving mutually exclusive exons with short read RNA-seq data, this category was not further assessed. It's worth noting that only 45 of the alternative splice events lead to significantly altered gene expression. This implies that the altered splicing events result in isoforms that aren't merely subject to NMD and underscore hnRNPU's involvement in diverse facets of mRNA metabolism.

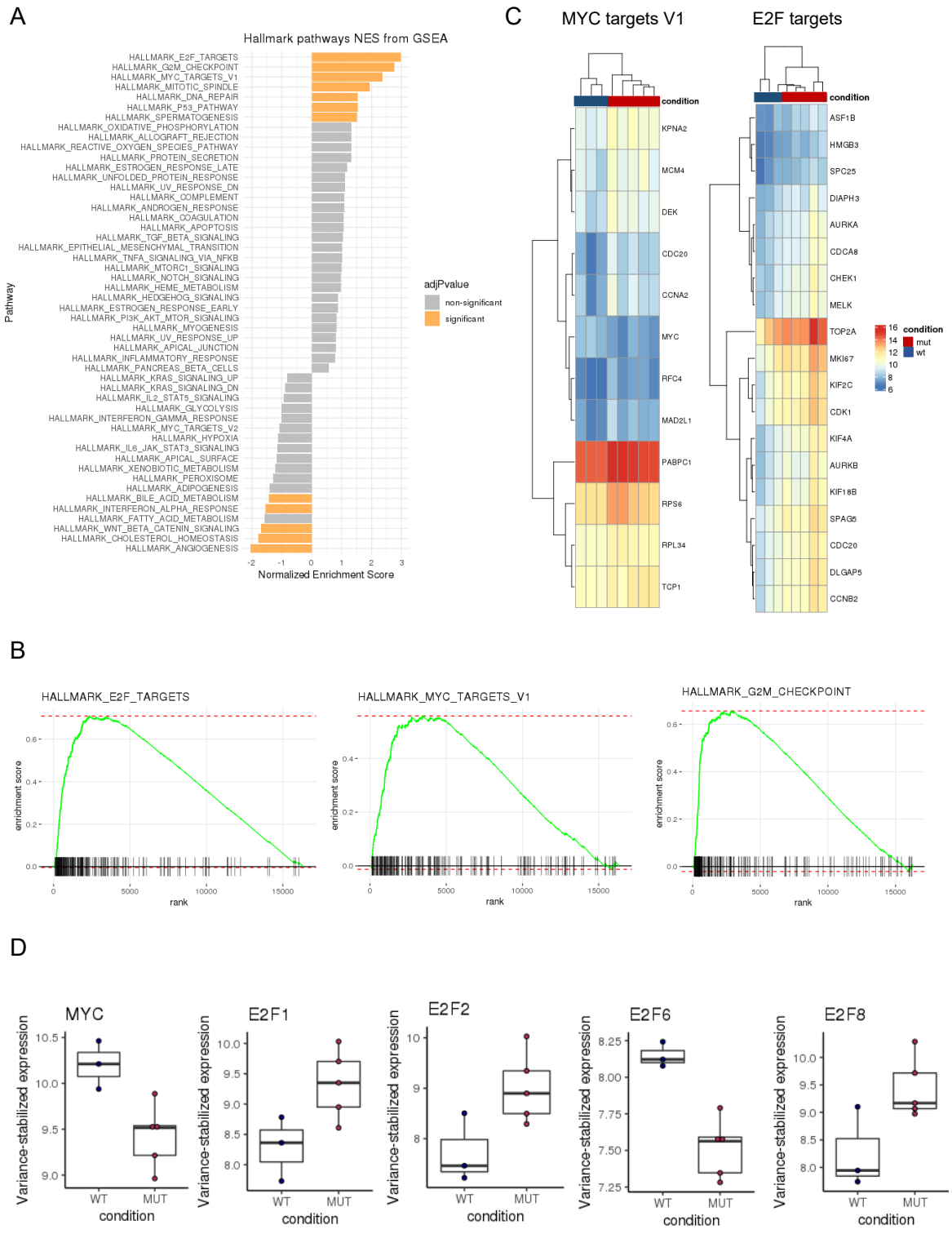


Figure 3-5. Significantly differentially expressed MYC and E2F family members in *HNRNPU* mutant CRISPR cell lines

A) Gene set enrichment analysis was performed using FGSEA with the hallmark pathways dataset. A positive Normalized Enrichment Score value indicates enrichment in the *HNRNPU* mutant phenotype, a negative NES indicates enrichment in the WT

phenotype. Significance threshold: adjPvalue>0.05. B) Heatmap showing expression of top 20 significantly differentially expressed MYC and E2F targets. C) Enrichment plots for top enriched pathways, showing the profile of the running Enrichment Score and positions of gene set members on the rank-ordered list. D) MYC and E2F family members that are found to be significantly differentially expressed with DESEQ, P-value <0.05, log2fc >0.58.

Next, we were interested in what cellular pathways were overrepresented in the alternatively spliced transcripts. We queried this list with the Reactome2022 dataset using EnrichR[171]. The most enriched cellular pathways amongst the alternatively spliced mRNAs were those involved in cell proliferation such as the G2/M transition, mitotic prometaphase, and organelle biogenesis (**Figure 3-6B**). Notably, a skipped exon event was observed in CDC25A, which is a protein phosphatase, and one of the most crucial cell cycle promoters, removing the inhibitory phosphorylation in various cyclin-dependent kinases important for G1/S transition[180]. We observed increased skipping of exon 6 in the *CDC25A* transcript of mutant cell lines (**Figure 3-6C**). *CDC25A* lacking exon 6 contributes to protein stabilization and activation based on reduced ubiquitination and phosphorylation at specific sites in exon 6[181]. Taken together, these findings suggest that *HNRNPU* mutations not only increase E2F expression, but also allows the cells to overcome inhibitory effects of Rb through the actions of CDC25A contributing to increase in proliferation associated genes observed in these cells.

The Rho GTPase cycle and signalling by Rho GTPases was also overrepresented in the altered splicing events. Rho GTPases are a family of small G-proteins of the Ras superfamily and are major regulators of actin polymerization[182]. In general terms, RHO GTPases are mostly involved in human tumourigenesis as downstream effectors of driver oncogenes and in hematological malignancies, they are implicated in tumour dissemination and invasion[183,184]. Rho GTPases have also been shown to regulate cell cycle entry and progression, in particular by regulating a number of genes involved in G1/S transition, e.g. cyclin D1[185]. Several Rho proteins have altered splicing in our mutant cell lines such as enhanced exon 3 skipping in *ARHGEF26,* a protein that activates RhoG and alters cytoskeletal reorganization[186]. Other Rho proteins with altered splicing include various transcripts encoding Rho GTPase activating proteins such as *ARHGAP44* and *ARHGAP25*, and *ARHGAP10.* These proteins bind to Rho GTPases to stimulate their activity. Enhanced exon skipping events in all three have

been observed in the *HNRNPU* mutant lines. In addition to the potential involvement of these proteins in cell cycle initiation, their influence on actin polymerization may be pertinent to *HNRNPU* mutations. Research has indicated that hnRNPU, when in complex with actin, exerts a regulatory function in the early stages of transcription activation by interacting with the C-terminal domain of RNA-polII[187]. Nevertheless, despite the documented instances of exon skipping in transcripts encoding Rho proteins and their annotation as such, the functional implications of these events remain largely unexplored. Although these exon-skipping events result in protein coding variants, the consequence of shifting from one isoform to the other is not well understood.

In addition, regulators of apoptosis also show differential splicing. CARD8, a member of the Caspase-associated recruitment domains (CARD) protein family was found to be differentially spliced in *HNRNPU* mutant cells. CARD8 triggers apoptosis in response to cell death signals by physically interacting with Caspase1[188]. Overexpression of CARD8 has been shown to induce apoptosis in transfected cells[189]. In the *HNRNPU* mutant cell lines, we observe a skipping event in exon 4 resulting in a transcript that is out of frame and degraded by NMD (**Figure 3-6D**). Exon skipping events in other transcripts encoding proteins involved in apoptosis such as MAPK8, CASP8, DNM1L were also observed although the functional impact of these altered splicing events are not understood.

Taken together, the gene expression and differential alternative splicing results collectively point to *HNRNPU* mutations altering the cell cycle dynamics of the B cells that harbour them to promote cell survival and proliferation.

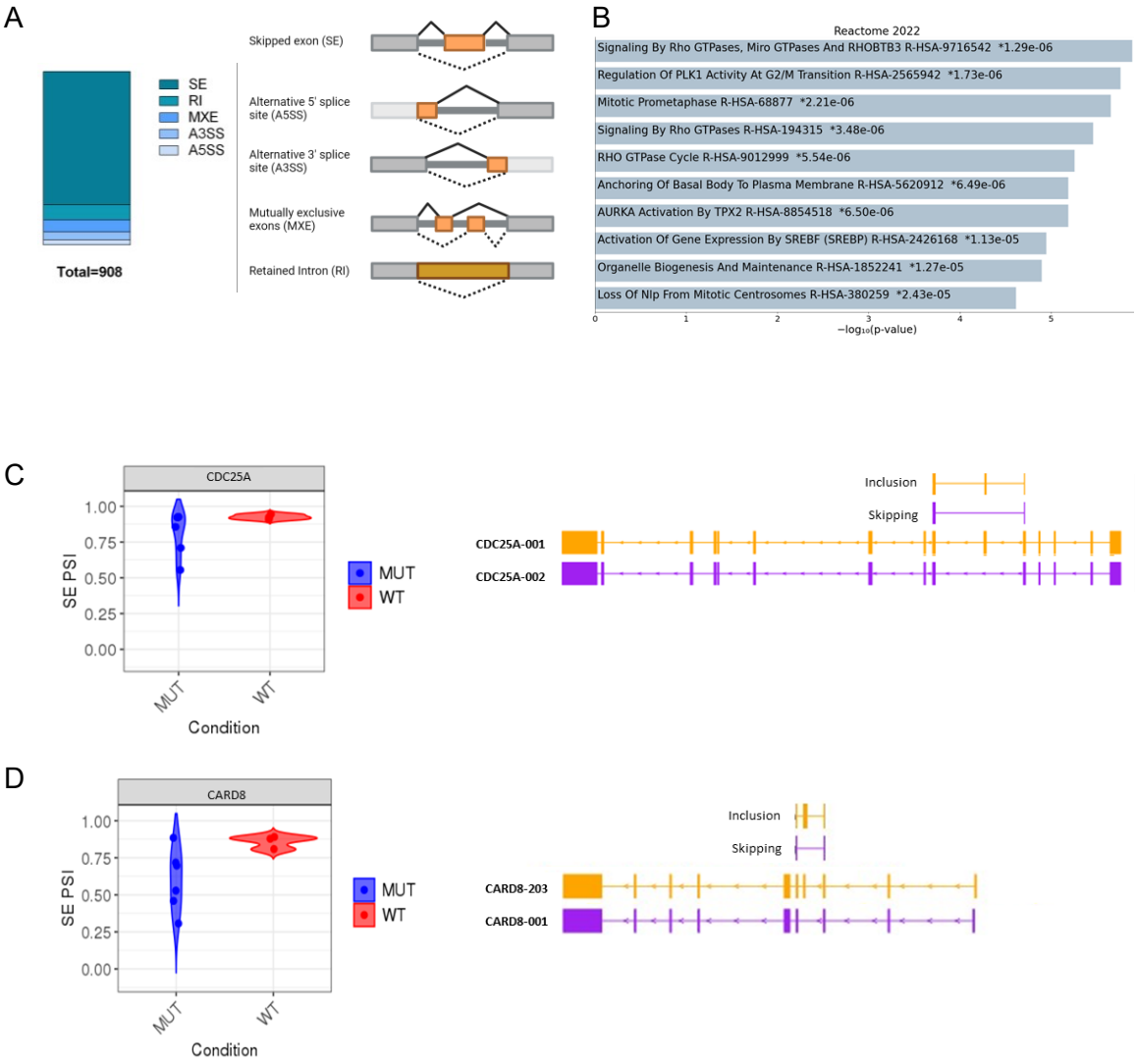


Figure 3-6. *HNRNPU* mutations alter the splicing landscape in Raji cells
A) A total of 908 alternative splice variants were identified from the RNA-seq data when we compare *HNRNPU* WT cell lines to mutants. FDR = 0.1 dPSI > 0.1. Amongst these, the most overrepresented class is that of skipped exon events. B) Pathway enrichment analysis performed with EnrichR using the Reactome2022 database. Genes involved in cell cycle progression are differentially spliced. C) Differential splice events in *CARD8* and *CDC25A*. dPSI calculated based on inclusion reads/exclusion reads.

3.4. *HNRNPU* mutations uncouple *MYC* and *E2F* activities allowing for cell cycle progression

Lymphomas with MYC dysregulation such as BL and HGBL-DH/TH exhibit oncogene addiction to MYC, meaning they are dependent on its expression for survival[190]. In Raji cells, it has been observed that *MYC* knockdown results in a significant decrease in cell viability[190,191]. For this reason, we were curious that the *HNRNPU* mutant Raji cell lines exhibited significantly decreased expression of *MYC* in the RNA-seq data (**Figure 3-5B**). We validated the decrease of *MYC* with both qPCR and western blot analysis. Notably, even though *MYC* was reduced to approximately 50% of WT at the mRNA level, with further reductions at the protein level, the cells maintained similar viability to WT (**Figure 3-7C**). We believe this might be a direct consequence of the upregulation of E2Fs and their target genes to promote cell proliferation. Upon evaluating the E2F pathway for major regulators exhibiting differential expression in mutant cells, it becomes apparent that there are limited alterations upstream of E2Fs in this pathway, except for the reduction in MYC. However, downstream targets of E2Fs, such as CDK1, CDC6, and ORC1, show upregulation, signifying that E2F upregulation plays a pivotal role as the mediator for cell cycle promotion (**Figure 3-7A**). We then validated the differentially expressed E2Fs from the RNA-seq results with qPCR and western blot (**Figure 3-7D-E**). With qPCR we observed the same trend where the canonical activators (E2F1 and E2F2) are upregulated, the canonical repressor E2F6 is downregulated and atypical repressor, E2F8 is upregulated. At the protein level we see E2F1 having the highest upregulation amongst E2Fs tested in both mutant cell lines. This is consistent with the mRNA expression, showing almost two-fold increase in E2F1 mRNA. E2F2 upregulation at the protein level is most apparent in CRISPR38 and E2F6 expression is significantly lower in both *HNRNPU* mutant lines. Changes in E2F8 protein are less conclusive. E2F1 is the primary E2F linked to MYC driven lymphomagenesis. E2F1 in particular is required for the initial entry to the cell cycle from a quiescent state and is required for the activation of other E2F genes[192]. It has been shown in cell lines that MYC is required to promote E2F1 mediated cell cycle entry. We hypothesise that *HNRNPU* mutations aid in uncoupling of the MYC-E2F axis, allowing for elevated E2F expression and cell cycle entry despite lower expression of MYC itself.

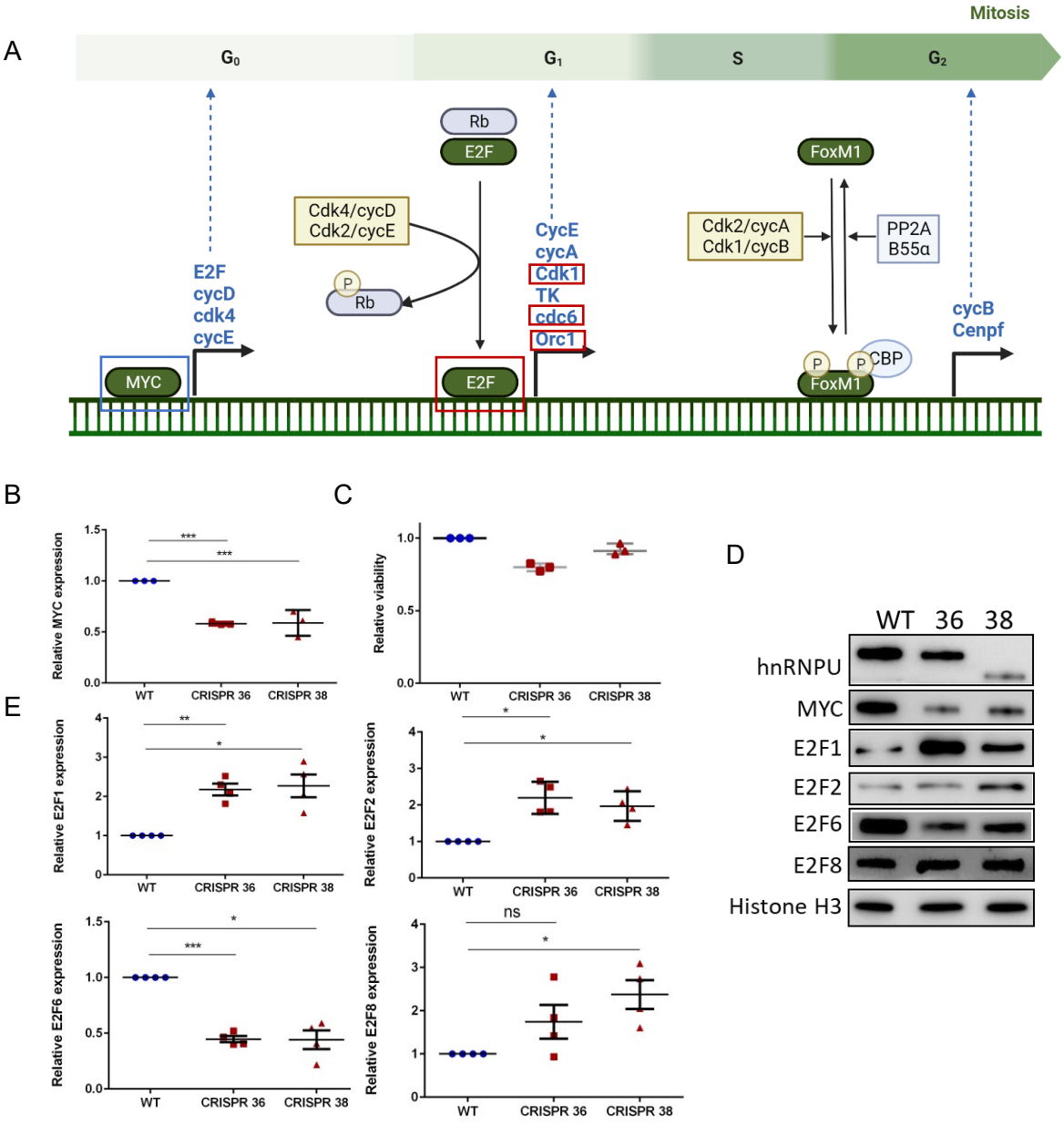


Figure 3-7. *HNRNPU* mutations uncouple MYC and E2F expression
(A) MYC and E2F pathway for cell cycle entry from G0 to G1/S phase. Blue boxes depict those that are significantly downregulated in the RNA-seq from the mutant lines relative to the WT, and red boxes indicate those that are significantly upregulated. (B) MYC expression measured with qPCR normalized to the geometric mean of Actin and GAPDH. Significant reduction of MYC is observed in the mutant lines relative to WT. (C) *HNRNPU* loss minimally alters cell viability. (D)Western blot showing reduction in hnRNPU protein in mutant lines, downregulation of MYC protein and induction of E2F1 protein. Histone H3 used as a loading control. (E) Altered mRNA expression of E2F family members in mutant lines, validated with qPCR. Significance was calculated by Anova ($p^* \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$). All qPCRs performed in three or four biological replicates; western blots shown as representative image from two biological replicates.

3.5. hnRNPU mediated modulation of E2Fs is observed in *HNRNPU* mutated HGBL-DH/TH tumours

We proceeded to investigate whether there was any observable evidence of E2F dysregulation in patient samples with *HNRNPU* mutant tumours. To do this, we utilized RNA-seq data from patients diagnosed with HGBL-DH/TH carrying *HNRNPU* mutations. Given that HGBL-DH/TH cases often involve translocations of MYC to non-Ig partners, we specifically selected a subset of samples with MYC-Ig rearrangements. The juxtaposition of *MYC* to the enhancers of non-Ig genes may result in distinct biological implications and altered *MYC* expression compared to that of *MYC*-Ig rearranged tumours, which is why we opted to exclude them. Subsequently, we compared the *HNRNPU* mutant samples to HGBL-DH/TH cases that did not have any mutations in splicing factors, ensuring no other splicing factors contributed to the observed changes. This comparison involved a total of 8 *HNRNPU* mutant samples and 28 splicing factor WT samples.

Through this analysis, we identified a similar pattern in patients as we did in the *HNRNPU* mutant cell lines. The dataset showed an enrichment of E2F target genes (**Figure 3-8A**), along with those associated with the G2M checkpoint. When we examined the expression of *MYC* and *E2F* family members that exhibited significant differences in the *HNRNPU* mutant Raji cells, we found that MYC expression was significantly reduced, consistent with what we observed in the mutant cell lines. While E2F1 expression showed an increase in the *HNRNPU* mutant HGBL-DH/TH tumours, it did not reach statistical significance, whereas E2F2 and E2F8 expression were significantly upregulated (**Figure 3-8B**). There were no noteworthy changes in E2F6 expression.

Given the limited sample size due to availability of RNA-seq data and filtering out non-Ig translocated cases as well as splicing factor mutated cases, it was encouraging to see a similar trend of E2F targets upregulated in the patient cell lines consistent with the findings in the *HNRNPU* mutant Raji cells.

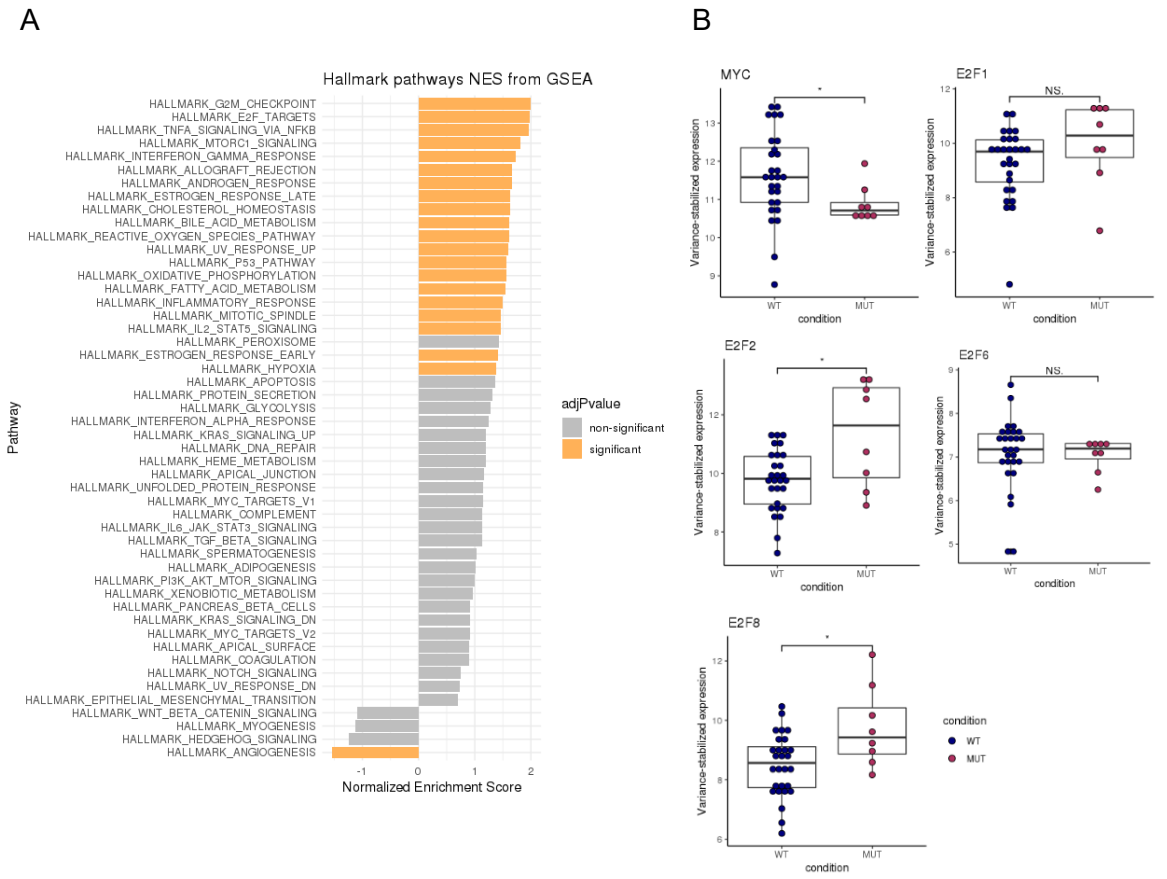


Figure 3-8. E2F targets and E2F expression is altered in *HNRNPU* mutated HGBL-DH/TH tumours

(A) RNA sequencing data from 8 *HNRNPU* mutants was compared to 28 splicing factor WT HGBL-DH/TH cases to obtain differentially expressed genes. Gene set enrichment analysis was performed with DESEQ2 and E2F targets and genes involved in the G2M checkpoint are the most significantly enriched. (B) Variance stabilized expression of E2F and MYC family members from the RNA-seq data. MYC and E2Fs differentially expressed in *HNRNPU* mutant lines were queried for differences in expression. Overall, MYC expression was significantly lower, consistent with findings in cell lines and E2F1 and E2F8 expression was significantly higher. Significance is determined with DEseq.

3.6. *HNRNPU* mediated modulation of E2Fs but not MYC is cell line specific

Subsequently, we aimed to determine whether the alterations in MYC and E2F expression were a direct outcome of *HNRNPU* depletion or if the cells had slowly adapted and reconfigured cellular dynamics to achieve this. To test this, we knocked down *HNRNPU* in Raji cells using an siRNA approach and validated the knockdown with both qPCR and western blot (**Figure 3-9A, C**). We then investigated the mRNA and protein expression of *E2F* and *MYC* 48 hours post transfection. Consistent with the cell line experiments, we observed that when *HNRNPU* was knocked down, there was a significant decrease in the mRNA expression of *MYC* (**Figure 3-9B, C**), accompanied by a reduction in its protein expression. Similarly, we observed consistent changes in *E2F1*, with an increase in its mRNA expression and the induction of the corresponding protein. As for *E2F2*, *E2F6*, and *E2F8*, although their mRNA expression patterns were consistent with what we observed in the *HNRNPU* mutant Raji cells (**Figure 3-9D**), we did not observe differences at the protein level (**Figure 3-9C**). This observation may suggest that E2F1 plays a pivotal role in the dysregulation of E2Fs.

We then performed the same experiment in HEK293 cells. From our initial cell screen (**Figure 3-2C**) we found that HEK293 cells, despite lacking *MYC* translocations, have high expression of *MYC* and *HNRNPU*. We knocked down *HNRNPU* in these cells and validated the knockdown with qPCR and western blot 48 hours post transfection. Here we see very robust decrease in *MYC* expression at almost 10% that of WT levels and very depleted *MYC* protein levels. When we examine expression of E2F family members we see a modest but significant decrease in all E2F family members however with no apparent changes at the protein level. We hypothesise that since MYC is responsible for the transcription of *E2F* family members under normal conditions, perhaps depleting it in HEK293 cells directly causes its expression to be lower.

The distinction between the two cell lines is intriguing and may be attributed to the fact that in certain cells, such as in MYC-driven B-cell lymphomas, *HNRNPU* loss instigates gene expression changes capable of uncoupling *MYC* expression from E2Fs, enabling the upregulation of the latter even in the absence of high *MYC* expression. *HNRNPU*-mediated control of *MYC* appears to be consistent across all cell lines tested. We had previously conducted a differential expression analysis using publicly available data from

the ENCODE project[165] in two cell lines, HEPG2 and K562 that had *HNRNPU* knocked down. When we inspect the common genes shared between *HNRNPU*-depleted HEPG2, K562, and Raji cells, we find MYC among them, but E2Fs do not appear in this intersection. Taken together we conclude that that hnRNPU's influence on E2F expression is specific to individual cell lines.

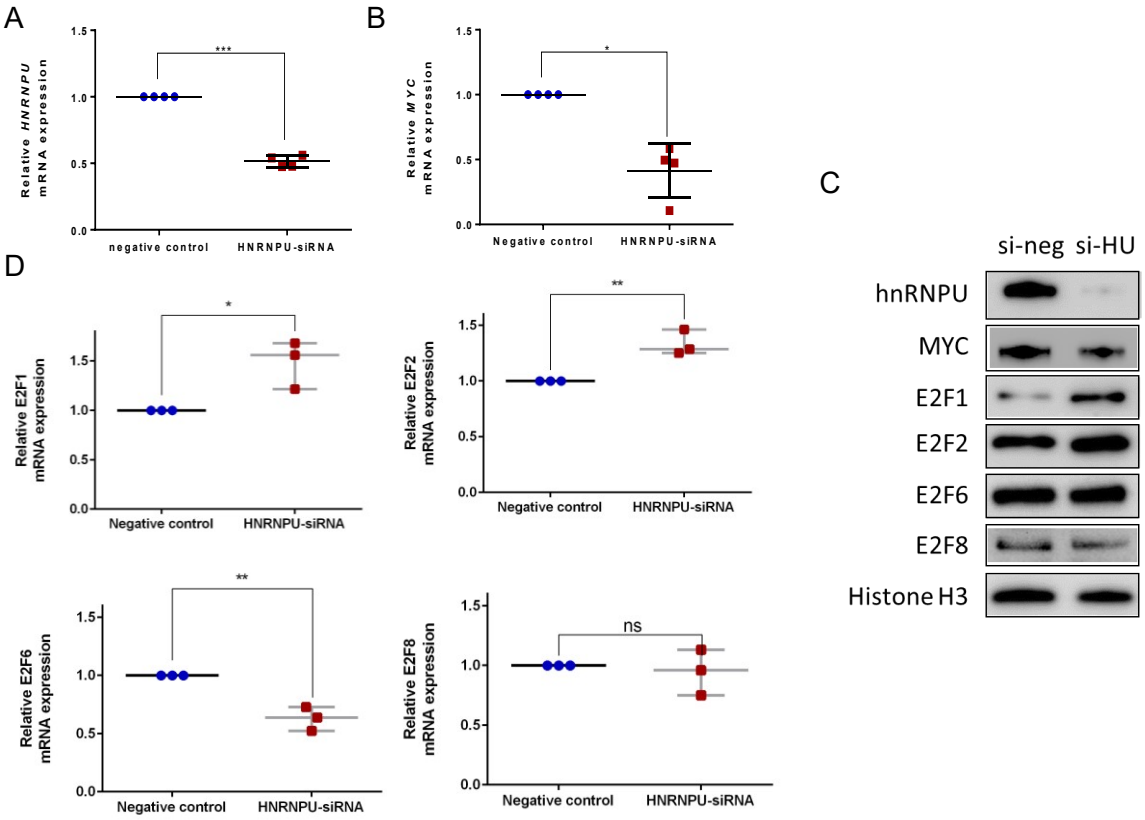


Figure 3-9. *HNRNPU* knockdown in Raji cells dysregulates expression of the MYC and E2F family transcription factors

(A) *HNRNPU* knockdown was validated in Raji cells using qPCR. Results showing normalized expression of *HNRNPU* to the geometric mean of *Actin* and *GAPDH*. and showed a 50% reduction in mRNA levels. (B) *HNRNPU* knockdown significantly lowered *MYC* expression. (C) Western blot for significantly altered MYC and E2F family members from the RNA-seq data. Activating E2F expression most noticeably induced. (D) qPCR results of significantly altered E2F family members follows similar trend to that of the RNA-seq data. Significance was calculated by Student's t-test ($p^* \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$). All qPCRs performed in three biological replicates; western blots shown as representative image from two biological replicates.

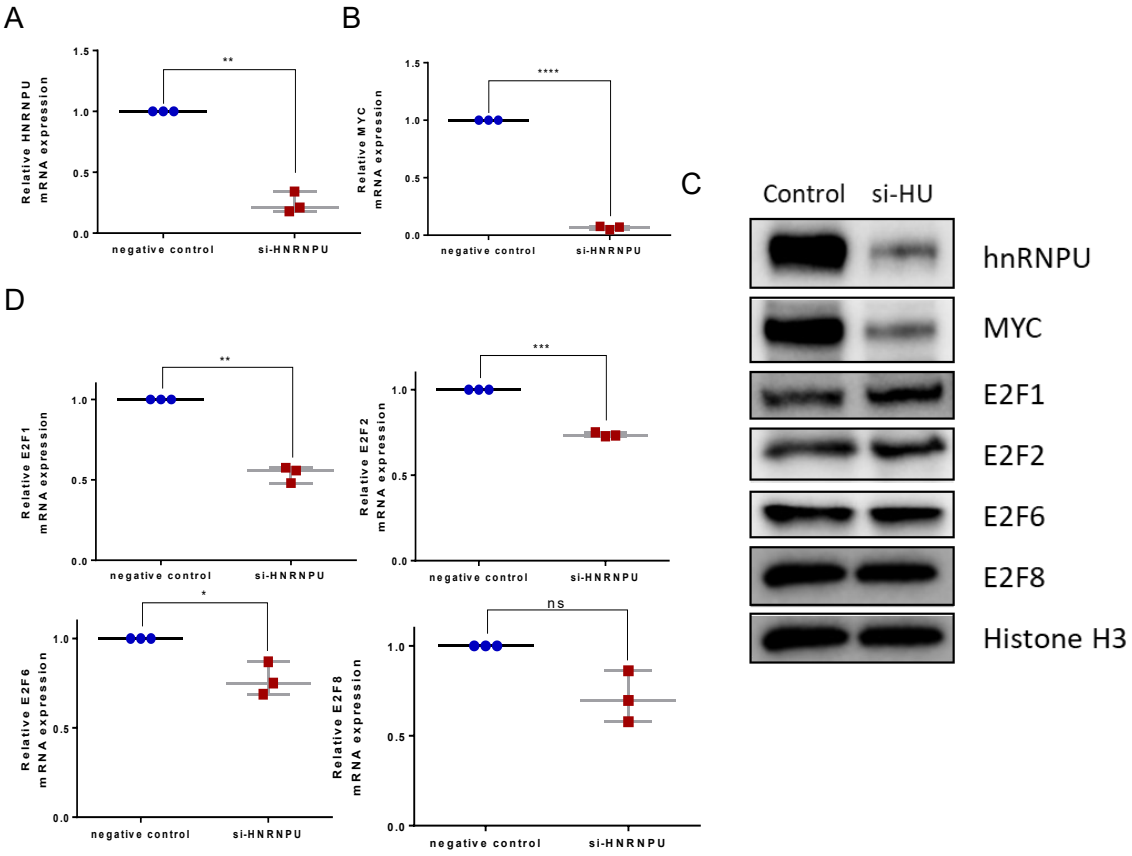


Figure 3-10. *HNRNPU* knockdown in HEK293 cells reduced MYC and E2F mRNA expression

(A) *HNRNPU* knockdown was validated with qPCR. (B) *HNRNPU* knockdown resulted in significant decrease in *MYC* mRNA expression. (C) *HNRNPU* knockdown resulted in decreased hnRNPU and MYC protein but no changes in E2F family proteins was observed. (D) Relative mRNA of all E2F family members was lower due to *HNRNPU* knockdown. Significance was calculated by Student's t-test ($p^* \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $p^{****} \leq 0.0001$). All qPCRs performed in three biological replicates; western blots shown as representative image from two biological replicates.

3.7. hnRNPU eCLIP-seq reveals its direct interaction with MYC and E2F1 mRNA

Given that *HNRNPU* mutations resulted in the dysregulated expression and splicing of transcripts encoding other RNA binding proteins and proteins with broad roles in gene expression and RNA metabolism, we were curious whether the observed changes in the E2F and MYC family members were under the direct control of hnRNPU or mediated through downstream intermediaries. We examined the splicing of MYC and E2F family transcripts, but we did not detect any alterations in their splicing patterns, indicating their expression was modulated by a different mechanism.

We hypothesised that by investigating whether hnRNPU directly binds to these transcripts, we could gain insights into how hnRNPU regulates their expression. To do this, we reanalyzed publicly available hnRNPU-enhanced crosslinking immunoprecipitation (eCLIP) data, followed by sequencing in two cell lines, HEPG2 and K562 cells from the ENCODE project. eCLIP techniques enable the precise detection of protein-RNA interaction sites, based on distinct events occurring at crosslink sites. We utilized PureCLIP to identify individual crosslink sites, considering both regions enriched in protein-bound fragments and eCLIP-specific truncation patterns, providing nucleotide-level precision on where the RNA binding protein is bound on the transcript[166]. Crosslink sites are then scored on a scale from 0 - 3. A rating of 0 signifies sites without crosslinks or enrichment, 1 indicates non-enriched sites with crosslinks, 2 refers to enriched sites without crosslinks, and 3 represents sites that are both enriched and crosslinked.

Using this approach, we analysed hnRNPU eCLIP targets relative to an IgG control to account for the extent of background binding and found that both *MYC* and *E2F1* had extensive hnRNPU binding in intron 1 of their transcripts in both K562 and HEPG2 cells. These were scored as 3 on the PureCLIP scale indicating that hnRNPU was directly bound to these transcripts and these regions were enriched in the IP. hnRNPU bound to E2F2 and E2F6 in K562 cells but not HEPG2 cells and no binding was observed for E2F8 (**Figure 3-11**). These findings provided evidence that hnRNPU may directly regulate *MYC* and *E2F1* expression by binding to their mRNA. Since no changes in their splicing was observed we hypothesised that hnRNPU might play a role in regulating the stability of these mRNA transcripts and/or their rate of translation into proteins.

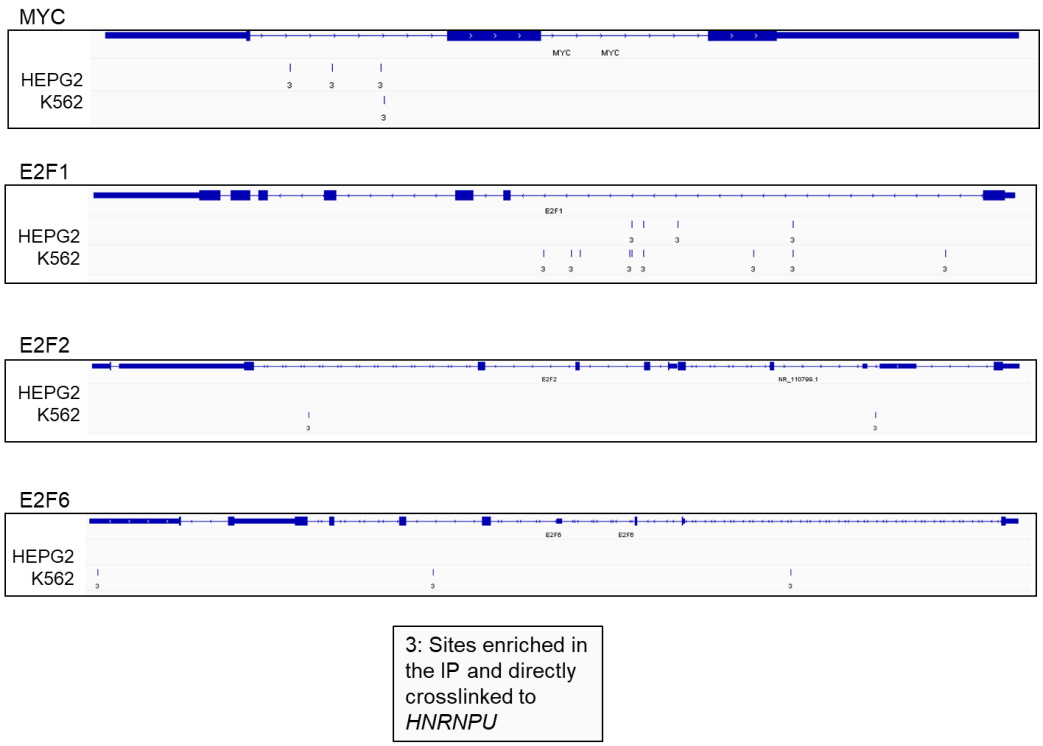


Figure 3-11. hnRNPU e-CLIP sites in HEPG2 and K562 for MYC and E2F family transcripts

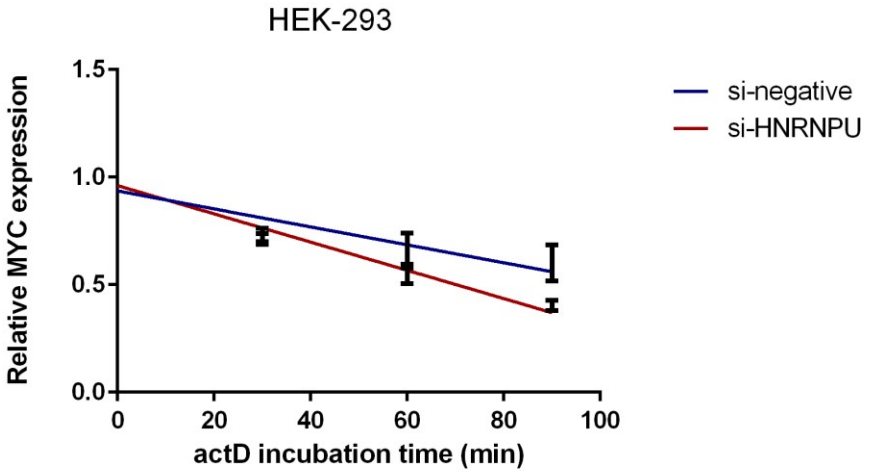
Using PureClip hnRNPU binding sites and co-immunoprecipitated RNA were analyzed. We found that hnRNPU binds extensively to intron 1 of the MYC and E2F1 targets in both HEPG2 and K562 cells but binds E2F2 and E2F6 transcripts to a lesser degree and this binding is only observed in K562 cells.

3.8. hnRNPU regulates MYC expression through enhancing *MYC* mRNA stability

Since dysregulated *MYC* expression is pertinent to BL and HGBL-DH/TH pathobiology, our next goal was to mechanistically understand how hnRNPU modulates the expression of MYC. In a recent study it was shown that hnRNPU enhances the expression of specific genes by binding to, and stabilizing their mRNA[193]. Given that hnRNPU knockdown causes a decrease in *MYC* expression at both the mRNA and protein levels and it binds directly to the *MYC* mRNA transcript, we hypothesised that this modulation was related to transcript stability. In a recent study conducted in U2OS osteosarcoma cells, it was shown that hnRNPU is associated with the mRNA binding IGF2BP1 complex promoting stabilization of the *MYC* mRNA via its coding region instability determinant (CRD)[194]. The CRD is located at the last 249 nucleotides of the *MYC* coding sequence and when bound to these proteins, shields it from endonucleolytic attack, stabilizing the mRNA[195].

To test this in our cell lines we conducted an actinomycin D pulse chase assay where we incubated WT, hnRNPU siRNA depleted HEK293 cells or *HNRNPU* mutant Raji cells with the transcriptional inhibitor, actinomycin D. Given that under normal conditions MYC mRNA half-life is roughly 10-30 minutes in various cell lines, we chased MYC mRNA expression every 30 minutes after the addition of actinomycin D[196,197]. Interestingly, our results show that MYC stability is higher than these estimates in both HEK293 and Raji cells at 104.5 minutes and 72.7 minutes, respectively. We found that MYC mRNA stability was lower in hnRNPU-depleted HEK293 cells as well as CRISPR38 cells. We did not notice any differences in MYC mRNA stability in CRISPR 36 cells. The half-life of MYC decreases from 104.5 minutes to 70.26 minutes in HEK293 cells (**Figure 3-12A**) and to 70.9 and 53.7 minutes in CRISPR 36 and 38 respectively (**Figure 3-12B**). This suggests that hnRNPU mediated regulation of *MYC* can at least in part be attributed to the decrease in its mRNA stability.

A



B

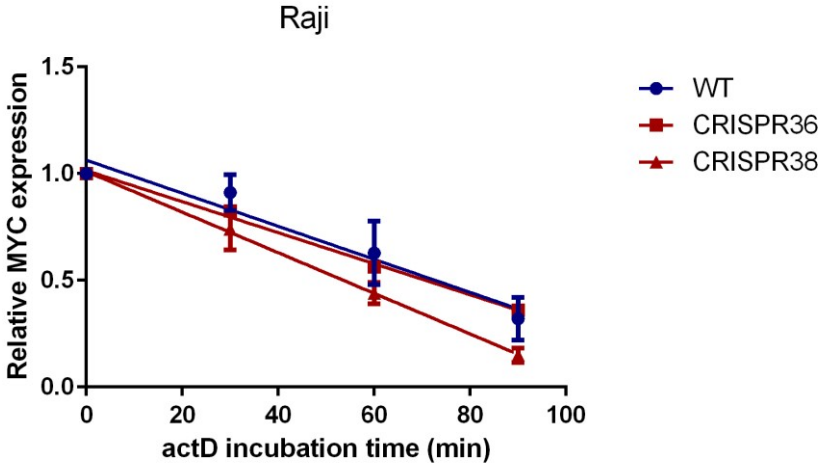


Figure 3-12. MYC half-life is lower in hnRNPU-depleted cells than hnRNPU WT cells

To determine MYC mRNA stability in hnRNPU-depleted cells we tested two cell lines **(A)** WT or *HNRNPU* knocked down HEK293 cells and **(B)** WT or *HNRNPU* mutant Raji cells. We incubated these cells with 4uM actinomycinD to block transcription and then harvested mRNA at 4 time points, 0,30,60 and 90 minutes. The relative mRNA expressions of MYC to that at time point 0 were then plotted and were subjected to linear regression analysis. MYC half-life decreases from 104.5 minutes to 70.26 minutes in HEK293 cells and from 70.9 to 53.7 minutes in *HNRNPU* mutant CRISPR 38.

3.9. Intron 1 of *MYC* is important for hnRNPU mediated modulation

As hnRNPU's regulation of MYC could not be fully attributed to regulating its mRNA stability, we remained intrigued by the pattern of hnRNPU binding to the *MYC* transcript, as indicated by the eCLIP-seq data. hnRNPU's role in the regulation of MYC stability was previously reported to be through its interaction with a MYC-CRD stabilization complex. This CRD is at the C-terminus of its coding sequence, however we observe direct hnRNPU binding in intron 1 of the *MYC* transcript through eCLIP-seq in HEPG2 and K562 cells (**Figure 3-13A**). In addition, even though MYC half-life was lower in the hnRNPU-depleted cells, the difference was only moderate. Based on these observations, we hypothesize that hnRNPU may exert influence on *MYC* expression through mechanisms that extend beyond regulation of its stability through the CRD. An additional confounder is that the region of MYC intron 1 where hnRNPU binding has been detected coincides with the region where a subset of MYC-IgH translocations occur.

MYC translocation breaks fall into three categories, translocations within intron 1 (Type I), those upstream of the MYC locus (Type-II), and those downstream of it (Type III) (**Figure 3-13B**). *HNRNPU* nonsense mutations only appear in tumours that have upstream or downstream *MYC* translocations but are never observed in patients with *MYC* intronic translocations **(Figure 3-13C**). Since the region 5' of the intron 1 breakpoint contains many regulatory motifs, we hypothesize that hnRNPU may have an additional role in MYC regulation by interacting with these regions. This together with the notion that hnRNPU has binding sites in intron 1 of the *MYC* transcript made us consider that perhaps this region is necessary for hnRNPU mediated modulation of *MYC* and in its absence *HNRNPU* mutations do not serve a functional advantage. Looking deeper at the crosslinking immunoprecipitation sites of hnRNPU on MYC we found that hnRNPU binds poly G-tracts in intron 1 that are capable of forming G-quadruplexes, a secondary structure that is a known hnRNPU target[198].

To explore this further, we cloned two minigene vectors in the PRL-TK backbone containing either the MYC locus, spanning from its 5'UTR, all exons and introns until the stop codon or a version that lacks intron 1 (**Figure 3-14A**). To test what effect loss of intron-1 has on hnRNPU mediated MYC regulation, we transfected the minigene

constructs into HEK293 cells that ether had WT *HNRNPU* (treated with a negative control siRNA) or cells with siRNA depleted *HNRNPU*. When comparing the RNA expression of *MYC* relative to a control that had no minigene expression (mock) we found that cells transfected with full-length minigene had on average 33% minigene expression relative to those with WT levels of hnRNPU (**Figure 3-14B-C**). Cells transfected with Δintron-1 minigene had 52% expression of the minigene in *HNRNPU*-depleted cells. This shows that while *HNRNPU* loss affects both intron-1 containing and intron-1 lacking transcripts, the effect is appreciably greater in those containing intron-1. This demonstrates that hnRNPU may have additional roles in regulating MYC through interactions with its intron-1 region, however further analyses need to be performed using these constructs in *HNRNPU* mutant Raji cells or in *HNRNPU* depleted lymphoma cell lines. Nevertheless, these results align with the hypothesis that an intron 1 translocation might diminish hnRNPU's regulatory control over MYC exerted through intron 1. Further investigations into this observation could offer insights into why instances involving intronic translocations never coincide with *HNRNPU* nonsense mutations.

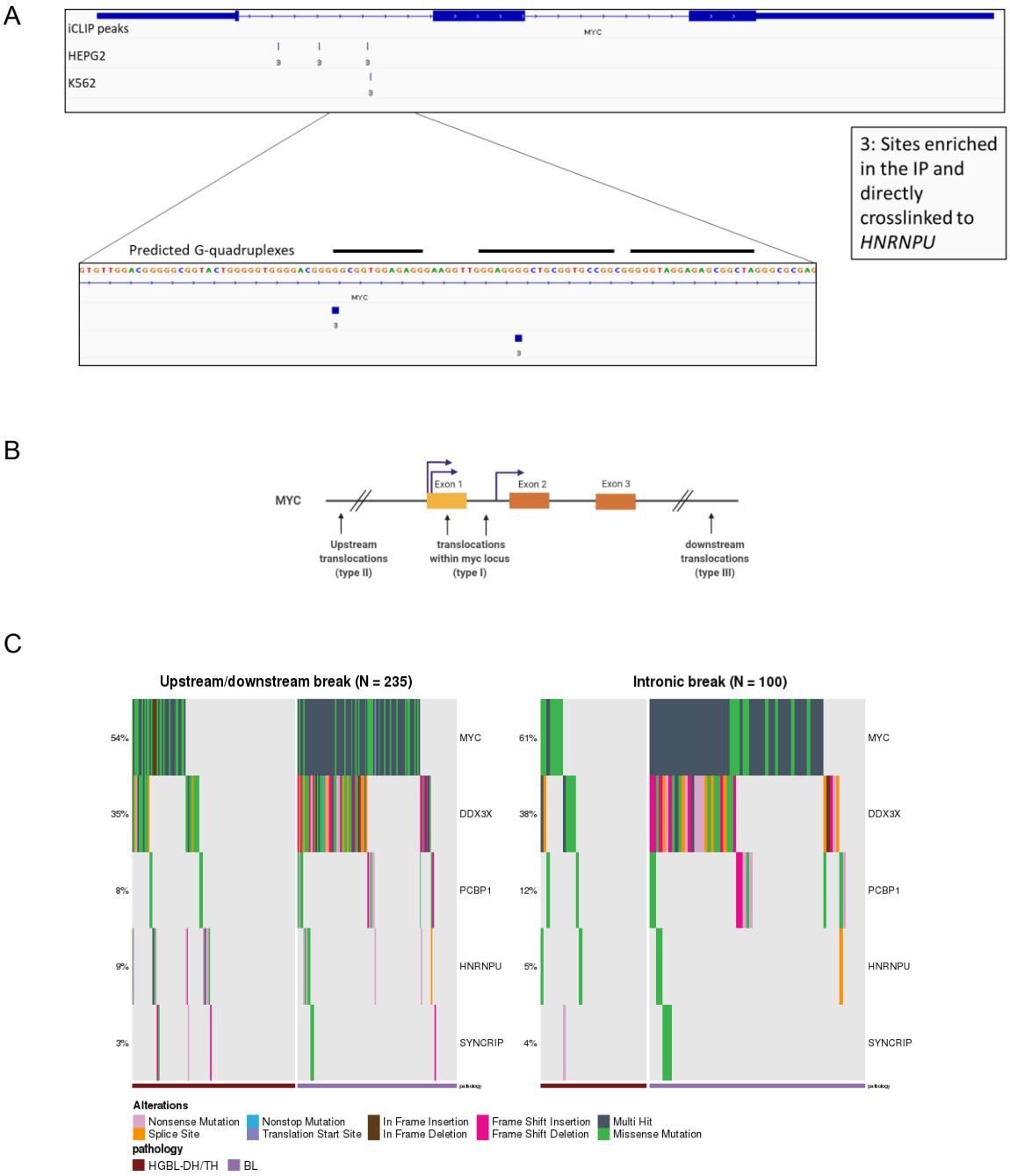


Figure 3-13. Mutual exclusivity of *HNRNPU* nonsense mutations and *MYC* intronic translocations suggest putative hnRNPU binding site in intron 1

(A) hnRNPU eCLIP-seq data resolved to single nucleotide resolution with PureCLIP. Data suggests hnRNPU binds directly to the *MYC* locus in intron-1 poly G-tracts predicted to fold into G-quadruplexes, a secondary structure that can be bound by hnRNPU. B) *MYC* translocation breakpoints. *MYC* translocations can occur at three distinct sites, either upstream, downstream, or within intron 1 of the gene. C) *HNRNPU* nonsense mutations do not appear in cases with *MYC* intronic breaks.

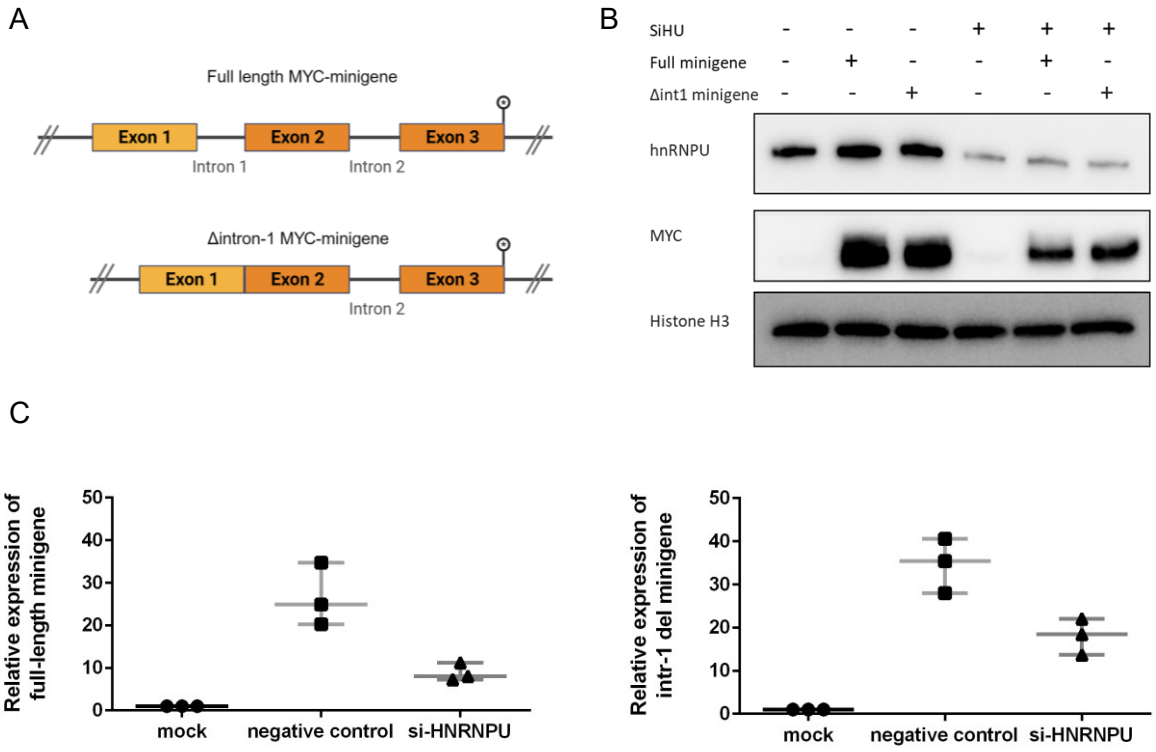


Figure 3-14. *HNRNPU* regulation of MYC is regulated partly through interactions with intron-1

(A) Schematic representation of the two *MYC* minigenes, either containing the entire MYC locus or lacking intron 1. Full length *MYC* minigene is depleted more at both the (B) protein and (C) mRNA level when *HNRNPU* is knocked down. All qPCRs performed in three biological replicates; western blots shown as representative image from three biological replicates.

3.10. *HNRNPU* mutations moderates MYC mediated proteotoxic stress

Considering that *MYC* serves as the primary driver of BL and HGBL-DH/TH, and dysregulated *MYC* expression is a characteristic associated with aggressiveness, we questioned why these tumours would exhibit *HNRNPU* mutations given that this decreases *MYC* expression. We postulated that this phenomenon might be related to the fact that the high expression of *MYC* can have potent and acute effects on programmed cell death, and one of the mechanisms of evading this is to reduce *MYC* levels to tolerable amounts, especially during early stages of lymphomagenesis. Since Raji cells are established lymphoma cell lines, they do not exhibit markers of active UPR and proteotoxic stress. We questioned if we overexpressed hnRNPU, and by consequence increase MYC expression, would we be able to trigger proteotoxic stress due to elevated MYC.

To test this, we first designed an hnRNPU-eGFP expression vector. The eGFP was cloned at the N-terminus of the coding sequence with considerations taken to ensure it does not disrupt the native folding of the protein. Since hnRNPU is a large protein, we wanted to validate that adding GFP to its N-terminus would still allow it to localize to the nucleus. We overexpressed hnRNPU in Raji cells and looked at eGFP expression in the transfected cells. Through this we confirmed the localization of hnRNPU to be diffuse nuclear consistent with what has previously been described (**Figure 3-15A**). After enriching for positive transfectants with FACS, we looked at expression of MYC with qPCR and western blot analysis. Indeed, overexpression of hnRNPU led to increased MYC mRNA and protein as expected (**Figure 3-15 B-C**). However, this increase led to a decrease in cell viability as shown in **Figure 3-15D**.

In a recent study by Gong *et. al*, it was reported that overexpression of MYC in primary germinal centre B cells results in an increased global protein synthesis which induces the unfolded protein response (UPR), suggestive of proteotoxic stress[105]. Physiologically, an imbalance between the folding load of nascent proteins entering the ER and the capacity of the ER to handle this load causes ER stress[199]. Intrinsic stresses in the tumour, such as oncogenic activation increase the levels of misfolded proteins in the ER, triggering the activation of UPR pathways[200]. Active UPR first attempts to restore protein folding homeostasis. If the stress is sustained UPR triggers apoptosis. It is

thought that UPR-mediated killing may eliminate cells expressing high oncogene activation and select those cells with either lower oncogene activation or defective death pathways[201]. We wanted to see if hnRNPU mediated overexpression of MYC induces the UPR in the Raji cell line. We analyzed the splicing of XBP-1, as a marker of activated UPR and ER stress. XBP-1 (X-box-binding protein 1) splicing is one the primary responses to ER stress in B cells, specifically observed in plasma cells that are secreting and producing large amounts of antibody[202]. Upon activation of the ER stress sensor IRE1, XBP-1 is spliced in the cytoplasm to produce a transcript encoding the XBP-1 protein aimed at restoring protein folding homeostasis[203]. Conversely, if the stress cannot be reversed, IRE1 also triggers apoptosis.

We designed a custom digital droplet PCR assay to quantify the abundance of spliced XBP-1 in the cells transfected with either the empty vector or hnRNPU-eGFP (**Figure 3-16A**). The splicing ratio was defined as the ratio of the spliced XBP-1 over the total abundance of all XBP-1 transcripts. In the hnRNPU overexpressed cells we noticed an increase in the splicing ratio suggesting an induction of the UPR and proteotoxic stress (**Figure 3-16B**). This proteotoxic stress may potentially be leading to decreased cell viability in the Raji cells transfected with *HNRNPU*. Based on these results, we propose a model where *HNRNPU* mutations moderate *MYC* expression, thus buffering MYC-induced apoptosis and proteotoxic stress in the B cells that harbour them.

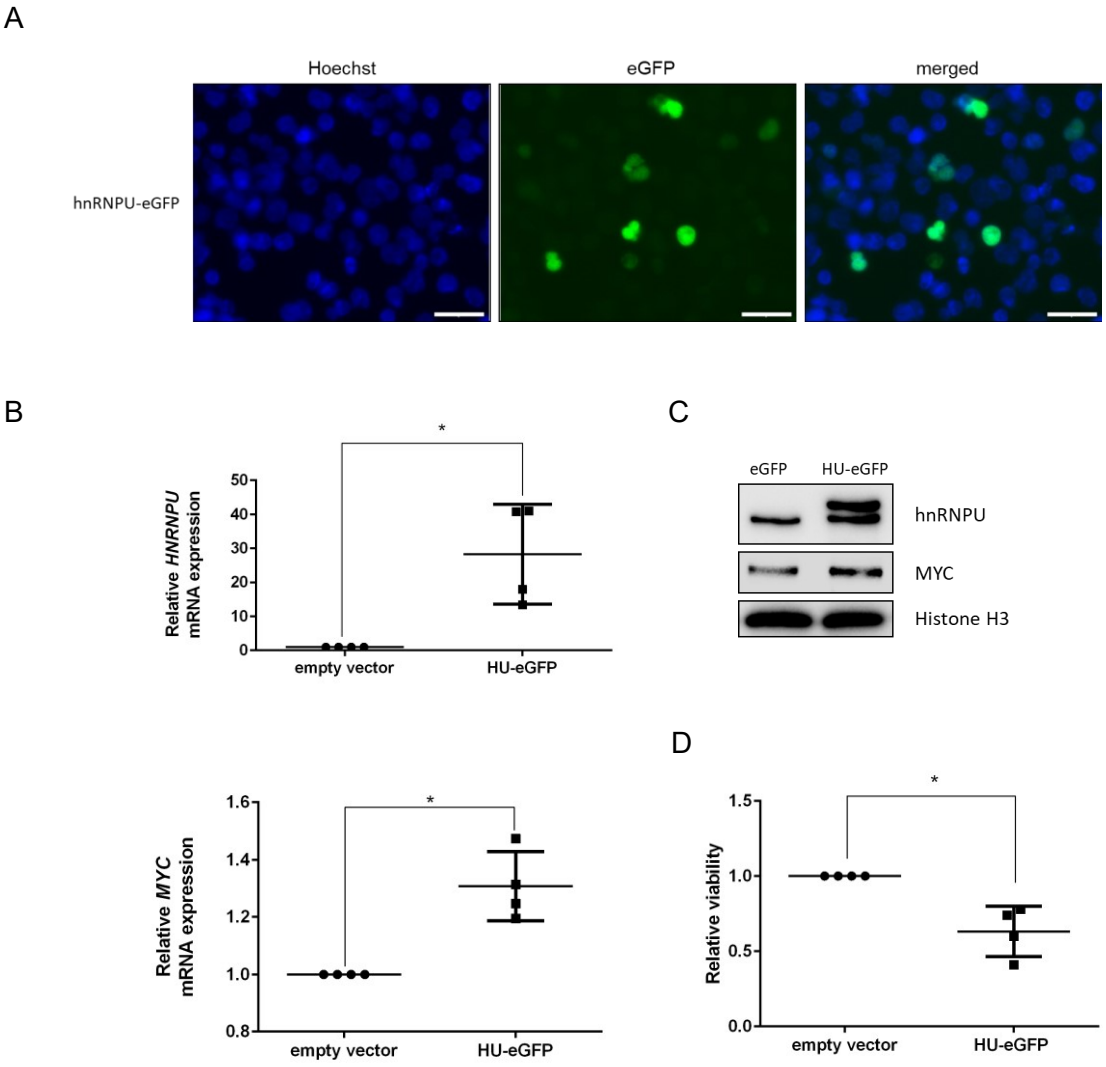


Figure 3-15. *HNRNPU*-eGFP overexpression results in significantly decreased cell viability

(A) Exogenous *HNRNPU* expression in Raji cells shows protein localizing to the nucleus. Scale bars correspond to 10um. *HNRNPU* overexpression shows increased (B) MYC mRNA and (C) protein expression (D) hnRNPU overexpression significantly decreases cell proliferation in overexpressed cells relative to those transfected with just the empty vector. Significance was calculated by Student's t-test (*$p \leq 0.05$). All qPCRs performed in four biological replicates; western blots shown as representative image from four biological replicates.

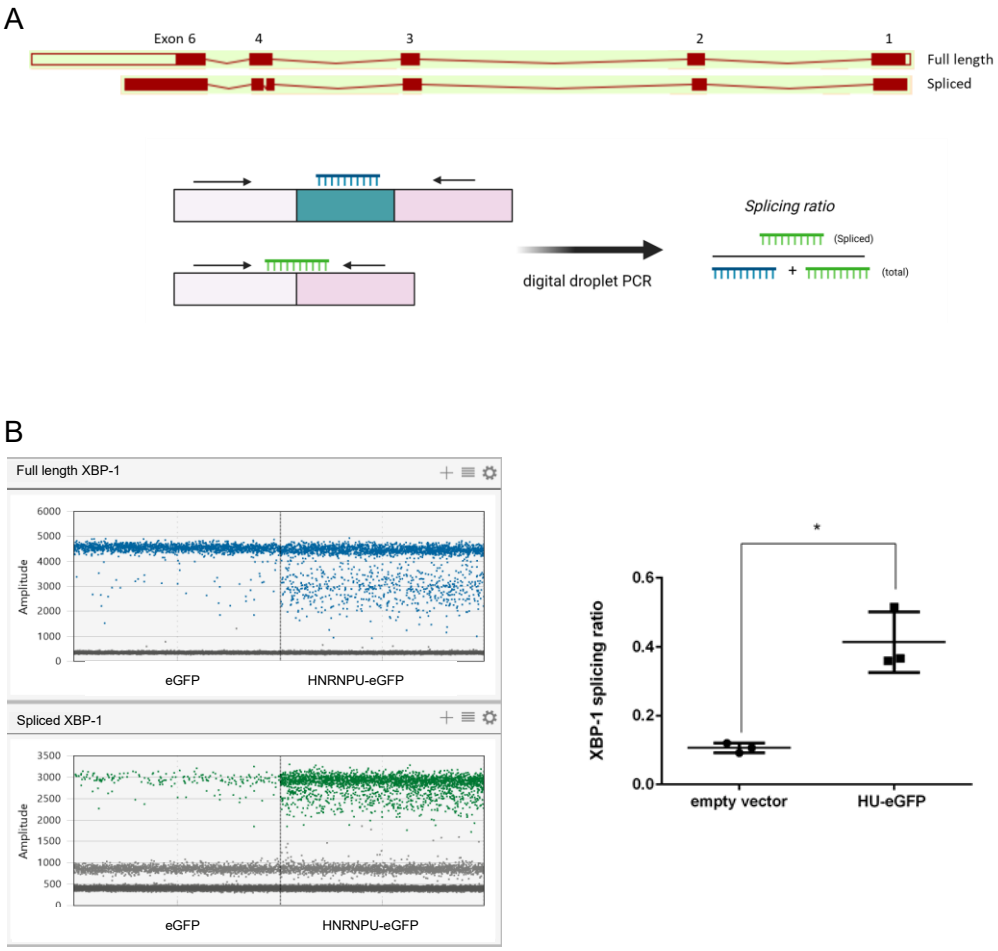


Figure 3-16. *HNRNPU* overexpression enhanced splicing of XBP-1 suggestive of proteotoxic stress

(A) A custom digital droplet PCR assay was designed to identify the abundance of spliced XBP-1 in the cells transfected with either the empty vector containing *eGFP* or *HNRNPU-eGFP*. The splicing ratio was calculated as the ratio of the spliced XBP-1 over all XBP-1 transcripts. (B) In the overexpressed cells we noticed an increase in the splicing of XBP-1 suggesting an induction of the UPR and proteotoxic stress. Significance was calculated by Student's t-test (*$p \leq 0.05$). ddPCR performed in three biological replicates.

3.11. *HNRNPU* mutant cell lines are sensitive to E2F inhibitors

Since we do not see a difference in cell viability between the *HNRNPU* mutant lines relative to WT, despite reduction of *MYC* expression, we were curious whether these cells were now even more dependent on E2Fs to promote cell cycle entry and promotion. To test this we used a pan-E2F inhibitor, HLM006474, which had shown efficacy in blocking cell growth in a melanoma culture model[204]. HLM006474 blocks DNA binding activity of all E2F complexes therefore we thought it would be an ideal drug candidate to use in our experiments. Concentrations ranging from 5-100 μM E2F inhibitor was added to the cells and then viability was measured relative to a DMSO control. We found that after 24 hours of incubation with the drug, there was lower cell viability in the two *HNRNPU* mutant cell lines. *HNRNPU* WT Raji cells also responded well to the drug with an IC50 of 31.6 μM. The IC50 of CRISPR36 and CRISPR38 was significantly lower at 18.57 μM and 13.8 μM respectively.

Since these cell lines are MYC driven yet can tolerate lower MYC expression due to increased E2F expression, we wondered if reducing MYC in these cell lines even further would result in cell death. To test this we used another inhibitor, Mycro3, which is a MYC-MAX dimerization inhibitor and showed efficacy in mouse models of pancreatic cancer[205]. When tested in WT and *HNRNPU* mutant Raji cells we do not see any notable differences in cell viability between WT and mutants. In fact, the drug did not appear to work very well with only partial cell death even at 100 μM. The IC50s were variable and very high therefore the efficacy of this drug was determined to be low in both WT and hnRNPU mutant cases.

In summary, these findings demonstrate that cells with *HNRNPU* mutations exhibit a greater reliance on E2F activities to propel the cell cycle, making them selectively more vulnerable to E2F small molecule inhibitors. Moreover, we observed that E2F inhibitors also effectively target WT Raji cells. Considering the limited efficacy of MYC inhibitors in these cell lines and the shared pathway between MYC and E2F for cell cycle initiation and proliferation, E2Fs emerge as promising therapeutic targets for these lymphomas.

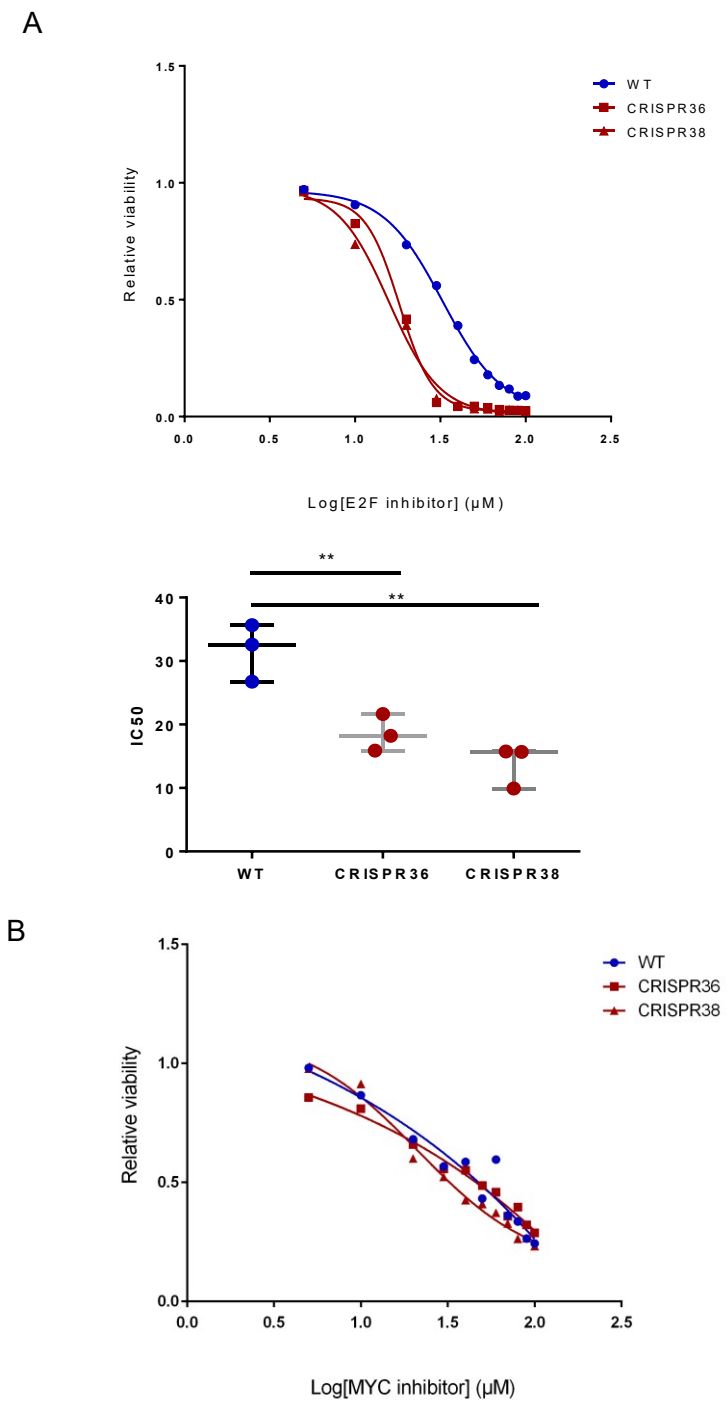


Figure 3-17. *HNRNPU* mutations render cells more sensitive to E2F inhibitors
Comparison of IC50 for Raji WT and *HNRNPU* CRISPR-mutant (36,38) cell lines for (A) HLM006474 (E2F inhibitor), (B) Mycro3 (MYC-inhibitor). Representative dose response curve for each drug is shown and IC50 calculated from three independent drug dose-curve assays. All data represent the mean of three independent experiments ± SD. Significance was calculated by Student's t-test (**p ≤ 0.01).

Chapter 4. Discussion

4.1. Summary of research findings

Here, we have characterized the role of *HNRNPU* mutations in *MYC*-driven B-cell lymphomas and show that they provide a selective advantage to the cells harbouring them. Through re-analysis of multiple large datasets, we have clarified the prevalence of *HNRNPU* mutations in BL and HGBL-DH/TH. Although these mutations predominantly occur in *MYC* translocated lymphomas, we observed rare examples in FL and DLBCL. This highlights the importance of understanding the functional role of hnRNPU and mutations that affect its expression in hematological malignancies.

Initial studies of *HNRNPU* knockout mice showed altered pre-mRNA splicing and post-natal development of the heart and brain[152,206]. Interestingly in these studies, hnRNPU targets and cellular functions were largely tissue specific and mice with knockout *HNRNPU* displayed early lethality (within 2 weeks of birth). In an essential gene screen conducted in B-cell lymphoma cell lines, *HNRNPU* was also found to be an essential gene in all cell lines tested[207]. It was thus apparent that *HNRNPU* mutations may have distinct effects in MYC-driven B cells and harbouring mutations in a single allele would play an important role in determining these functions.

4.1.1. *HNRNPU* is haploinsufficient and mutations dysregulate cell cycle dynamics

The classification of a gene as either an oncogene or a tumor suppressor has long been a cornerstone of cancer biology. However, as our understanding of these genes deepens, this categorization has become increasingly challenging as they do not always reflect the reality of aberrant gene expression. Many RNA binding proteins, particularly hnRNPs, have the capacity to regulate both tumor suppressive and oncogenic pathways, and both their overexpression and knockdown results in cell proliferation and apoptotic defects[208]. This together with hnRNPs having tissue specific functional roles suggest that they might have different target repertoire as well as functional affects depending on the tissue type. As an example, hnRNP K's role as tumor suppressor has recently been described in acute myeloid leukemia and demonstrated in a haploinsufficient mouse model[209]. In contrast, data from other clinical

correlation studies suggest that hnRNP K may be more fittingly described as an oncogene, due to its increased expression in a variety of other malignancies[210,211]. Similar to hnRNPK, hnRNPU's involvement in cancer to date has primarily featured its role as an oncogene, where its increased expression is associated with tumor progression[159,160]. It was thus intriguing that we observed an enrichment of nonsense mutations in a single allele of mutant tumors. To understand the role of these mutations further and in the context of B cells, I introduced heterozygous frameshift mutations in the genome of the BL cell line Raji. Through this we found that frame shift deletions in a single allele reduces the amount of hnRNPU protein in B cells that harbour them. Through RNA-seq, and subsequent transcriptomic analysis relative to cells with wild type *HNRNPU*, we observed that *HNRNPU* mutations result in significant alterations to the transcriptomic and splicing landscapes of the mutant cells. This is consistent with the model that *HNRNPU* is haploinsufficient in these cell lines as a single copy of the functional allele at the *HNRNPU* locus in heterozygous combination with the variant allele is insufficient to produce the wild-type phenotype. Overall, most of these changes indicated a propensity toward cell cycle promotion by upregulation of genes that facilitate the G1S and G2M transition.

4.1.2. *HNRNPU* modulates E2F expression in B cells

Our transcriptomic analysis revealed altered expression of several transcription, cell proliferation, and cell cycle related genes. Notably, we found enrichment of members of the E2F transcription factor family and of cell cycle related targets of the E2F family. While we do not observe elevated expression of E2F activators or lower expression of E2F repressors at the mRNA level, the overexpression of E2Fs' downstream targets implies that the root of dysregulation within this pathway likely stems from the E2Fs themselves.

So far, eight mammalian E2F family members have been identified. The expression of various E2Fs is differentially regulated throughout the cell cycle, and some E2Fs are expressed in a cell type-specific manner[212]. The family is divided into two subfamilies: E2Fs 1–3A are activators of transcription, whereas E2Fs 3B–8 act as repressors[213]. The current model of E2F function in the cell cycle is that the oscillatory nature of cell cycle-dependent gene expression is driven by sequential binding of E2F activators and repressors to target promoters[214]. This results in the transcription of

numerous target genes involved in DNA replication and cell cycle progression[215]. Under normal conditions, activating E2Fs allow for transcription of cell cycle progression genes promoting G1/S phase of the cell cycle. E2F6 then initiates transcriptional repression of E2F targets during S phase. During replicative stress, E2F6 is inhibited prolonging E2F activity and E2F target expression in S phase, enabling cells to recover from replicative stress[214]. In the *HNRNPU* mutant cell lines, we uncover an increased mRNA and protein expression of activating E2Fs, E2F1 and E2F2, alongside reduced expression of the repressive E2F, E2F6. We validated this modulation through *HNRNPU* knockdown in Raji cells where we observed a similar result. However, it is worth noting that while we observe similarities in E2F expression in the knockdown cells, it is not as pronounced as in the *HNRNPU* mutants, potentially suggesting that this dysregulation intensifies due to significant cellular adaptation. In addition, this regulatory pattern seems to be specific to certain tissues, as we do not observe similar regulatory patterns in HEK293 cells. This observation could offer insights into why hnRNPU mutations that decrease its expression have not been reported in malignancies beyond lymphomas. The precise mechanism behind the dysregulated expression of E2Fs is not fully understood. However, preliminary evidence from crosslinking immunoprecipitation in K562 and HEPG2 cells (**Figure 3-11**) suggests that hnRNPU may directly bind to the activating E2F, E2F1, and affect its expression through this interaction. We hypothesise that the exact mechanism of this modulation is through a combination of direct interactions with the E2F transcripts and downstream effects on gene expression and alternative splicing changes, which collectively impact various transcription regulators.

Accumulating evidence has implied that E2Fs are closely linked with tumorigenesis in a variety of cancer types and were found to be closely associated with poor prognosis of cancer patients[216,217]. In MYC-driven lymphomas, particularly Burkitt lymphoma, E2F1, is highly expressed relative to non-transformed B cells and reduction in its expression decreases tumor formation and proliferation[93]. Ectopic expression of E2F1 is sufficient to drive quiescent cells to S phase[218]. In plasmablastic lymphoma, the upregulation of E2F1, E2F3 and E2F8 correlates significantly with an increased expression of MKI67 (Ki67), indicating that E2F exerts a crucial role in the abnormal proliferation[219]. Generally, E2F pathway dysregulation in cancers is a direct result of alterations in the RB1 pathway however we see no alterations in the RB1 mRNA highlighting a possible RB1-independent mechanism for E2F dysregulation in these

lymphomas. We do however see altered splicing in CDC25A, a crucial cell cycle regulator. We observe enhanced skipping of exon 6 in CDC25A in the mutant cell lines relative to the wildtype. This event removes two poly-ubiquitination sites, ultimately increasing in its protein stability. CDC25A removes inhibitory phosphorylation from various cyclins that phosphorylate the Rb repressor rendering its dissociation from E2F, and promoting G1/S transition[180]. We hypothesised that the combined effects of hnRNPU elevating E2F expression as well as altered expression of cell cycle regulators such as CDC25A contribute to increased cell cycle progression in the *HNRNPU* mutant cells.

4.1.3. Putative role of hnRNPU mutations in buffering MYC induced cell stress

Considering the overall increase in the expression of genes associated with the cell cycle and the elevation of genes that are known MYC targets, it was striking that *HNRNPU* mutant cell lines exhibited a reduction in *MYC* expression itself. This observation raises curiosity, especially in the context of *MYC*-driven lymphomas like BL and HGBL-DH/TH, as to why mutations that lower *MYC* expression would be enriched in these tumors. Several studies have shown that targeting MYC expression in MYC-driven lymphoma cells results in diminished tumor cell proliferation through cell cycle arrest, necrosis and apoptosis[190,220]. We find that despite the reduced *MYC* levels in *HNRNPU* mutants, we still see similar viability to that of WT. This could be attributed to the upregulation of E2Fs, which may serve to compensate for the reduced MYC expression.

Studies have shown that in contrast to normal cells, MYC-overexpressing human cancer cells need E2F activity for their survival[92]. As E2Fs are recognized as transcriptional targets of *MYC* and several CDKs that facilitate E2F-mediated cell cycle initiation[218], we propose that hnRNPU modifies the gene expression profile in a manner that uncouples MYC from E2F expression. This uncoupling would enable the cell cycle to proceed, even with reduced MYC levels. This phenomenon is thought to be cell specific as it is not observed in HEK293 cells. In the case of HEK293 cells, the decrease in *HNRNPU* results in diminished expression of both MYC and all E2Fs, aligning with expectations when MYC regulates their expression.

When considering the reasons behind the reduced MYC expression in these cells, it is worth noting that MYC has intrinsic tumor suppressive role in these lymphomas where its overexpression has potent and acute effects on programmed cell death. This is thought to be mediated in part through triggering proteotoxic stress[100,221]. Elevated expression of MYC in primary germinal centre B-cells has been shown to upregulate genes involved in the unfolded protein response, indicative that the cells are undergoing proteotoxic stress[105]. During tumor development, the protein synthesis rate is tightly regulated to sustain cell survival. Increased protein synthesis requires concomitant increased folding capacity to avoid proteotoxicity[222]. MYC activation constitutes a global increase in protein synthesis and intrinsic stress that places further weight on protein synthesis and secretion[100]. If the cell fails to restore protein folding homeostasis, the resulting chronic ER stress and persistent UPR signaling leads to cell death[223]. In MYC-driven lymphomas, loss-of-function mutations have been observed in *DDX3X*, a RNA-binding protein that promotes global protein synthesis[105]. These mutations are thought to dampen MYC induced proteotoxic stress by lowering global protein synthesis. Similarly, another way of lowering MYC induced cell stress is by lowering *MYC* expression itself. It was recently shown through *in vivo* studies that modest elevation of MYC enhances transformation as expected, however, robust overexpression of *MYC* leads to a dramatic increase in apoptosis[106]. The requirement to keep activated oncogenes at low levels to avoid engaging tumor suppression is likely an important selective pressure governing the early stages of tumor microevolution. We hypothesize that *HNRNPU* mutations lower MYC expression to a tolerable amount to help buffer MYC induced proteotoxic stress. In addition, we have not observed instances of *HNRNPU* and *DDX3X* nonsense mutations in the same tumor likely pointing to their similar roles.

In addition, data suggests that overexpression of *HNRNPU* in the BL cell line Raji results in elevated *MYC* expression and is accompanied by proteotoxic stress. While we can not attribute this proteotoxicity to MYC alone, it is important to note that it results in lower cell viability. Because this is an established B-cell NHL cell line, it has undergone substantial cellular adaptations to endure proteotoxic stress and handle elevated MYC expression. As a result, we do not observe active UPR in WT cells. However, when we modify this condition, by increasing MYC expression, it may lead to an imbalance in the mechanisms in place to regulate protein folding homeostasis and consequently, lead to

increased proteotoxic stress. Given their existing low levels of proteotoxic stress as they have presumably already stabilized the level of proteotoxic stress induced by MYC, the reduced expression of hnRNPU in the *HNRNPU* mutants does not exhibit any additional decreases in the activation of the UPR pathway.

It has been reported that reducing E2Fs expression affects MYC's ability to induce proliferation, while its capacity to trigger apoptosis remains unaffected[90]. Taken together we propose a model where *HNRNPU* mutations might be an early event in lymphomagenesis where they would provide a selective advantage to cells by lowering *MYC* expression to a tolerable amount to help buffer MYC induced cell stress yet sill allowing for the cells to enter the cell cycle by means of upregulated E2F activity (**Figure 4-1**).

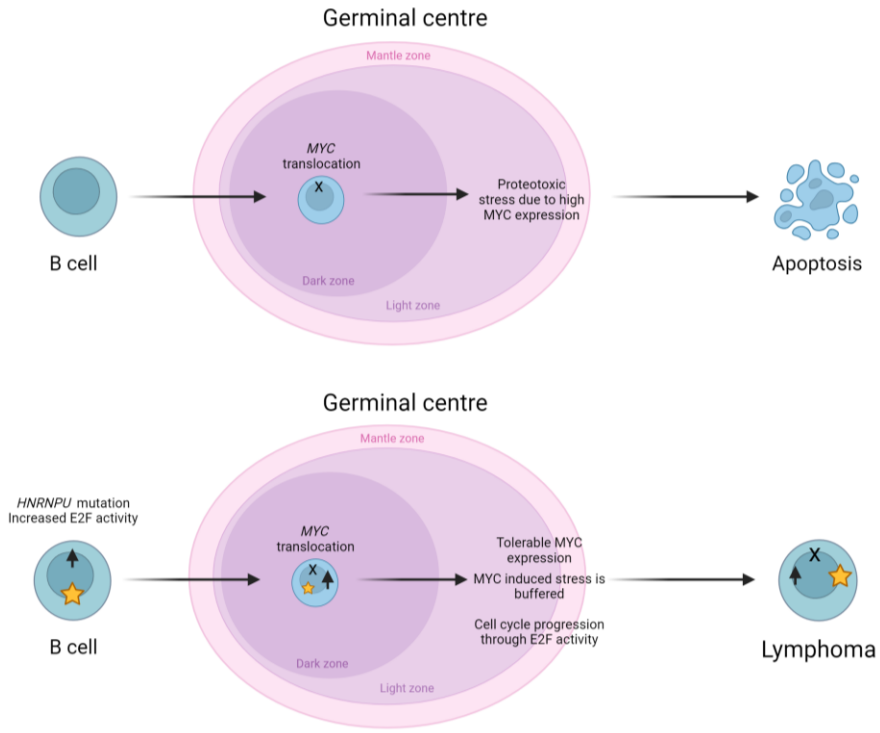


Figure 4-1. Model for hnRNPU mediated regulation of MYC and E2Fs allowing for lymphoma progression
(A) Aberrant upregulation of MYC results in global proteotoxicity that can be harmful to tumor cells during early stages of tumor development. (B) Proposed mechanism of HNRNPU *mutations* and subsequent decrease in its expression, buffering MYC induced proteotoxic stress and activation of the E2F pathway to allow for cell cycle progression.

4.2. Potential clinical relevance

Current therapies for MYC driven lymphomas are based on highly aggressive short-term combination chemotherapy regimens and the toxicities associated with these therapies can be severe. In addition, patients with MYC translocations such as HGBL DH/TH have a higher tendency to be refractory to conventional therapy options or experience early relapse. Such patients would therefore benefit from new therapeutic strategies that exploit molecular vulnerabilities specific to MYC-driven lymphomas. We found *HNRNPU* mutations in 5.6% of patients with Burkitt and 12% patients with HGBL-DH/TH. Our results highlight hnRNPU-mediated regulation of MYC and its downstream effects as a new layer of regulation particularly relevant in MYC-driven lymphomas. We show that *HNRNPU* mutations reduce expression of MYC, thereby increasing dependence on E2Fs for tumor cell survival. We hypothesised that this would result in the mutant cells being more sensitive to E2F inhibitors. Indeed, we observed a greater sensitivity to E2F inhibitors in the *HNRNPU* mutant cell lines relative to WT. Interestingly we also noticed that the WT cell line was also sensitive to E2F inhibition with an IC50 of 31.6 µM. This is consistent with previous reports of E2F dependence in MYC driven lymphomagenesis where it was shown that loss of a single or both alleles of E2F1 slowed MYC driven proliferation and lymphomagenesis in mice[90].

Considering the reduced MYC expression in the *HNRNPU* mutant cell lines, we hypothesised that this lowered expression might be at a level where the cells can evade MYC-driven cellular stress while maintaining their viability. In such a scenario, we postulated that this reduction could make the cells more responsive to MYC inhibition. MYC was historically thought to be “undruggable” owing to its lack of traditional small molecule binding pockets and its intrinsically disordered structure[224]. Over the years, various strategies to inhibit MYC have been presented, including inhibition of either MYC expression or MYC/MAX dimerization. The role of MAX as an obligate partner of MYC has led to significant efforts towards the development of MYC-MAX heterodimer inhibitors. We used a MYC-MAX dimerization inhibitor, Mycro3, to test whether the *HNRNPU* mutant cell lines were more sensitive to MYC inhibition. Mycro3 showed efficacy in pre-clinical trials where daily treatment by oral gavage increased survival in a mouse model of pancreatic ductal carcinoma and xenografts of human pancreatic cancer cells[225]. We tested efficacy of this compound to inhibit proliferation of Raji and

HNRNPU mutant Raji cells, however the response was variable and associated with unacceptably high IC50s, indicating that its efficacy was limited. Another MYC/MAX dimerization inhibitor, OMO-103, has shown efficacy in inhibiting lymphoma cell proliferation in vitro, This compound is currently being evaluated in a stage I clinical trial for pancreatic, bowel, and non-small cell lung cancers[226,227]. We aim to test the efficacy of OMO-103 directly or in combination with E2F inhibitors in future experiments in both MYC driven lymphoma cell lines and *HNRNPU* mutants.

Aside from direct targeting of MYC or E2Fs in lymphomas, many other indirect targeting avenues also exist. Many of hnRNPs, like hnRNPU, are essential genes therefore therapeutic windows may be limited due to toxicity at higher doses[228–230]. However, in tumours with loss of heterozygosity events in essential genes, monoallelic inactivation of the allele retained in tumors can selectively kill cancer cells but not somatic cells, which retain both alleles[231]. It should therefore be considered that lower concentrations of targeted therapeutics to essential hnRNPs, could confer increased tumor sensitivity in some cases.

Another feature of hnRNPU is its propensity to interact with G-rich sequences of RNA capable of folding into stable tertiary structures known as G-quadruplexes (rG4). Dynamic changes in rG4 folding and unfolding dictate the biological significance of these structures in regulating RNA metabolism[232]. The changes in folded and unfolded state of rG4s have been causally linked with malignancy[233]. Herviou *et al.* found an over-representation of unfolded rG4s in hnRNP H/F-binding sites at translational regulatory regions of mRNAs involved in pathways associated to genome instability and DNA damage[234]. To date, around 1000 small molecules that target rG4s have been reported, affording another opportunity to perturb hnRNP-mediated regulatory processes[235]. Specific small molecule rG4 inhibitors such as pyridostatin blocks the interaction between hnRNP H1 and a rG4 within the pre-mRNA of the EWSR1 protein involved in sarcoma translocations[236]. In addition, the small molecule emetine binds and disrupts G-quadruplexes, globally inhibiting rG4 dependent alternative splicing[237]. Our data warrant exploration of the potential utility of G4-inhibition in B-cell lymphomas, particularly those with *HNRNPU* mutations.

Further, since many hnRNPs are shown to contribute to oncogenesis by means of aberrant alternative splicing, splice-switch oligonucleotides (SSO) may also be

promising therapeutic strategies. SSOs are antisense oligonucleotides, typically 15–30 nucleotides long, capable of base-pairing with the pre-mRNA target site to sterically block the binding of splicing factors.[238] An important feature of SSOs are their chemically modified bases allowing pre-mRNA-SSO complex to preclude RNase H cleavage.[239] This feature allows SSOs to modify splicing without necessarily altering the abundance of the mRNA transcript. We identified a potential pathogenic splice variant in CDC25A, an important regulator of cell cycle progression. Further understanding of this splicing event and its consequences to the cell may be helpful in exploring the potential for SSOs as a therapeutic approach.

4.3. Ongoing work and future directions

Thus far, my analyses have relied mostly on the BL cell line Raji. The results have demonstrated that *HNRNPU* mutations in this specific cell line led to gene expression changes consistent with those observed in patients. This positive correlation has motivated us to extend our investigation by evaluating the efficacy of E2F inhibition in cell lines carrying *HNRNPU* mutations. In my ongoing research, I have identified two HGBL-DH/TH cell lines, namely SU-DHL-4 and Seraphina, both harbouring monoallelic deletions in the *HNRNPU* gene locus. Experiments testing the efficacy of E2F inhibitors on these cell lines in comparison to other lymphoma cell lines and correlating these with the gene expression of E2Fs should therefore be pursued. Given that Mycro3 did not prove to be the optimal inhibitor for MYC, testing the sensitivity of *HNRNPU* mutant cell lines to MYC inhibition using a different inhibitor, OMO-103, will provide further insights. This inhibitor has demonstrated greater efficacy in studies of MYC-driven lymphomas. To support these findings, examining the effects of depleting MYC in these cell lines via siRNA to determine whether they exhibit heightened sensitivity or resistance to MYC depletion should also be pursued. In addition, since *HNRNPU* is an essential gene in MYC driven cell lines, further reducing its expression could selectively kill the cancer cells. Cell lines with the *HNRNPU* mutation could therefore be used to perform a small molecule library screen to find lead compounds targeting specific vulnerabilities in these cell lines.

Secondly, in this thesis I have made use of multiple large data sets from the ENCODE project consortium to mechanistically understand how hnRNPU modulates both MYC and E2F expression. I have found preliminary evidence of hnRNPU binding to

both MYC and E2F1 transcripts in HEPG2 and K562 cells. However, as is apparent from the data presented in this thesis, hnRNPU's cellular targets and functions are largely tissue specific. While it is encouraging to see evidence of hnRNPU binding MYC and E2F1 in two distantly related cell lines, it does not escape the possibility that this might not be the case in BL cell lines. Unbiased experiments to globally delineate the hnRNPU interactome in specific BL cell lines will provide important context. Similarly, the relationship between hnRNPU binding of MYC and E2Fs should be studied in additional BL and HGBL-DH/TH cell lines to determine the generality of the current findings.

A HyperTRIBE approach would therefore be beneficial in globally understanding hnRNPU's interactome. HyperTRIBE expresses a fusion protein consisting of an RBP and the catalytic domain of the RNA-editing enzyme ADAR (adenosine deaminase acting on RNA) (ADARcd), which marks target RNA transcripts by converting adenosine to inosine near the RBP binding sites[240]. These marks can be subsequently identified via high-throughput sequencing. I have cloned an hnRNPU-ADAR construct and a version that has a catalytically dead version of ADAR as a control. The control would bind RNA transcripts through hnRNPU but would lack RNA editing. I have expressed these constructs in mammalian cells and have confirmed their expression and localization to the nucleus. Expressing these constructs in MYC driven cell lines and subsequent RNA-seq will help resolve the interactome of hnRNPU in these cells. For local binding analysis, an RNA-IP strategy can be implemented where RNA that co-immunoprecipitates with hnRNPU can be analysed using quantitative PCR methods. Implementing this strategy with the MYC minigene constructs may help resolve whether hnRNPU binding is increased or diminished due to the presence or absence of intron 1.

Lastly, the *HNRNPU* mutations were created in mature and established B-cell lymphoma cell lines. These typically carry an extensive and biased mutational repertoire and had presumably learned to adapt to MYC induced stress, having undergone many rounds of clonal evolution and cellular adaptation. Since we hypothesise that *HNRNPU* mutations are early events, allowing the cells to handle MYC induced cell stress, it would be interesting to observe the effects of hnRNPU loss in primary human GC B cells grown ex vivo in a co-culture designed to mimic the GC microenvironment. Dan Hodson's group has recently described an optimized methodology that facilitates proliferation and efficient transduction of non-malignant, primary, human GC B cells ex vivo[241]. Using such a strategy would allow us to model the initial stages of human lymphomagenesis in

a manner that is not possible in established cell lines. We can then directly test if overexpressing *MYC* in *HNRNPU* mutant GC B cells results in lower levels of cell death and proteotoxicity.

4.4. Conclusions

HNRNPU mutations are novel recurrent driver mutations that appear specific to MYC-translocated B-cell lymphomas. In the cell of origin, *HNRNPU* acts as a modulator of MYC expression, and these mutations appear to diminish this role. We propose a model wherein *HNRNPU* mutations buffer the negative effect of excessive MYC expression caused by a translocation, thus reducing MYC-induced proteotoxic stress and apoptosis. Introduction of *HNRNPU* mutations in MYC-translocated lymphomas leads to a reliance on E2Fs to promote cell cycle progression. These relationships can be explored to identify novel therapeutic strategies for *HNRNPU*-mutant tumors. This thesis also emphasizes the utilization of next-generation sequencing technologies for an unbiased identification of transcriptomic changes attributed to mutated regulatory proteins. It underscores the significance of considering the cell-type-specific functions of these proteins and studying their roles within specific contexts. The results obtained from these analyses ultimately offer insights into the mechanisms contributing to disease, supporting ongoing efforts to establish in vitro and in vivo models for future therapeutic investigations.

References

1. Faguet, G. An Historical Overview: From Prehistory to WWII. in *The Conquest of Cancer: A distant goal* (ed. Faguet, G.) 13–33 (Springer Netherlands, 2015). doi:10.1007/978-94-017-9165-6_2.
2. Brenner, D. R. *et al.* Projected estimates of cancer in Canada in 2020. *CMAJ* **192**, E199–E205 (2020).
3. Hanahan, D. & Weinberg, R. A. The Hallmarks of Cancer. *Cell* **100**, 57–70 (2000).
4. Hanahan, D. Hallmarks of Cancer: New Dimensions. *Cancer Discov.* **12**, 31–46 (2022).
5. Casás-Selves, M. & DeGregori, J. How cancer shapes evolution, and how evolution shapes cancer. *Evolution* **4**, 624–634 (2011).
6. Seferbekova, Z., Lomakin, A., Yates, L. R. & Gerstung, M. Spatial biology of cancer evolution. *Nat. Rev. Genet.* **24**, 295–313 (2023).
7. Guan, Y.-F. *et al.* Application of next-generation sequencing in clinical oncology to advance personalized treatment of cancer. *Chin. J. Cancer* **31**, 463–470 (2012).
8. Raphael, B. J., Dobson, J. R., Oesper, L. & Vandin, F. Identifying driver mutations in sequenced cancer genomes: computational approaches to enable precision medicine. *Genome Med.* **6**, 5 (2014).
9. Only three driver gene mutations are required for the development of lung and colorectal cancers. https://www.pnas.org/doi/10.1073/pnas.1421839112 doi:10.1073/pnas.1421839112.
10. Stratton, M. R., Campbell, P. J. & Futreal, P. A. The cancer genome. *Nature* **458**, 719–724 (2009).

11. Driver and Passenger Mutations in Cancer | Annual Review of Pathology: Mechanisms of Disease. https://www.annualreviews.org/doi/10.1146/annurev-pathol-012414-040312.

12. Todd, R. & Wong, D. T. Oncogenes. *Anticancer Res.* **19**, 4729–4746 (1999).

13. Bell, J. C. Oncogenes. *Cancer Lett.* **40**, 1–5 (1988).

14. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nat. Med.* **10**, 789–799 (2004).

15. Oncogenes in development | Development | The Company of Biologists. https://journals.biologists.com/dev/article/99/4/449/51826/Oncogenes-in-development.

16. Chen, K., Zhang, Y., Qian, L. & Wang, P. Emerging strategies to target RAS signaling in human cancer therapy. *J. Hematol. Oncol.J Hematol Oncol* **14**, 116 (2021).

17. Hinds, P. W. & Weinberg, R. A. Tumor suppressor genes. *Curr. Opin. Genet. Dev.* **4**, 135–141 (1994).

18. Ozaki, T. & Nakagawara, A. Role of p53 in Cell Death and Human Cancers. *Cancers* **3**, 994–1013 (2011).

19. Rivlin, N., Brosh, R., Oren, M. & Rotter, V. Mutations in the p53 Tumor Suppressor Gene. *Genes Cancer* **2**, 466–474 (2011).

20. Stehr, H. *et al.* The structural impact of cancer-associated missense mutations in oncogenes and tumor suppressors. *Mol. Cancer* **10**, 54 (2011).

21. Sharma, S. & Lichtenstein, A. Premature Termination Codon Mutations in Chronic Lymphocytic Leukemia (CLL): Identification of Potential Tumor Suppressors in CLL Including E-Cadherin. *Blood* **106**, 2935 (2005).

22. Iengar, P. An analysis of substitution, deletion and insertion mutations in cancer genes. *Nucleic Acids Res.* **40**, 6401–6413 (2012).

23. Kazanets, A., Shorstova, T., Hilmi, K., Marques, M. & Witcher, M. Epigenetic silencing of tumor suppressor genes: Paradigms, puzzles, and potential. *Biochim. Biophys. Acta BBA - Rev. Cancer* **1865**, 275–288 (2016).

24. Haploinsufficiency for tumour suppressor genes: when you don't need to go all the way - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S0304419X04000022?via%3Dihub.

25. Carbone, A. Cancer Classification at the Crossroads. *Cancers* **12**, 980 (2020).

26. Saggioro, M., D'Angelo, E., Bisogno, G., Agostini, M. & Pozzobon, M. Carcinoma and Sarcoma Microenvironment at a Glance: Where We Are. *Front. Oncol.* **10**, 76 (2020).

27. Gilbert, N. F., Cannon, C. P., Lin, P. P. & Lewis, V. O. Soft-tissue sarcoma. *J. Am. Acad. Orthop. Surg.* **17**, 40–47 (2009).

28. Whiteley, A. E., Price, T. T., Cantelli, G. & Sipkins, D. A. Leukaemia: a model metastatic disease. *Nat. Rev. Cancer* **21**, 461–475 (2021).

29. Eaton, M. & Fox, R. Surgical biopsy in lymphoma. *ANZ J. Surg.* **75**, 810–812 (2005).

30. Roth, S. L. *et al.* Contiguous pattern spreading in patients with Hodgkin's disease. *Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol.* **47**, 7–16 (1998).

31. Berger, M. F. & Mardis, E. R. The emerging clinical relevance of genomics in cancer medicine. *Nat. Rev. Clin. Oncol.* **15**, 353–365 (2018).

32. Armitage, J. O., Gascoyne, R. D., Lunning, M. A. & Cavalli, F. Non-Hodgkin lymphoma. *Lancet Lond. Engl.* **390**, 298–310 (2017).

33. Lee, S. Cancer statistics. *Canadian Cancer Society* https://cancer.ca/en/research/cancer-statistics.

34. Ansell, S. M. Non-Hodgkin Lymphoma: Diagnosis and Treatment. *Mayo Clin. Proc.* **90**, 1152–1163 (2015).

35. Jamil, A. & Mukkamalla, S. K. R. Lymphoma. in *StatPearls* (StatPearls Publishing, 2023).

36. Küppers, R. & Hansmann, M.-L. The Hodgkin and Reed/Sternberg cell. *Int. J. Biochem. Cell Biol.* **37**, 511–517 (2005).

37. Sapkota, S. & Shaikh, H. Non-Hodgkin Lymphoma. in *StatPearls* (StatPearls Publishing, 2023).

38. Lewis, W. D., Lilly, S. & Jones, K. L. Lymphoma: Diagnosis and Treatment. *Am. Fam. Physician* **101**, 34–41 (2020).

39. Neelapu, S. S. Non-Hodgkin Indolent B-Cell Lymphoma. in *60 Years of Survival Outcomes at The University of Texas MD Anderson Cancer Center* 241–250 (Springer, New York, NY, 2013). doi:10.1007/978-1-4614-5197-6_22.

40. Jeong, S. H. Treatment of indolent lymphoma. *Blood Res.* **57**, 120–129 (2022).

41. Sehn, L. H. Introduction to a review series: the paradox of indolent B-cell lymphoma. *Blood* **127**, 2045–2046 (2016).

42. Cabanillas, F. Can we cure indolent lymphomas? *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **3**, 2655–2659 (1997).

43. A Clinical Evaluation of the International Lymphoma Study Group Classification of Non-Hodgkin's Lymphoma. *Blood* **89**, 3909–3918 (1997).

44. Pathobiology of follicular lymphoma - UpToDate. https://www.uptodate.com/contents/pathobiology-of-follicular-lymphoma.

45. Gisselbrecht, C. [Aggressive lymphomas]. *Rev. Prat.* **43**, 1648–1653 (1993).

46. Young, R. M. *et al.* Taming the Heterogeneity of Aggressive Lymphomas for Precision Therapy. *Annu. Rev. Cancer Biol.* **3**, 429–455 (2019).

47. Nowakowski, G. S. & Czuczman, M. S. ABC, GCB, and Double-Hit Diffuse Large B-Cell Lymphoma: Does Subtype Make a Difference in Therapy Selection? *Am. Soc. Clin. Oncol. Educ. Book* e449–e457 (2015) doi:10.14694/EdBook_AM.2015.35.e449.

48. Frazer, J. K. *et al.* Excellent outcomes in children and adolescents with CNS+ Burkitt lymphoma or other mature B-NHL using only intrathecal and systemic chemoimmunotherapy: results from FAB/LMB96 and COG ANHL01P1. *Br. J. Haematol.* **185**, 374–377 (2019).

49. Čubranić, A. *et al.* BURKITT LYMPHOMA IN GASTROINTESTINAL TRACT: A REPORT OF TWO CASES. *Acta Clin. Croat.* **58**, 386–390 (2019).

50. Lynch, D. T., Koya, S., Acharya, U. & Kumar, A. Mantle Cell Lymphoma. in *StatPearls* (StatPearls Publishing, 2023).

51. Bhatt, R. & Desai, D. S. Plasmablastic Lymphoma. in *StatPearls* (StatPearls Publishing, 2023).

52. De Silva, N. S. & Klein, U. Dynamics of B cells in germinal centres. *Nat. Rev. Immunol.* **15**, 137–148 (2015).

53. Huang, C. Germinal Center Reaction. *Adv. Exp. Med. Biol.* **1254**, 47–53 (2020).

54. Gitlin, A. D. *et al.* T cell help controls the speed of the cell cycle in germinal center B cells. *Science* **349**, 643–646 (2015).

55. Calado, D. P. *et al.* MYC is essential for the formation and maintenance of germinal centers. *Nat. Immunol.* **13**, 1092–1100 (2012).

56. Calado, D. P. *et al.* The cell-cycle regulator c-Myc is essential for the formation and maintenance of germinal centers. *Nat. Immunol.* **13**, 1092–1100 (2012).

57. Shats, I. *et al.* Expression level is a key determinant of E2F1-mediated cell fate. *Cell Death Differ.* **24**, 626–637 (2017).

58. Béguelin, W. *et al.* EZH2 enables germinal centre formation through epigenetic silencing of CDKN1A and an Rb-E2F1 feedback loop. *Nat. Commun.* **8**, 877 (2017).

59. Ranuncolo, S. M., Polo, J. M. & Melnick, A. BCL6 represses CHEK1 and suppresses DNA damage pathways in normal and malignant B-cells. *Blood Cells. Mol. Dis.* **41**, 95–99 (2008).

60. Phan, R. T. & Dalla-Favera, R. The BCL6 proto-oncogene suppresses p53 expression in germinal-centre B cells. *Nature* **432**, 635–639 (2004).

61. Martin, A. & Scharff, M. D. Somatic hypermutation of the AID transgene in B and non-B cells. *Proc. Natl. Acad. Sci.* **99**, 12304–12308 (2002).

62. Khodabakhshi, A. H. *et al.* Recurrent targets of aberrant somatic hypermutation in lymphoma. *Oncotarget* **3**, 1308–1319 (2012).

63. Pasqualucci, L. *et al.* BCL-6 mutations in normal germinal center B cells: Evidence of somatic hypermutation acting outside Ig loci. *Proc. Natl. Acad. Sci.* **95**, 11816–11821 (1998).

64. Peng, H.-Z. *et al.* Nonimmunoglobulin Gene Hypermutation in Germinal Center B Cells. *Blood* **93**, 2167–2172 (1999).

65. Okazaki, I., Kinoshita, K., Muramatsu, M., Yoshikawa, K. & Honjo, T. The AID enzyme induces class switch recombination in fibroblasts. *Nature* **416**, 340–345 (2002).

66. Zahn, A. *et al.* Activation induced deaminase C-terminal domain links DNA breaks to end protection and repair during class switch recombination. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E988-997 (2014).

67. Insight into the origin and clonal history of B-cell tumors as revealed by analysis of immunoglobulin variable region genes - Stevenson - 1998 - Immunological Reviews - Wiley Online Library. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-065X.1998.tb01446.x?casa_token=KPVZpOX6lxQAAAAA:FqjV_ftBKLDxuSy2ig2R3SfOk4TQDntF_iP05f0tYelpe7dJn0j07-slD6xpXPhP4rQYoZaec0oqeg.

68. Pasqualucci, L. *et al.* Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* **412**, 341–346 (2001).

69. Bödör, C. *et al.* Aberrant somatic hypermutation and expression of activation-induced cytidine deaminase mRNA in mediastinal large B-cell lymphoma. *Br. J. Haematol.* **129**, 373–376 (2005).

70. Okazaki, I. *et al.* Constitutive expression of AID leads to tumorigenesis. *J. Exp. Med.* **197**, 1173–1181 (2003).

71. Hernández-Verdin, I. *et al.* Pan-cancer landscape of AID-related mutations, composite mutations, and their potential role in the ICI response. *Npj Precis. Oncol.* **6**, 1–14 (2022).

72. Lenz, G. *et al.* Aberrant immunoglobulin class switch recombination and switch translocations in activated B cell–like diffuse large B cell lymphoma. *J. Exp. Med.* **204**, 633–643 (2007).

73. Guikema, J. E. J., Schuuring, E. & Kluin, P. M. Structure and Consequences of IGH Switch Breakpoints in Burkitt Lymphoma. *JNCI Monogr.* **2008**, 32–36 (2008).

74. Sun, R.-F., Yu, Q.-Q. & Young, K. H. Critically dysregulated signaling pathways and clinical utility of the pathway biomarkers in lymphoid malignancies. *Chronic Dis. Transl. Med.* **4**, 29–44 (2018).

75. Knoepfler, P. S. Myc Goes Global: New Tricks for an Old Oncogene. *Cancer Res.* **67**, 5061–5063 (2007).

76. Walhout, A. J., Gubbels, J. M., Bernards, R., van der Vliet, P. C. & Timmers, H. T. c-Myc/Max heterodimers bind cooperatively to the E-box sequences located in the first intron of the rat ornithine decarboxylase (ODC) gene. *Nucleic Acids Res.* **25**, 1493–1501 (1997).

77. Wang, C. *et al.* Alternative approaches to target Myc for cancer treatment. *Signal Transduct. Target. Ther.* **6**, 1–14 (2021).

78. Wasylishen, A. R. & Penn, L. Z. Myc. *Genes Cancer* **1**, 532–541 (2010).

79. Smith, S. M., Anastasi, J., Cohen, K. S. & Godley, L. A. The impact of MYC expression in lymphoma biology: Beyond Burkitt lymphoma. *Blood Cells. Mol. Dis.* **45**, 317–323 (2010).

80. Nguyen, L., Papenhausen, P. & Shao, H. The Role of c-MYC in B-Cell Lymphomas: Diagnostic and Molecular Aspects. *Genes* **8**, 116 (2017).

81. Rosenwald, A. *et al.* Prognostic Significance of MYC Rearrangement and Translocation Partner in Diffuse Large B-Cell Lymphoma: A Study by the Lunenburg Lymphoma Biomarker Consortium. *J. Clin. Oncol.* **37**, 3359–3368 (2019).

82. Johnson, N. A. *et al.* Lymphomas with concurrent BCL2 and MYC translocations: the critical factors associated with survival. *Blood* **114**, 2273–2279 (2009).

83. Bretones, G., Delgado, M. D. & León, J. Myc and cell cycle control. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1849**, 506–516 (2015).

84. The Roles of Cyclin-Dependent Kinases in Cell-Cycle Progression and Therapeutic Strategies in Human Breast Cancer - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7139603/.

85. Dong, P. *et al.* Division of labour between Myc and G1 cyclins in cell cycle commitment and pace control. *Nat. Commun.* **5**, 4750 (2014).

86. Attwooll, C., Denchi, E. L. & Helin, K. The E2F family: specific functions and overlapping interests. *EMBO J.* **23**, 4709–4716 (2004).

87. García-Gutiérrez, L., Delgado, M. D. & León, J. MYC Oncogene Contributions to Release of Cell Cycle Brakes. *Genes* **10**, 244 (2019).

88. Myc/Max/Mad regulate the frequency but not the duration of productive cell cycles | EMBO reports. https://www.embopress.org/doi/full/10.1093/embo-reports/kve251.

89. Leone, G. *et al.* Myc requires distinct E2F activities to induce S phase and apoptosis. *Mol. Cell* **8**, 105–113 (2001).

90. Baudino, T. A. *et al.* Myc-Mediated Proliferation and Lymphomagenesis, but Not Apoptosis, Are Compromised by E2f1 Loss. *Mol. Cell* **11**, 905–914 (2003).

91. Swier, L. J. Y. M., Dzikiewicz-Krawczyk, A., Winkle, M., van den Berg, A. & Kluiver, J. Intricate crosstalk between MYC and non-coding RNAs regulates hallmarks of cancer. *Mol. Oncol.* **13**, 26–45 (2019).

92. Santoni-Rugiu, E. *et al.* E2F activity is essential for survival of Myc-overexpressing human cancer cells. *Oncogene* **21**, 6498–6509 (2002).

93. Molina-Privado, I. *et al.* E2F1 expression is deregulated and plays an oncogenic role in sporadic Burkitt's lymphoma. *Cancer Res.* **69**, 4052–4058 (2009).

94. Pelengaris, S., Khan, M. & Evan, G. c-MYC: more than just a matter of life and death. *Nat. Rev. Cancer* **2**, 764–776 (2002).

95. Death Pathways Triggered by Activated Ras in Cancer Cells - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3098755/.

96. Wu, Z., Zheng, S. & Yu, Q. The E2F family and the role of E2F1 in apoptosis. *Int. J. Biochem. Cell Biol.* **41**, 2389–2397 (2009).

97. Evan, G. I. *et al.* Induction of apoptosis in fibroblasts by c-myc protein. *Cell* **69**, 119–128 (1992).

98. Hoffman, B. & Liebermann, D. A. Apoptotic signaling by c-MYC. *Oncogene* **27**, 6462–6472 (2008).

99. Curti, L. & Campaner, S. MYC-Induced Replicative Stress: A Double-Edged Sword for Cancer Development and Treatment. *Int. J. Mol. Sci.* **22**, 6168 (2021).

100. Zhang, T., Li, N., Sun, C., Jin, Y. & Sheng, X. MYC and the unfolded protein response in cancer: synthetic lethal partners in crime? *EMBO Mol. Med.* **12**, e11845 (2020).

101. Vecchio, E. *et al.* Insights about MYC and Apoptosis in B-Lymphomagenesis: An Update from Murine Models. *Int. J. Mol. Sci.* **21**, 4265 (2020).

102. Wang, X. J. *et al.* P53 expression correlates with poorer survival and augments the negative prognostic effect of MYC rearrangement, expression or concurrent MYC/BCL2 expression in diffuse large B-cell lymphoma. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **30**, 194–203 (2017).

103. Salam, D. S. D. A. *et al. C-MYC, BCL2* and *BCL6* Translocation in B-cell Non-Hodgkin Lymphoma Cases. *J. Cancer* **11**, 190–198 (2020).

104. MYC as a regulator of ribosome biogenesis and protein synthesis | Nature Reviews Cancer. https://www.nature.com/articles/nrc2819.

105. Gong, C. *et al.* Sequential inverse dysregulation of the RNA helicases DDX3X and DDX3Y facilitates MYC-driven lymphomagenesis. *Mol. Cell* **81**, 4059-4075.e11 (2021).

106. Murphy, D. J. *et al.* Distinct thresholds govern Myc's biological output in vivo. *Cancer Cell* **14**, 447–457 (2008).

107. Major, A. & Smith, S. M. DA-R-EPOCH vs R-CHOP in DLBCL: How Do We Choose? *Clin. Adv. Hematol. Oncol. HO* **19**, 698–709 (2021).

108. Ma'koseh, M. *et al.* Treatment of adult Burkitt lymphoma with the CALGB 1002 protocol: a single center experience in Jordan. *Blood Res.* **56**, 279–284 (2021).

109. Clipson, A. *et al.* The prognosis of MYC translocation positive diffuse large B-cell lymphoma depends on the second hit. *J. Pathol. Clin. Res.* **1**, 125–133 (2015).

110. Cai, Q., Medeiros, L. J., Xu, X. & Young, K. H. MYC-driven aggressive B-cell lymphomas: biology, entity, differential diagnosis and clinical management. *Oncotarget* **6**, 38591–38616 (2015).

111. Sun, P. *et al.* R-split-EPOCH plus high dose methotrexate in untreated diffuse large B cell lymphoma with MYC rearrangement or double expression of MYC and BCL-2. *J. Cancer* **12**, 2059–2064 (2021).

112. Short, N. J. *et al.* Outcomes of adults with relapsed or refractory Burkitt and high-grade B-cell leukemia/lymphoma. *Am. J. Hematol.* **92**, E114–E117 (2017).

113. Soucek, L. & Evan, G. I. THE UPS AND DOWNS OF MYC BIOLOGY. *Curr. Opin. Genet. Dev.* **20**, 91 (2010).

114. Qin, H. *et al.* RNA-binding proteins in tumor progression. *J. Hematol. Oncol.J Hematol Oncol* **13**, 90 (2020).

115. Dreyfuss, G., Swanson, M. S. & Piñol-Roma, S. Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends Biochem. Sci.* **13**, 86–91 (1988).

116. Piñol-Roma, S. & Dreyfuss, G. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* **355**, 730–732 (1992).

117. Michael, W. M., Choi, M. & Dreyfuss, G. A nuclear export signal in hnRNP A1: a signal-mediated, temperature-dependent nuclear protein export pathway. *Cell* **83**, 415–422 (1995).

118. Michael, W. M., Eder, P. S. & Dreyfuss, G. The K nuclear shuttling domain: a novel signal for nuclear import and nuclear export in the hnRNP K protein. *EMBO J.* **16**, 3587–3598 (1997).

119. Jain, N., Lin, H.-C., Morgan, C. E., Harris, M. E. & Tolbert, B. S. Rules of RNA specificity of hnRNP A1 revealed by global and quantitative analysis of its affinity distribution. *Proc. Natl. Acad. Sci.* **114**, 2206–2211 (2017).

120. Caputi, M. & Zahler, A. M. Determination of the RNA Binding Specificity of the Heterogeneous Nuclear Ribonucleoprotein (hnRNP) H/H′/F/2H9 Family *. *J. Biol. Chem.* **276**, 43850–43859 (2001).

121. Görlach, M., Burd, C. G. & Dreyfuss, G. The determinants of RNA-binding specificity of the heterogeneous nuclear ribonucleoprotein C proteins. *J. Biol. Chem.* **269**, 23074–23078 (1994).

122. Russo, A. *et al.* hnRNP H1 and intronic G runs in the splicing control of the human rpL3 gene. *Biochim. Biophys. Acta BBA - Gene Regul. Mech.* **1799**, 419–428 (2010).

123. Pararajalingam, P. *et al.* Coding and noncoding drivers of mantle cell lymphoma identified through exome and genome sequencing. *Blood* **136**, 572–584 (2020).

124. McGlincy, N. J. *et al.* Expression proteomics of UPF1 knockdown in HeLa cells reveals autoregulation of hnRNP A2/B1 mediated by alternative splicing resulting in nonsense-mediated mRNA decay. *BMC Genomics* **11**, 565 (2010).

125. Decorsière, A., Cayrel, A., Vagner, S. & Millevoi, S. Essential role for the interaction between hnRNP H/F and a G quadruplex in maintaining p53 pre-mRNA 3′-end processing and function during DNA damage. *Genes Dev.* **25**, 220–225 (2011).

126. Svitkin, Y. V. *et al.* Control of Translation and miRNA-Dependent Repression by a Novel Poly(A) Binding Protein, hnRNP-Q. *PLOS Biol.* **11**, e1001564 (2013).

127. Svitkin, Y. V., Ovchinnikov, L. P., Dreyfuss, G. & Sonenberg, N. General RNA binding proteins render translation cap dependent. *EMBO J.* **15**, 7147–7155 (1996).

128. Hussey, G. S. *et al.* Identification of an mRNP complex regulating tumorigenesis at the translational elongation step. *Mol. Cell* **41**, 419–431 (2011).

129. Gamarnik, A. V. & Andino, R. Two functional complexes formed by KH domain containing proteins with the 5’ noncoding region of poliovirus RNA. *RNA N. Y. N* **3**, 882–892 (1997).

130. Gaudreau, M.-C. *et al.* Heterogeneous Nuclear Ribonucleoprotein L is required for the survival and functional integrity of murine hematopoietic stem cells. *Sci. Rep.* **6**, 27379 (2016).

131. Zandhuis, N. D., Nicolet, B. P. & Wolkers, M. C. RNA-Binding Protein Expression Alters Upon Differentiation of Human B Cells and T Cells. *Front. Immunol.* **12**, (2021).

132. Ullrich, S. & Guigó, R. Dynamic changes in intron retention are tightly associated with regulation of splicing factors and proliferative activity during B-cell development. *Nucleic Acids Res.* **48**, 1327–1340 (2020).

133. Monzón-Casanova, E. *et al.* The RNA binding protein PTBP1 is necessary for B cell selection in germinal centers. *Nat. Immunol.* **19**, 267–278 (2018).

134. Hu, W., Begum, N. A., Mondal, S., Stanlie, A. & Honjo, T. Identification of DNA cleavage- and recombination-specific hnRNP cofactors for activation-induced cytidine deaminase. *Proc. Natl. Acad. Sci.* **112**, 5791–5796 (2015).

135. Yin, Z. *et al.* RNA-binding motifs of hnRNP K are critical for induction of antibody diversification by activation-induced cytidine deaminase. *Proc. Natl. Acad. Sci.* **117**, 11624–11635 (2020).

136. Mondal, S., Begum, N. A., Hu, W. & Honjo, T. Functional requirements of AID's higher order structures and their interaction with RNA-binding proteins. *Proc. Natl. Acad. Sci.* **113**, E1545–E1554 (2016).

137. Refaat, A. M. *et al.* HNRNPU facilitates antibody class-switch recombination through C-NHEJ promotion and R-loop suppression. *Cell Rep.* **42**, 112284 (2023).

138. Habib, T. *et al.* Myc stimulates B lymphocyte differentiation and amplifies calcium signaling. *J. Cell Biol.* **179**, 717–731 (2007).

139. Veis, D. J., Sorenson, C. M., Shutter, J. R. & Korsmeyer, S. J. Bcl-2-deficient mice demonstrate fulminant lymphoid apoptosis, polycystic kidneys, and hypopigmented hair. *Cell* **75**, 229–240 (1993).

140. Bunting, K. L. & Melnick, A. M. New effector functions and regulatory mechanisms of BCL6 in normal and malignant lymphocytes. *Curr. Opin. Immunol.* **25**, 339–346 (2013).

141. Sadri, N., Lu, J.-Y., Badura, M. L. & Schneider, R. J. AUF1 is involved in splenic follicular B cell maintenance. *BMC Immunol.* **11**, 1 (2010).

142. Lim, M.-H. *et al.* Effect of Modulation of hnRNP L Levels on the Decay of bcl-2 mRNA in MCF-7 Cells. *Korean J. Physiol. Pharmacol. Off. J. Korean Physiol. Soc. Korean Soc. Pharmacol.* **14**, 15–20 (2010).

143. Chang, X., Li, B. & Rao, A. RNA-binding protein hnRNPLL regulates mRNA splicing and stability during B-cell to plasma-cell differentiation. *Proc. Natl. Acad. Sci.* **112**, E1888–E1897 (2015).

144. Xie, Y., Luo, X., He, H., Pan, T. & He, Y. Identification of an individualized RNA binding protein-based prognostic signature for diffuse large B-cell lymphoma. *Cancer Med.* **10**, 2703–2713 (2021).

145. Shi, Y. *et al.* IL-6–Induced Stimulation of c-Myc Translation in Multiple Myeloma Cells Is Mediated by Myc Internal Ribosome Entry Site Function and the RNA-Binding Protein, hnRNP A1. *Cancer Res.* **68**, 10215–10222 (2008).

146. Gallardo, M. *et al.* Uncovering the Role of RNA-Binding Protein hnRNP K in B-Cell Lymphomas. *JNCI J. Natl. Cancer Inst.* **112**, 95–106 (2019).

147. Wagener, R. *et al.* Cryptic insertion of MYC exons 2 and 3 into the IGH locus detected by whole genome sequencing in a case of MYC-negative Burkitt lymphoma [published online ahead of print May 2019]. *Haematologica* **10.3324/haematol.2018.208140**, (2019).

148. Grande, B. M. *et al.* Genome-wide discovery of somatic coding and noncoding mutations in pediatric endemic and sporadic Burkitt lymphoma. *Blood* **133**, 1313–1324 (2019).

149. Schmitz, R. *et al.* Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature* **490**, 116–120 (2012).

150. Love, C. *et al.* The genetic landscape of mutations in Burkitt lymphoma. *Nat. Genet.* **44**, 1321–1325 (2012).

151. Kishor, A., Ge, Z. & Hogg, J. R. hnRNP L-dependent protection of normal mRNAs from NMD subverts quality control in B cell lymphoma. *EMBO J.* **38**, e99128 (2019).

152. Ye, J. *et al.* hnRNP U protein is required for normal pre-mRNA splicing and postnatal heart development and function. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E3020–E3029 (2015).

153. Romig, H., Fackelmayer, F. O., Renz, A., Ramsperger, U. & Richter, A. Characterization of SAF-A, a novel nuclear DNA binding protein from HeLa cells with high affinity for nuclear matrix/scaffold attachment DNA elements. *EMBO J.* **11**, 3431–3440 (1992).

154. Podgornaya, O. I. Nuclear organization by satellite DNA, SAF-A/hnRNPU and matrix attachment regions. *Semin. Cell Dev. Biol.* **128**, 61–68 (2022).

155. Sapir, T. & Reiner, O. HNRNPU's multi-tasking is essential for proper cortical development. *BioEssays* **45**, 2300039 (2023).

156. Thomas, N. *et al.* GENETIC SUBGROUPS INFORM ON PATHOBIOLOGY IN ADULT AND PEDIATRIC BURKITT LYMPHOMA. *Blood* blood.2022016534 (2022) doi:10.1182/blood.2022016534.

157. Balasubramanian, M. HNRNPU-Related Neurodevelopmental Disorder. in *GeneReviews*® (eds. Adam, M. P. et al.) (University of Washington, Seattle, 1993).

158. Dugger, S. A. *et al.* Neurodevelopmental deficits and cell-type-specific transcriptomic perturbations in a mouse model of HNRNPU haploinsufficiency. 2020.05.01.072512 Preprint at https://doi.org/10.1101/2020.05.01.072512 (2021).

159. Liang, Y., Fan, Y., Liu, Y. & Fan, H. HNRNPU promotes the progression of hepatocellular carcinoma by enhancing CDK2 transcription. *Exp. Cell Res.* **409**, 112898 (2021).

160. HNRNPU promotes the progression of triple-negative breast cancer via RNA transcription and alternative splicing mechanisms | Cell Death & Disease. https://www.nature.com/articles/s41419-022-05376-6.

161. Qureshi, Q. U. A., Audas, T. E., Morin, R. D. & Coyle, K. M. Emerging roles for heterogeneous ribonuclear proteins in normal and malignant B cells. *Biochem. Cell Biol.* **101**, 160–171 (2023).

162. Thomas, N. *et al.* Genetic subgroups inform on pathobiology in adult and pediatric Burkitt lymphoma. *Blood* **141**, 904–916 (2023).

163. GRIDSS: sensitive and specific genomic rearrangement detection using positional de Bruijn graph assembly - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5741059/.

164. Chen, X. *et al.* Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinforma. Oxf. Engl.* **32**, 1220–1222 (2016).

165. An integrative ENCODE resource for cancer genomics | Nature Communications. https://www.nature.com/articles/s41467-020-14743-w.

166. Krakau, S., Richard, H. & Marsico, A. PureCLIP: capturing target-specific protein–RNA interaction footprints from single-nucleotide CLIP-seq data. *Genome Biol.* **18**, 240 (2017).

167. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

168. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).

169. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

170. Korotkevich, G. *et al.* Fast gene set enrichment analysis. 060012 Preprint at https://doi.org/10.1101/060012 (2021).

171. Chen, E. Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).

172. Chen, J. L., Steele, T. W. J. & Stuckey, D. C. Metabolic reduction of resazurin; location within the cell for cytotoxicity assays. *Biotechnol. Bioeng.* **115**, 351–358 (2018).

173. Giulietti, M. *et al.* SpliceAid-F: a database of human splicing factors and their RNA-binding sites. *Nucleic Acids Res.* **41**, D125–D131 (2013).

174. Velavan, T. P. Epstein-Barr virus, malaria and endemic Burkitt lymphoma. *EBioMedicine* **39**, 13–14 (2018).

175. Crombie, J. L. & LaCasce, A. S. Epstein Barr Virus Associated B-Cell Lymphomas and Iatrogenic Lymphoproliferative Disorders. *Front. Oncol.* **9**, (2019).

176. Tang, C., Bagnara, D., Chiorazzi, N., Scharff, M. D. & MacCarthy, T. AID Overlapping and Polη Hotspots Are Key Features of Evolutionary Variation Within the Human Antibody Heavy Chain (IGHV) Genes. *Front. Immunol.* **11**, (2020).

177. Parekh, V. J., Węgrzyn, G., Arluison, V. & Sinden, R. R. Genomic Instability of G-Quadruplex Sequences in Escherichia coli: Roles of DinG, RecG, and RecQ Helicases. *Genes* **14**, 1720 (2023).

178. Exploring genomic alteration in pediatric cancer using ProteinPaint | Nature Genetics. https://www.nature.com/articles/ng.3466.

179. Lv, Y., Xiao, J., Liu, J. & Xing, F. E2F8 is a Potential Therapeutic Target for Hepatocellular Carcinoma. *J. Cancer* **8**, 1205–1213 (2017).

180. Shen, T. & Huang, S. The role of Cdc25A in the regulation of cell proliferation and apoptosis. *Anticancer Agents Med. Chem.* **12**, 631–639 (2012).

181. Liu, X. *et al.* SRSF10 stabilizes CDC25A by triggering exon 6 skipping to promote hepatocarcinogenesis. *J. Exp. Clin. Cancer Res.* **41**, 353 (2022).

182. Brakebusch, C. Rho GTPase Signaling in Health and Disease: A Complex Signaling Network. *Cells* **10**, 401 (2021).

183. Li, H., Peyrollier, K., Kilic, G. & Brakebusch, C. Rho GTPases and cancer. *BioFactors Oxf. Engl.* **40**, 226–235 (2014).

184. Voena, C. & Chiarle, R. RHO Family GTPases in the Biology of Lymphoma. *Cells* **8**, 646 (2019).

185. David, M., Petit, D. & Bertoglio, J. Cell cycle regulation of Rho signaling pathways. *Cell Cycle* **11**, 3003–3010 (2012).

186. Gauthier-Rouvière, C. *et al.* RhoG GTPase Controls a Pathway That Independently Activates Rac1 and Cdc42Hs. *Mol. Biol. Cell* **9**, 1379–1394 (1998).

187. Kukalev, A., Nord, Y., Palmberg, C., Bergman, T. & Percipalle, P. Actin and hnRNP U cooperate for productive transcription by RNA polymerase II. *Nat. Struct. Mol. Biol.* **12**, 238–244 (2005).

188. CARD8 is a negative regulator for NLRP3 inflammasome, but mutant NLRP3 in cryopyrin-associated periodic syndromes escapes the restriction | Arthritis Research & Therapy | Full Text. https://arthritis-research.biomedcentral.com/articles/10.1186/ar4483.

189. Razmara, M. *et al.* CARD-8 Protein, a New CARD Family Member That Regulates Caspase-1 Activation and Apoptosis *. *J. Biol. Chem.* **277**, 13952–13958 (2002).

190. Poole, C. J. *et al.* Targeting the MYC Oncogene in Burkitt Lymphoma through HSP90 Inhibition. *Cancers* **10**, 448 (2018).

191. Yang, Y. *et al.* c-Myc regulates the CDK1/cyclin B1 dependent-G2/M cell cycle progression by histone H4 acetylation in Raji cells. *Int. J. Mol. Med.* **41**, (2018).

192. Leung, J. Y., Ehmann, G. L., Giangrande, P. H. & Nevins, J. R. A role for Myc in facilitating transcription activation by E2F1. *Oncogene* **27**, 4172–4179 (2008).

193. Yugami, M., Kabe, Y., Yamaguchi, Y., Wada, T. & Handa, H. hnRNP-U enhances the expression of specific genes by stabilizing mRNA. *FEBS Lett.* **581**, 1–7 (2007).

194. Weidensdorfer, D. *et al.* Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. *RNA N. Y. N* **15**, 104–115 (2009).

195. Lemm, I. & Ross, J. Regulation of c-myc mRNA Decay by Translational Pausing in a Coding Region Instability Determinant. *Mol. Cell. Biol.* **22**, 3959–3969 (2002).

196. Gregory, M. A. & Hann, S. R. c-Myc Proteolysis by the Ubiquitin-Proteasome Pathway: Stabilization of c-Myc in Burkitt's Lymphoma Cells. *Mol. Cell. Biol.* **20**, 2423–2435 (2000).

197. Dani, C. *et al.* Extreme instability of myc mRNA in normal and transformed human cells. *Proc. Natl. Acad. Sci.* **81**, 7046–7050 (1984).

198. Izumi, H. & Funa, K. Telomere Function and the G-Quadruplex Formation are Regulated by hnRNP U. *Cells* **8**, 390 (2019).

199. Brancolini, C. & Iuliano, L. Proteotoxic Stress and Cell Death in Cancer Cells. *Cancers* **12**, 2385 (2020).

200. Tameire, F., Verginadis, I. I. & Koumenis, C. Cell Intrinsic and Extrinsic Activators of the Unfolded Protein Response in Cancer: Mechanisms and Targets for Therapy. *Semin. Cancer Biol.* **33**, 3–15 (2015).

201. The role of the unfolded protein response in cancer progression: From oncogenesis to chemoresistance - Madden - 2019 - Biology of the Cell - Wiley Online Library. https://onlinelibrary.wiley.com/doi/full/10.1111/boc.201800050.

202. Maestre, L. *et al.* Expression pattern of XBP1(S) in human B-cell lymphomas. *Haematologica* **94**, 419–422 (2009).

203. Chen, Y. & Brandizzi, F. IRE1: ER stress sensor and cell fate executor. *Trends Cell Biol.* **23**, 10.1016/j.tcb.2013.06.005 (2013).

204. Ma, Y. *et al.* A SMALL MOLECULE E2F INHIBITOR BLOCKS GROWTH IN A MELANOMA CULTURE MODEL. *Cancer Res.* **68**, 6292–6299 (2008).

205. Whitfield, J. R. & Soucek, L. The long journey to bring a Myc inhibitor to the clinic. *J. Cell Biol.* **220**, e202103090 (2021).

206. Sapir, T. *et al.* Heterogeneous nuclear ribonucleoprotein U (HNRNPU) safeguards the developing mouse cortex. *Nat. Commun.* **13**, 4209 (2022).

207. Wang, T. *et al.* Identification and characterization of essential genes in the human genome. *Science* **350**, 1096–1101 (2015).

208. Gallardo, M. *et al.* Aberrant hnRNP K expression: All roads lead to cancer. *Cell Cycle* **15**, 1552–1557 (2016).

209. Gallardo, M. *et al.* hnRNP K is a haploinsufficient tumor suppressor that regulates proliferation and differentiation programs in hematologic malignancies. *Cancer Cell* **28**, 486–499 (2015).

210. Single-cell RNA binding protein regulatory network analyses reveal oncogenic HNRNPK-MYC signalling pathway in cancer | Communications Biology. https://www.nature.com/articles/s42003-023-04457-2.

211. Li, M. *et al.* Heterogeneous nuclear ribonucleoprotein K promotes the progression of lung cancer by inhibiting the p53-dependent signaling pathway. *Thorac. Cancer* **13**, 1311–1321 (2022).

212. Rowland, B. D. & Bernards, R. Re-Evaluating Cell-Cycle Regulation by E2Fs. *Cell* **127**, 871–874 (2006).

213. Morgunova, E. *et al.* Structural insights into the DNA-binding specificity of E2F family transcription factors. *Nat. Commun.* **6**, 10050 (2015).

214. Kent, L. N. & Leone, G. The broken cycle: E2F dysfunction in cancer. *Nat. Rev. Cancer* **19**, 326–338 (2019).

215. Xie, D., Pei, Q., Li, J., Wan, X. & Ye, T. Emerging Role of E2F Family in Cancer Stem Cells. *Front. Oncol.* **11**, (2021).

216. Yu, H., Li, Z. & Wang, M. Expression and prognostic role of E2F transcription factors in high-grade glioma. *CNS Neurosci. Ther.* **26**, 741–753 (2020).

217. Wang, H., Wang, X., Xu, L., Zhang, J. & Cao, H. Integrated analysis of the E2F transcription factors across cancer types. *Oncol. Rep.* **43**, 1133–1146 (2020).

218. Matsumura, I., Tanaka, H. & Kanakura, Y. E2F1 and c-Myc in Cell Growth and Death. *Cell Cycle* **2**, 332–335 (2003).

219. Verdú Bou, M. *et al.* Lymphomagenic Roles of MYC, E2F and Mir-150-5p in Plasmablastic Lymphoma: Therapeutic and Prognostic Implications. *Blood* **140**, 9266–9268 (2022).

220. Winkler, R. *et al.* Targeting the MYC interaction network in B-cell lymphoma via histone deacetylase 6 inhibition. *Oncogene* **41**, 4560–4572 (2022).

221. Ruan, H. *et al.* Targeting Myc-driven stress vulnerability in mutant KRAS colorectal cancer. *Mol. Biomed.* **3**, 10 (2022).

222. Harding, H. P., Zhang, Y., Bertolotti, A., Zeng, H. & Ron, D. Perk is essential for translational regulation and cell survival during the unfolded protein response. *Mol. Cell* **5**, 897–904 (2000).

223. Cell death induced by the ER stressor thapsigargin involves death receptor 5, a non-autophagic function of MAP1LC3B, and distinct contributions from unfolded protein response components | Cell Communication and Signaling | Full Text. https://biosignaling.biomedcentral.com/articles/10.1186/s12964-019-0499-z.

224. Henley, M. J., Doyle, S. K. & Koehler, A. N. Nailing Down a Notoriously Elusive Cancer Target: Direct Inhibition of MYC by a Covalent Small Molecule. *Cell Chem. Biol.* **28**, 1–3 (2021).

225. Therapeutic Effects of an Anti-Myc Drug on Mouse Pancreatic Cancer | JNCI: Journal of the National Cancer Institute | Oxford Academic. https://academic.oup.com/jnci/article/106/12/dju320/920545.

226. Garralda, E. *et al.* Dose escalation study of OMO-103, a first in class Pan-MYC-Inhibitor in patients (pts) with advanced solid tumors. *Eur. J. Cancer* **174**, S5–S6 (2022).

227. Demma, M. J. *et al.* Omomyc Reveals New Mechanisms To Inhibit the MYC Oncogene. *Mol. Cell. Biol.* **39**, (2019).

228. Gallardo, M. *et al.* Aberrant hnRNP K expression: All roads lead to cancer. *Cell Cycle* **15**, 1552–1557 (2016).

229. Liu, T.-Y. *et al.* Muscle developmental defects in heterogeneous nuclear Ribonucleoprotein A1 knockout mice. *Open Biol.* **7**, 160303 (2017).

230. Sapir, T. *et al.* Heterogeneous nuclear ribonucleoprotein U (HNRNPU) safeguards the developing mouse cortex. *Nat. Commun.* **13**, 4209 (2022).

231. Nichols, C. A. *et al.* Loss of heterozygosity of essential genes represents a widespread class of potential cancer vulnerabilities. *Nat. Commun.* **11**, 2517 (2020).

232. Kharel, P., Becker, G., Tsvetkov, V. & Ivanov, P. Properties and biological impact of RNA G-quadruplexes: from order to turmoil and back. *Nucleic Acids Res.* **48**, 12534–12555 (2020).

233. Biffi, G., Tannahill, D., Miller, J., Howat, W. J. & Balasubramanian, S. Elevated Levels of G-Quadruplex Formation in Human Stomach and Liver Cancer Tissues. *PLOS ONE* **9**, e102711 (2014).

234. Herviou, P. *et al.* hnRNP H/F drive RNA G-quadruplex-mediated translation linked to genomic instability and therapy resistance in glioblastoma. *Nat. Commun.* **11**, 2661 (2020).

235. Sanchez-Martin, V., Soriano, M. & Garcia-Salcedo, J. A. Quadruplex Ligands in Cancer Therapy. *Cancers* **13**, 3156 (2021).

236. Neckles, C. *et al.* HNRNPH1-dependent splicing of a fusion oncogene reveals a targetable RNA G-quadruplex interaction. *RNA* **25**, 1731 (2019).

237. Zhang, J., Harvey, S. E. & Cheng, C. A high-throughput screen identifies small molecule modulators of alternative splicing by targeting RNA G-quadruplexes. *Nucleic Acids Res.* **47**, 3667–3679 (2019).

238. Havens, M. A. & Hastings, M. L. Splice-switching antisense oligonucleotides as therapeutic drugs. *Nucleic Acids Res.* **44**, 6549 (2016).

239. Scoles, D. R., Minikel, E. V. & Pulst, S. M. Antisense oligonucleotides: A primer. *Neurol. Genet.* **5**, (2019).

240. Rahman, R., Xu, W., Jin, H. & Rosbash, M. Identification of RNA-binding protein targets with HyperTRIBE. *Nat. Protoc.* **13**, 1829–1849 (2018).

241. Caeser, R. *et al.* Genetic modification of primary human B cells to model high-grade lymphoma. *Nat. Commun.* **10**, 4543 (2019).