The Role of KDM2B in Regulating Ribosome Biogenesis and mRNA Translation

A thesis submitted by

Vollter Anastas

in partial fulfillment of the requirements for the degree of

Ph.D.

in

Genetics

Tufts University

Graduate School of Biomedical Sciences

May 2024

Advisor: Philip N. Tsichlis, M.D.

Abstract

Ribosome biogenesis is a highly orchestrated process essential for all life, and its dysfunction is linked to multiple human diseases. Previous work in the lab had identified that the lysine demethylase KDM2B was critical for stemness and growth of basal-like breast cancer. Here, we provide new evidence that KDM2B regulates ribosome biogenesis and, in turn, controls translation both globally and of specific mRNAs. RNA-seq of KDM2B depleted MDA-MB-231 cells revealed significant downregulation of genes encoding ribosomal biogenesis factors (e.g., *TCOF1, NAT10, FBL, UTP14A,* and *MYC*) and components of the translation machinery (e.g., *RPS28, EIF5,* and *EEF2*). Proteomics analysis and western blotting confirmed this transcriptional repression impacted corresponding protein levels after the knockdown of KDM2B. Furthermore, polysome profiling demonstrated that KDM2B-depleted cells had reduced free subunits, monosomes, and polysomes. Similarly, the reduction of ribosome biogenesis was confirmed by tracking the synthesis of 40S and 60S ribosomal subunits in control and KDM2B-depleted MDA-MB-231 cells by editing the endogenous *RPS3* and *RPL29* loci with a Halo-tag. Acute labeling using a fluorescent Halo-ligand and subsequent tracking of Halo-tagged 40S and 60S subunits showed a reduction in nascent ribosome biogenesis in KDM2B knock-down cells. Consistent with the decrease in available ribosomes for translation, we measured a 60% reduction in global protein synthesis levels by L-homopropargylglycine labeling in shKDM2B-transduced cells. Characterization of rRNA processing by northern blotting revealed a significant defect in 18S rRNA biogenesis, particularly the processing event between the 30S to 26S rRNA pre-processing. To investigate the effect of KDM2B depletion on the translation of specific transcripts, we

pursued Ribo-seq in control and shKDM2B-transduced cells. Translation of mRNAs encoding positive regulators of the S/G2/M phases of the cell cycle displayed the most dependence on KDM2B levels, providing some insight into the observed cell cycle defects in KDM2B-depleted cells. In summary, we provide evidence that KDM2B is an upstream transcriptional regular of ribosomal biogenesis, and its loss in basal-like breast cancer cell lines leads to rRNA processing defects and decreased ribosomal abundance. These changes in ribosome biogenesis correspond to global and transcript-specific translation changes, demonstrating how KDM2B impacts multiple levels of gene regulation beyond transcription.

Dedication

To my parents, Anila and Arben, who left their own family, friends, and home to start all over in a foreign country for the chance that their son might have a brighter future.

And to my partner, Christian, whose unwavering support has made any insurmountable task seem doable.

Acknowledgments

First and foremost, I want to extend my gratitude to Dr. Philip Tsichlis. He has an infectious love of science and will never think that a pursuit is too far-reaching to chase. No matter what direction a project may go, Philip will continue to follow it into uncharted territory, trying to find answers. His office door was always open, and he thrived on people coming in to share data or discuss interesting papers they had read. This extends beyond lab and molecular biology, as he will discuss at length any historical or literary factoids he has learned with any passerby. Philip's philosophy of "chasing a question anywhere" provided a unique opportunity to grow as a scientist. I never felt that I had any limits in building my thesis work. I would have never forayed into this research without his openness and scientific stride. His scholarly nature allowed me to build a project I could truly call my own, and for that, I am incredibly thankful.

I would also like to acknowledge my committee members, Drs. Claire Moore, Karl Munger, and Phil Hinds. Their support, both inside and outside committee meetings, during my entire graduate school career, has been critical to my success. I could not have built this work without their teaching of how to think of multiple hypotheses when pursuing questions and how to appraise data to find limitations critically.

I want to thank all the lab members in the Tsichlis lab with whom I have had the pleasure of working. It was very easy to feel a sense of community among everyone, making the challenges of working in a lab much more manageable. Thank you to Elia Chavdoula, my side-by-side KDM2B investigator and bench-mate. Having someone who worked so closely on the project made it a wonderful scientific experience, allowing us to bounce ideas back and forth and drive our projects forward. Thank you to Alessandro La

Ferlita for keeping his door open to discuss how best to tackle many bioinformatic questions; he was always willing to collaborate with me on challenging problems with analysis. Thank you to Burak Soysal, a newcomer to the lab who dedicated himself to being a great scientist and was always willing to help with the project. I also want to thank my colleagues with whom I had the pleasure of meeting and working. Science is rarely done in isolation and isn't done just by interacting with those working in the same lab. Thank you to Dr. Dario Palmieri, who was always willing to hear my ideas about what I wanted to do next and offer feedback about my experimental plans. Thank you to Dr. Lara Rizzotto, who I could call at any time to explain some new data I had obtained and hear her affirmation or critique. Thank you to Dr. Michael Kearse, whose passion for translation made me more interested in the field and whose advocacy for scientific trainees extends beyond his lab.

Thank you to everyone at Tufts who always made me feel welcome and involved despite being far away during my graduate training. Everyone always brought up such insightful points during journal clubs or student talks that I always walked away from seminars thinking about something new. I want to especially thank Julia Hisey and Abe Bayer, who opened their homes when I had to travel back and were willing to spend long hours on the phone discussing my work.

Thank you to all my family members, the Albanian community, and my friends. I feel so blessed to have had such staunch support at every stage. Regardless of the difficult aspects of the journey, I have always felt that multiple people were cheering me on from behind. My mom and dad have always prioritized education above all and have made many sacrifices to ensure that was a priority. I would not be where I am today without

their tireless investment in promoting my success. I knew I could always call, and immediately, my parents would pick up and listen to anything I had to say. To my partner Christian, I do not know how I would have made it through this process without having you by my side. You always listened to everything occurring in the lab and asked me to explain the science more, regardless of whether an experiment failed or succeeded. Thank you for always being there for me through the difficulties.

Table of Contents

Title Page i

Abstract ii

Dedication iv

Acknowledgments v

Table of Contents viii

List of Figures xi

List of Copyrighted Materials Used xiii

List of Abbreviations xiv

Chapter 1: Introduction 1
1.1. Ribosomes as regulatory entities 3
1.2. Upstream Regulators of Ribosomal Biogenesis 5
1.3. Ribosome Alterations in Cancer 7
1.4. Regulation of Gene Expression by KDM2B 10
1.5. KDM2B regulation of ribosomal biogenesis 12
1.6. KDM2B role in cancer 15
1.7. Relevance of the proposed study 16
1.8. Contributions 16

Chapter 2: Materials and Methods 17
2.1. Cell Culture 17
2.1.1. Cell Growth and Media Conditions 17
2.1.2. Cell Counting 17
2.1.3. Cell Line Validation 17
2.1.4. Mycoplasma Testing of Cell Lines 18
2.2. Transient Transfection and Lentiviral Transduction 18
2.2.1. Transient Transfection 18
2.2.2. Lentiviral Production 19
2.2.3. Lentiviral Transduction 20
2.2.4. Antibiotic Selection after Lentiviral Transduction 20
2.3. Protein Extraction and Western Blotting 21
2.3.1. Protein Extraction 21
2.3.2. Western Blotting 21
2.4. RNA sequencing and analysis 23
2.4.1. RNA Extraction and Sample Quality Control 23
2.4.2. Library Preparation and Sequencing 23
2.4.3. RNA Sequencing Analysis 23
2.4.4. Gene Set Enrichment Analysis 24

2.4.5. Exon Intron Split Analysis 25
2.5. Chromatin Immunoprecipitation Sequencing and Analysis 25
2.5.1. Lysis and Immunoprecipitation 25
2.5.2. Library Preparation and Sequencing 26
2.5.3. Sequencing Alignment and Analysis 27
2.6. Assay for Transposase Accessible Chromatin Sequencing and Analysis 27
2.6.1. Nuclear pelleting, tagmentation reaction, and library preparation 27
2.6.2. Sequencing 28
2.6.3. Sequencing Alignment and Analysis 28
2.7. Breast Cancer Patient database analysis 28
2.7.1. Data Acquisition 28
2.7.2. Pre-Ranked GSEA on Correlations 29
2.8. CRISPR Cas9 Gene Editing 31
2.9. Halo-Pulse Chase Labeling and Flow Cytometry Analysis 32
2.10. Polysome Profiling 33
2.11. TMT Proteomics and Analysis 35
2.12. Northern Blotting 36
2.12.1. Probe Synthesis 36
2.12.2. RNA Extraction and Northern Blotting 37
2.13. Biorthogonal Amino Acid Nascent Peptide Labeling 39
2.13.1. Methionine Starvation, Pulse-chase, Cellular Lysis 39
2.13.2. Click-Chemistry Reaction 40
2.14. Paired RNA sequencing and Ribosome Protected Fragment sequencing and analysis 41
2.14.1. Pellet Freezing and Cell Lysis 41
2.14.2. Paired Ribosome profiling 42
2.14.3. Paired RNA-seq 43
2.14.4. Identifying a canonical transcriptome of MDA-MB-231 cells 43
2.14.5. Paired RNA-seq and Ribo-seq Analysis 44
2.14.6. Codon Analysis 45
2.14.7. Ribosomal Heterogeneity Prediction Analysis 45
2.15. Propidium Iodine staining and Flow Cytometry Analysis for Cell Cycle 46
2.16. Statistical software and analysis 46
2.17. Contributions 46

Chapter 3: Results 48
3.1. Overview 48
3.2. Rationale 51
3.3. Results 52
3.3.1. KDM2B is expressed in higher levels in basal-like breast cancer. 52
3.3.2. KDM2B knockdown impacts the expression of genes involved in ribosomal biogenesis and translation. 54
3.3.3. KDM2B demonstrated high co-occupancy with MYC at target gene promoters. 58
3.3.4. KDM2B expression correlates with the expression of genes encoding ribosomal proteins and ribosomal biogenesis factors in basal-like breast cancer 62

3.3.5. The knockdown of KDM2B decreases ribosome numbers by lowering the rate of ribosomal biogenesis. 65
3.3.6. The KDM2B knockdown disproportionally inhibits the biogenesis of the 40S ribosomal subunit. 67
3.3.7. The knockdown of KDM2B interferes with pre-ribosomal RNA processing, primarily the processing leading to the mature 18S ribosomal RNA. 69
3.3.8. The knockdown of KDM2B inhibits global mRNA translation. 76
3.3.9. The inhibition of the global translation rate in the shKDM2B cells is independent of the Integrated Stress and Ribotoxic Stress Responses 78
3.3.10. KDM2B knockdown dysregulates transcript specific translation. 79
3.3.11. KDM2B does not alter the overall ribosomal composition. 85
3.4. Acknowledgements 86
3.5. Contributions 86

Chapter 4: Discussion 88
4.1. Summary and Implication 88
4.2. Mechanism of the regulation by KDM2B 90
4.3. KDM2B as an oncogene 94
4.4. Alterations of ribosomal abundance and processing 95
4.5. Decreased levels of eIF2α 97
4.6. Ribosomes and Cell Cycle 98
4.7. How can ribosomal concentration affect transcript-specific translation? 100
4.8. Future Directions 105
4.9. Contributions 109

Chapter 5: Appendix 110

Chapter 6: Bibliography 115

List of Figures

Figure 1.1 - Mechanisms of translational regulation. 1
Figure 1.2 - Ribosomal Biogenesis and Cancer 8
Figure 1.3 - RNAi screen targetting KDM2B across multiple cell lines 14
Figure 2.1 - CRISPR Cas9 Editing of Homologous Loci 30
Figure 2.2 - Ribosomal Halo Pulse Chase 32
Figure 2.3 - Methodology of biorthogonal amino acid labeling and nascent peptide production. 40
Figure 3.1 - KDM2B is highly expressed in basal-like breast cancer. 53
Figure 3.2 - Knockdown of KDM2B globally downregulates ribosomal biogenesis..... 56
Figure 3.3 - KDM2B primarily regulates ribosomal biogenesis factors transcriptionally. 57
Figure 3.4 - KDM2B binds to active genes in the TSS along with MYC. 59
Figure 3.5 - Regulation of ribosomal biogenesis does not depend on the PRC2..... 60
Figure 3.6 - Chromatin accessibility changes after knockdown of KDM2B..... 61
Figure 3.7 - *KDM2B* RNA levels correlates with RNA levels of upstream ribosomal biogenesis factors in breast cancer patients 62
Figure 3.8 - KDM2B globally correlates with ribosomal biogenesis genes in basal-like breast cancer..... 64
Figure 3.9 - Knockdown of KDM2B reduces ribosomal abundance..... 65
Figure 3.10 - KDM2B reduces the rate of ribosomal biogenesis. 67
Figure 3.11 - KDM2B alters the proteome impacting ribosomal subunit abundance. 68
Figure 3.12 - KDM2B more potently repressed small ribosomal proteins as opposed to large..... 69
Figure 3.13 - Overview of rRNA processing steps and intermediates. 70
Figure 3.14 - ITS1 Northern blot probe after knockdown of KDM2B..... 71
Figure 3.15 - 5'ETS Northern blot probe after knockdown of KDM2B..... 72
Figure 3.16 – ITS2 Northern blot probe after knockdown of KDM2B. 74
Figure 3.17 - 3'ETS Northern blot probe after knockdown of KDM2B..... 75
Figure 3.18 - KDM2B reduces global translational rates 77
Figure 3.19 - KDM2B does not hyperactivate the ISR or RSR..... 78
Figure 3.20- KDM2B impacts translational efficiency of multiple transcripts. 80
Figure 3.21 - Knockdown of KDM2B does not alter codon usage 81
Figure 3.22 - KDM2B translationally regulated transcripts have differences in 5' UTR composition and CDS length. 83
Figure 3.23 - KDM2B alters the translation of cell-cycle related transcripts. 84
Figure 3.24 - KDM2B does not significantly alter ribosomal composition. 85
Figure 4.1 - Overall model of KDM2B regulation of ribosomal biogenesis. 89
Figure 4.2 - High Re-initiation rates are able to predict ribosomal density on short and long transcripts. 103
Figure 5.1- Confirmation of Editing of *RPS3* and *RPL29* loci 113
Figure 5.2 - P-site density of Ribo-seq across all samples. 114
Figure 5.3 Proportion of IRES-containing transcripts in different translational categories 114

List of Tables

Table 5.1 - Table of Constructs..... 110
Table 5.2 - Table of DNA Oligos 111
Table 5.3 - Table of Antibodies for Western Blots..... 112
Table 5.4 - Table of Antibodies for ChIP-seq..... 113

List of Copyrighted Materials Used

An, H., Ordureau, A., Körner, M. et al. Systematic quantitative analysis of ribosome inventory during nutrient stress. Nature 583, 303–309 (2020). https://doi.org/10.1038/s41586-020-2446-y

Pelletier, J., Thomas, G. & Volarević, S. Ribosome biogenesis in cancer: new players and therapeutic avenues. Nat Rev Cancer 18, 51–63 (2018). https://doi.org/10.1038/nrc.2017.104

Rogers DW, Böttcher MA, Traulsen A, Greig D (2017) Ribosome reinitiation can explain length-dependent translation of messenger RNA. PLoS Comput Biol 13(6): e1005592. https://doi.org/10.1371/journal.pcbi.1005592

List of Abbreviations

ATAC-seq	Assay for Transposase Accessible Chromatin Sequencing
BRCA	Breast Cancer
CDS	Coding Sequence
ChIP-seq	Chromatin Immunoprecipitation Sequencing
CHX	Cycloheximide
CPTAC	Clinical Proteomic Tumor Analysis Consortium
ETS	External Transcribed Spacer
FBS	Fetal-Bovine Serum
gDNA	Genomic DNA
GO	Gene Ontology
GSEA	Gene Set Enrichment Analysis
ISR	Integrated Stress Response
ITS	Internal Transcribed Spacer
KD	Knockdown
L-HPG	L-homopropargylglycine
LRR	Leucine-Rich Repeat
LTR	Long Terminal Repeat
mRNA	Messenger RNA
mtRNA	Mitochondrial RNA
mtrRNA	Mitochondrial Ribosomal RNA
ORF	Open reading frame
PRC	Polycomb Repressor Complex
Ribo-seq	Ribosome Footprint Sequencing
RNA-seq	RNA sequencing
RPF	Ribosome Protected Fragment
rRNA	Ribosomal RNA
RSR	Ribotoxic Stress Response
SDS	Sodium dodecyl sulfate
sgRNA	CRISPR guide RNA
STR	Short Tandem Repeat

Chapter 1: Introduction

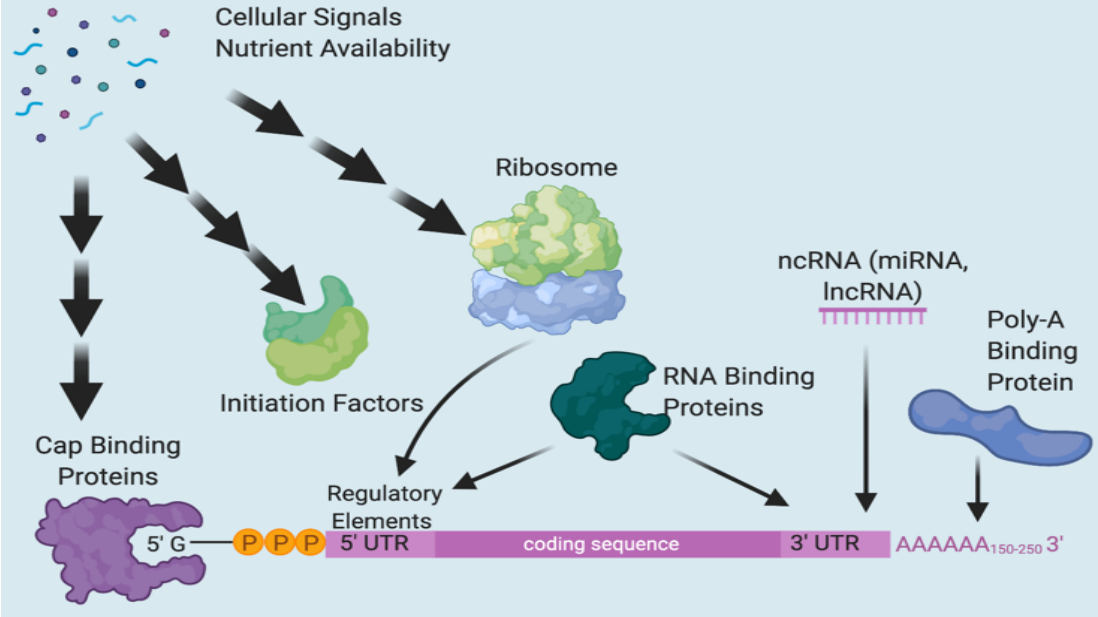


Figure 1.1 - Mechanisms of translational regulation.
Before the ribosome translates the coding sequencing of the mRNA for proper peptide production, multiple factors coordinate in the regulation of translation to ensure it is adequately controlled and responsive when cellular resources are abundant or scarce. In brief, key signaling factors that respond to growth factors and nutrient availability (in particular, the mTOR pathway (*2*)) are required to bring about proper mRNA cap binding and initiation factor recruitment (*3*). Additionally, these signals impact ribosomal biogenesis and formation and, thus, also determine the ribosomal abundance in the cell. (*4, 5*). Additionally, the Poly-A Binding protein binds to tracts of poly-adenosine with initiation factors to bring about efficient translation. (*6*). Regulatory elements in the 5' Untranslated Region (UTR), such as uORFs, can stall/halt translation and even lead to translation of short peptide products that can further stall the translation of the primary coding sequencing. Regulatory elements in the 3' UTR can be targeted by non-coding RNAs (ncRNA), including microRNA (miRNA) and long non-coding RNA (lncRNA), and impact translation (*7*). Finally, RNA-binding proteins can bind in multiple mRNA regions and impact translation positively and negatively. (*8*). The figure was made with Biorender (https://biorender.com/)

Regulation of gene expression occurs at multiple levels, from the chromatin structure of DNA to post-translational modifications on proteins to finely tune a cell's inner workings. This multi-level regulation allows each level to be responsive to any environmental changes (*9, 10*). This also allows different time courses of responses depending on the level of regulation, with chromatin structure typically being more

associated with slower, permanent cell identity changes and post-translational mechanisms like phosphorylation being more associated with transient, rapid responses. Despite this complexity, RNA transcript levels are often used as the sole markers of gene expression changes between conditions (*11*). Utilizing them as the sole markers, however, is limited. While mRNA levels and their corresponding protein levels correlate moderately well (*12*), their expressions can be discordant, highlighting the role of post-transcriptional mechanisms modulating final gene expression (*11, 13*).

Discordance between RNA and protein levels is, in part, due to regulatory mechanisms that dictate which transcripts to translate and which not to translate. This translational regulation represents an essential aspect of post-transcriptional gene expression control. It is critical in regulating multiple biological processes, including the cellular stress response (*14*), cell-cycle progression (*15, 16*), cellular differentiation (*17*), viral responses (*18*), and cellular migration (*19*). Translational regulation occurs through various types of regulators that alter the pattern of mRNAs being translated by enhancing or repressing ribosomal recruitment onto the main coding sequence of a transcript. Some well-studied translational regulatory mechanisms include RNA regulatory elements on the mRNA itself (*7, 20*), RNA-binding proteins that recognize certain elements and impact translation (*21*), non-coding RNAs (*22*), and ribosomal binding proteins that alter ribosomal function (*23*). These mechanisms represent only a fraction of the diverse arsenal for regulating gene expression at the translational level (*11*) (Figure 1.1). The existence of such complex translational regulatory machinery coincides with the fact that mRNA levels alone only partially determine final protein expression. In contrast, a much larger proportion of protein expression is better predicted by assessment of the translation

of mRNAs (*11, 12, 24-27*). However, the proper ribosomal machinery must be present in the cell before translation can even occur. As such, the primary limiting factors of translation are the production and maturation of functional ribosomes and the production of initiation, elongation, and termination factors. As such, it is no surprise that regulation of translation occurs also through controlling the core machinery involved in the process, the ribosomal subunits.

Ribosomal biogenesis is an energy-expensive process that involves multiple steps occurring simultaneously to lead to the production and maturation of functional ribosomes (*28*). The RNA component, the rRNA, is transcribed, processed/cleaved into its functional components (28S, 18S, 5.8S, and 5S). In parallel, the protein components, the ribosomal proteins, are translated in the cytoplasm from the corresponding encoding mRNAs and are then imported into the nucleolus/nucleus. The protein components are then assembled with the rRNA components to make the 40S and 60S ribosomal subunits. After assembly in the nucleolus, the subunits are exported out into the cytoplasm to complete maturation and perform their role in the cell to translate mRNAs (*29*). Many accessory factors in this process are not part of the functional ribosomal subunits but are still critical in their biogenesis. This includes non-coding RNAs, non-ribosomal accessory proteins, and import/export machinery (*30*). Given that multiple components are required for each step of biogenesis, regulatory control of this process can occur at any of these steps and impact the total level of functional ribosomes in the cell.

1.1. Ribosomes as regulatory entities

Beyond just serving as factories for protein production, ribosomes themselves can function as regulatory bodies that select which transcripts get translated or not.

Ribosomes are able to achieve this through two principal methods: 1. Qualitative changes of the ribosomal composition, and 2. Quantitative changes in ribosomal abundance. Firstly, ribosomes have a degree of functional specialization due to heterogeneity in their composition. This heterogeneity is generated by alternative modifications in the rRNA, differential composition of ribosomal protein subunits, and post-translational modifications on select subunits. These changes bring about qualitative changes to ribosomal composition that potentially alter the ribosome's affinity for specific RNAs (*31*), alter their subcellular localization (*32*), or impact their activity (*33*). Recent evidence has highlighted ribosomal heterogeneity as a central mechanism leading to alternative translational regulation (*32, 34*). This is achieved as specific mRNAs are translated more or less efficiently due to selectivity by structurally different ribosomes (*32, 34-37*). Notably, even within the same cell, ribosomes have heterogeneous compositions made of different ribosomal subunits, and these ribosomes selectively translate mRNAs (*34, 35, 38*). Secondly, the availability of ribosomes themselves can act selectively on mRNA translation. Consequently, quantitative changes in ribosomal amount impart translational changes in the cell both on a global scale and a transcript-specific scale. Selective mRNA translation has been reported to be impacted by the concentration of ribosomes and availability of ribosomes (*39-42*). Thus, a hypothesis of ribosome concentration leading to transcript specificity has been generated (*39*). However, it is challenging to exclusively manipulate ribosomal abundance, and many experiments rely on the knockdown of multiple individual ribosomal factors that demonstrate similar results across all knockdowns as proof that abundance is the likely regulatory mechanism for eliciting translation specificity (*42*). mRNAs with highly

structured untranslated regions (UTRs) or translational inhibitory sequences in their 5' UTR are predicted to be less efficient at recruiting ribosomes to their coding region in their main open reading frame (ORF). Thus, one can imagine when the ribosome number is low, the overall protein production of those specific mRNAs is inefficient even if mRNA levels remain stable. Alternatively, mRNAs that are very efficient at recruiting ribosomes can recruit multiple ribosomes to their ORF to make protein; however, when too many ribosomes are recruited, ribosomes may collide with one another, falling off without translating the transcript, paradoxically causing high levels of ribosomes present to decrease protein production of those specific mRNAs (*39, 43*). In summary, the cell can affect transcript-specific translation even by regulating ribosomal abundance. While recent literature has highlighted the ribosome as a regulatory nexus in translational regulation, the factors that establish ribosomal heterogeneity or abundance are less known. As such, it is critical to understand the upstream regulators of ribosomal biogenesis to understand how cells generate alternative ribosomal environments for translational regulation.

1.2. Upstream Regulators of Ribosomal Biogenesis

Upstream regulators such as MYC or signaling pathways such as the mTOR pathway (Figure 1.1, Figure 1.2) are some of the known regulators of ribosomal biogenesis and promote cell growth and survival (*44*). mTORC1 canonically functions by promoting the translation of mRNAs containing 5' terminal-oligo-pyrimidine tracts (5' TOP), most of which are mRNAs encoding ribosomal proteins (*45-48*). mTORC1 can also directly promote the production of precursor rRNA as it phosphorylates and activates RRN3, a transcription factor involved in RNA polymerase I activity (*49*). In line with

this, mTORC1 also promotes the phosphorylation of UBTF, which promotes the transcription of rRNA (*50*). These dual roles of the mTORC1 pathway demonstrate a concerted effort of the pathway to promote the translation of the individual ribosomal proteins and promote the production of the rRNA to ensure increased ribosomal output.

MYC acts similarly as a concerted factor to promote ribosomal biogenesis but relies on direct transcriptional activation of the genes involved (*51, 52*). Recent data has highlighted MYC as promoting pro-survival aspects in cells and that MYC appears to bind to most active genes, suggesting it has a more general role in transcription as opposed to gene-specific effects (*53, 54*). However, this idea remains controversial as more specific experiments investigating nascent transcription suggest that MYC primarily impacts transcripts involved in ribosomal biogenesis, tRNA synthesis, and *de novo* purine synthesis (*55*). Regardless, as ribosomal proteins are among the highest expressed transcripts in the cell, they are more likely to depend on high transcriptional output to increase their levels meaningfully. Thus, whether MYC is a global regulator (*53, 54*) or a more gene-specific one (*55*), it has a potent effect in promoting ribosomal biogenesis (*44*). Additionally, MYC has also been shown to promote the activity of RNA polymerase I and directly promote the expression of rRNA (*56*). It also increases the activity of RNA polymerase III to directly promote the expression of the 5S rRNA, small nucleolar RNAs involved in ribosomal biogenesis, and tRNA (*57-59*). This leads to the upregulation of all translational machinery needed. Additionally, MYC regulates other protein factors involved in rRNA processing, demonstrating that MYC also promotes efficient rRNA processing beyond just rRNA production (*60*). In summary, similar to

mTORC1, MYC appears to regulate multiple stages of ribosomal biogenesis to promote global upregulation.

The concerted impact on ribosome biogenesis is a common theme among many such upstream regulators of ribosomes. Ribosome production is costly in terms of energy (*28*). Thus, it must be tightly regulated in response to energy demands. As such, mechanisms have evolved to ensure that all necessary components are upregulated if the cell is committed to producing more ribosomes. Many factors participate in ribosomal production, including >70 ribosomal proteins, multiple ribosomal biogenesis factors, small non-coding RNAs, and the rRNA itself. Thus, it is no surprise that factors have to be regulated in a coordinated manner to ensure no energy expenditure is wasted (*61*). As such, multi-pronged effects impacting multiple layers of ribosomal production, similar to mTORC1 or MYC, are common and often co-regulated together (*62, 63*). However, these multi-level effects of MYC or mTORC1 make it challenging to specifically target or study them as genetic or pharmacological manipulation leads to pleiotropic effects across all cell types as all levels of ribosomal biogenesis would be impacted (*64, 65*). Thus, identifying other upstream factors that regulate ribosome biogenesis that have less pleiotropic effects or are more cell-type specific is critical to modulating ribosomes therapeutically (*61*).

1.3. Ribosome Alterations in Cancer

Ribosomal abundance dramatically impacts a cell's ability to grow and divide, and, expectedly, upregulation of ribosomal growth is necessary for cancer cell growth. As discussed, ribosomes not only alter overall protein output but also alter translational fidelity to specific transcripts (*29, 66*). As expected, many oncogene activations, such as

MYC or activations of the mTOR1 pathway, promote ribosomal biogenesis as part of their pro-oncogenic function (*29, 67-69*) (Figure 1.2).

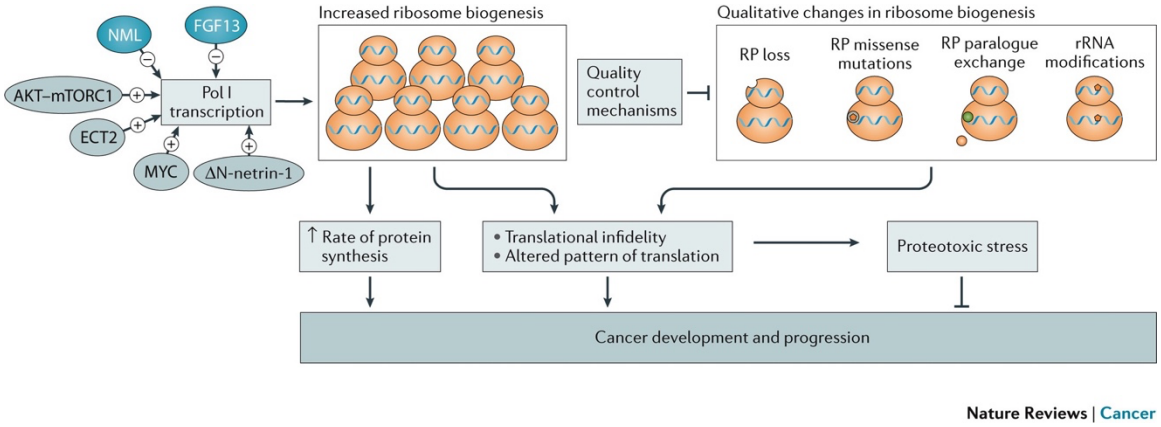


Figure 1.2 - Ribosomal Biogenesis and Cancer
Reprinted with permission from Pelletier, J., Thomas, G. & Volarević, S. Ribosome biogenesis in cancer: new players and therapeutic avenues. Nat Rev Cancer 18, 51–63 (2018). https://doi.org/10.1038/nrc.2017.104

Oncogene activating mutations are not the only things that can alter ribosomal biogenesis; tumor suppressor gene mutations also impact ribosomal biogenesis. *TP53* is the most commonly mutated gene in cancer (*70*) and functions not only as a part of cell-cycle control but also as a part of ribosomal quality control. Normally, aberrations in ribosomal production lead to activation of the nuclear stress response pathway, which in turn activates p53 to halt and control the process (*71*). In *TP53*-mutated cancer cells, the nucleolar stress responses are blunted, opening the way for aberrations in ribosomal production to occur as the quality control is gone. In fact, the loss of ribosomal proteins in Diamond-Blackfan Anemia (DBA) induces nucleolar stress activation due to p53 activation from imbalances in ribosomal proteins, and this leads to apoptosis (*72, 73*). Initially, ribosomal haploinsufficiency leads to cellular death through activation of p53; however, eventually, ribosomal imbalances confer a selective advantage for inactivating mutations or loss of *TP53* in order to bypass the nucleolar stress presented by ribosomal

dysregulation (*74*). Notably, ribosomal haploinsufficiency after loss of *RPS6* causes translational changes in the cell. These changes, interestingly, depend on the *TP53* status in the cell. Specifically, haploinsufficiency of *RPS6* in mice leads to changes in limb development and alterations in the translational efficiency of specific transcripts. However, this is ameliorated when *TP53* is not present in the cells (*75*). This highlights that some translational changes resulting as a primary cause of ribosomal protein expression changes are dependent on activation by p53. Additionally, this demonstrates the strong selection pressures that the nucleolar stress response induces on cells with altered ribosomes to drive *TP53* mutations.

All in all, this demonstrates that most commonly altered oncogenes (*RAS, MYC*) and tumor suppressor genes (*TP53*) regulate ribosomal production and translation. Cancer cells often promote dysregulated ribosomal production through both increased production and loss of quality control mechanisms. Interestingly, loss of quality control offers a selective advantage to avoid the consequences of dysregulated biogenesis, as evidenced in DBA (*72-74*). This is suggestive of why alterations like *MYC* copy-number/expression increases often co-occur with *TP53* inactivating mutations (*76, 77*).

Upregulation of even single ribosomal factors in cancer cells confers a selective advantage in cells to promote tumor growth or metastasis (*78*). For example, overexpression of *RPLP1* (*79*), *RPL36A* (*80*), or *RPS6* (*81*) in cancer cells promotes cell-cycle progression, diminishes replicative senescence, and promotes stemness, all benefiting the cancer cell tumor growth. Further providing evidence for this, recently, a CRISPR activation screen in MDA-MB-231 cell lines was performed to characterize which genes, if activated, would enhance the metastatic rate of the tumor cells.

Intriguingly, the top hits from the screen came from ribosomal proteins, particularly *RPL15* (*82*). This further lends credence to the fact that traditional oncogenes converge to act on the upregulation of ribosomal factors. Even in these individual factor studies, it is likely that these individual factors promote general ribosomal and protein production. This general upregulation of ribosome biogenesis leads to enhanced tumorigenesis and metastasis, as opposed to the factors acting as solitary effectors themselves. In fact, a comparison of *RPL15* overexpressed cell lines versus control demonstrated that *RPL15* overexpression enhances the translation of ribosomal proteins and other growth factors. This shows that the upregulation of *RPL15* leads to an overall higher ribosomal signature, which likely provides a selective advantage for the cancer cell line (*82*). This is also reflected clinically with analysis of patient breast cancer revealing that high overall expression of ribosomal subunit proteins is associated with poor prognosis (*82*).

To conclude, the most commonly altered tumor suppressors and oncogenes act in regulating ribosomal biogenesis. Tumors have selective advantages imparted when promoting overall ribosomal biogenesis and inhibiting quality control mechanisms. This is evidenced by the fact that high ribosomal abundance correlates with poor prognosis in cancer (*82*). These changes in ribosomal abundance promote selective translation of key transcripts that are critical for tumor growth.

1.4. Regulation of Gene Expression by KDM2B

KDM2B is a histone H3K36me1/2, H3K4me3, and H3K79me2/3 demethylase that is important in stem cell maintenance (*83, 84*), cancer cell biology (*85-88*), and cellular senescence (*89, 90*). It has been described primarily as a transcriptional regulator of protein-encoding genes (*87, 91*). Independent of its demethylase activity, KDM2B,

through its CXXC domain, is able to bind to unmethylated CpG islands and recruit transcriptional complexes such as the variant polycomb repressor complex (PRC). The best described cellular action of KDM2B is through the PRC1 complex (*86, 92-97*). The PRC1 complex is a multi-protein complex that is recruited to target genomic loci and alter their gene expression. Traditionally, it is thought to facilitate the repression of genes at target loci; however, new literature has revealed that it also participates in gene activation by recruiting active chromatin remodelers or activating transcription factors to specific sites (*98-100*). The PRC1 is defined by RING1B as the common factor that forms a heterodimeric complex with one of the PCGF proteins to act as an E3-ubiquitin ligase targeting and monoubiquitylating H2A K119. There are 6 PCGF proteins (PCGF1-6), and thus, there are 6 unique PRC1 complexes that function in gene regulation numbered PRC1.1-PRC1.6. Canonically, the PRC1.2 and PRC1.4 complexes identify H3K27me3 marks placed by the PRC2 complex and then target those loci through the H3K27me3 mark to and lead to H2AK119 ubiquitylation to further repress gene expression and confer a closed chromatin state. Thus, in the canonical PRC pathway, the PRC2 complex, which consists of H3K27me3 methyltransferases EZH1 and EZH2, deposits the repressive marks first and then the PRC1.2 or PRC.1.4 commits to further repression through ubiquitination of H2AK119. However, the other PRC complexes (PRC1.1, 1.3, 1.5, and 1.6) are able to target genomic loci independently of H3K27me3 marks and thus are labeled as non-canonical (*101*). Most of these non-canonical PRC1 complexes target their respective loci through factors that directly target DNA or factors that recognize histone marks other than H3K27me3. The PRC1.1 complex, of which KDM2B is the DNA targeting component, recognizes unmethylated CpG islands and

brings the complex to those loci. These loci tend to be enriched for genes involved in cell survival and metabolism and are rarely repressed (*102*), hinting a role of the non-canonical PRC1.1 in active transcription as opposed to transcriptional repression as in the canonical PRC1.2 and PRC1.4. In line with this, recent evidence also suggests that KDM2B is able to bind to CpG islands and protect them from hypermethylation by DNA methyltransferases (*103*) and that this function is critical for embryonic development. Interestingly, mutations in KDM2B, particularly those in its DNA-binding domain, induce a hypermethylation signature in CpG islands and are associated with a neurodevelopmental disorder with variable penetrance, further highlighting its role in active transcription and not just transcriptional repression (*104*).

Overall, KDM2B is a complex multidomain protein and has been challenging to investigate mechanistically. Given its multiple reported gene regulatory mechanisms, it may be that KDM2B's function differs across different cell types depending on the context and availability of other co-factors. For example, in a cell type where other PRC1.1 members are more highly expressed, KDM2B would be predicted to function mainly through PRC1.1. While in cell types where the other PRC1.1 components are rate limiting, KDM2B might act through solitary regulatory functions such as maintaining CpG island hypomethylation or demethylating histone marks.

1.5. KDM2B regulation of ribosomal biogenesis

Recently, methyltransferases and demethylases similar to KDM2B have been shown to play important roles in ribosomal and translation regulation (*105-109*). These factors not only act as upstream epigenetic regulators of ribosomal biogenesis (*110*) but are also able to directly target ribosomal proteins or ribosomal processing factors (*105*).

As histone methyltransferases and demethylases are critical in cancer development, with small-molecule inhibitors being developed as therapeutics, understanding these translational and ribosomal functions will be essential in evaluating treatment potentials (*111*).

Few reports exist detailing the role of KDM2B in ribosomal biogenesis. However, some groups do report that KDM2B impacts ribosomes by inhibiting ribosomal biogenesis (*112*) and inhibiting cell growth (*113, 114*). These two groups claim that KDM2B regulates rDNA genomic loci through epigenetic regulation and renders the loci more repressive. This is based on an early report that KDM2B is nucleolar (*112*). However, more recent evidence casts doubt on their initially proposed mechanism of inhibition (*93*). Firstly, the initial work by Frescas *et al.* utilized an in-house antibody of KDM2B generated from an immunogen containing an arginine-lysine rich portion of KDM2B which they predicted to be the nucleolar localization signal of KDM2B. The antibody led to an intense nucleolar signal in HELA cells, leading them to erroneously label KDM2B as a nucleolar protein that represses rRNA transcription (*112*). However, this arginine-lysine region is common among all nucleolar proteins, and thus, the antibody likely would have higher binding to the more highly expressed nucleolar proteins as opposed to KDM2B. More recently, validated antibodies against more specific domains/regions of KDM2B, in addition to mass spectrometry experiments characterizing its binding partners, find that KDM2B is nuclear and has no nucleolar signal (*83, 93, 112, 115*). Additionally, reports that loss of KDM2B leads to increased cell growth (*113, 114*) through the same previously reported mechanism (*112*) stands in stark contrast to the body of literature characterizing it as oncogenic and even

overexpressed in a variety of cancer subtypes (*85, 97, 102, 116-120*). Thus, it is unlikely that the primary function of KDM2B is to downregulate ribosomal biogenesis, which would be uncharacteristic for a more oncogenic protein. Overall, as one might expect, it is rare for the loss of a gene product in a cancer cell line to lead to a selective growth advantage as, over time, gene products that are deleterious to a cell would be lost, leaving only ones that provide a growth advantage. As such, one can predict that if loss of KDM2B conferred a growth advantage over time, it would be expressed at lower levels in cancer cell lines. This has been confirmed in the cancer cell dependency map, where most downregulated genes by CRISPR knockout (KO) or RNAi knockdown (KD) are neutral or deleterious (*121, 122*). Specifically, for KDM2B, RNAi inhibition is mainly negative (Figure 1.3). In the studied cell line, MDA-MB-231, KD of KDM2B results in a deleterious effect (Figure 1.3), calling into question their original findings (*113, 114*).

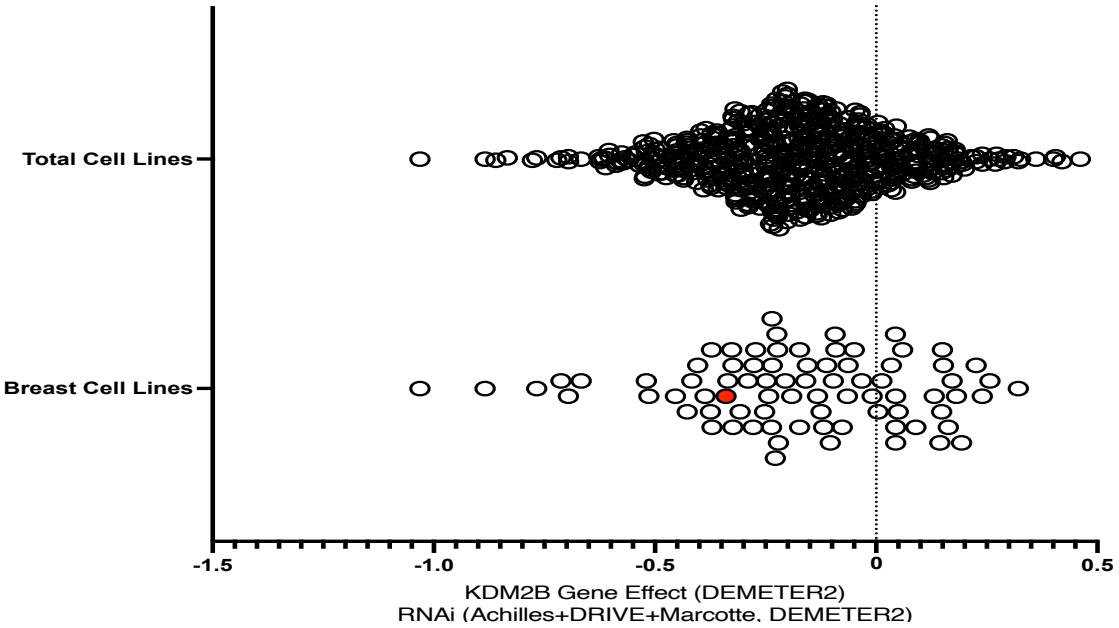


Figure 1.3 - RNAi screen targetting KDM2B across multiple cell lines
Calculated DEMETER2 score of gene effect after KDM2B knockdown with 0 being neutral, negative effect being deleterious, and positive being beneficial. Red dot signifies the MDA-MB-231 cell line. Data was obtained from the Cell Dependency map (https://depmap.org/portal/) and plotted using Graphpad Prism.

1.6. KDM2B role in cancer

KDM2B is overexpressed and displays oncogenic activity in a variety of cancers (*117*). This is attributed to its epigenetic effects modulating cell differentiation, cell metabolism, and cell proliferation. Its mechanism of promoting cancer varies with reports. KDM2B has been shown to impact cell cycle regulators (*86, 119, 123*), stemness factors (*124*), and miRNAs (*120, 125, 126*). A recent study in T-cell leukemia revealed that loss of KDM2B can enhance the initiation of cancer, but once cancer has already developed, then KDM2B is critical in promoting growth (*127*). This demonstrates that KDM2B might have a pro-oncogenic role in supporting tumor maintenance as opposed to being an initiating event in cancer. Notably, KDM2B is overexpressed in triple-negative breast cancer (TNBC) relative to other breast cancer subtypes and contributes to pathogenesis/relapse through the maintenance of a cancer stem cell like fate (*85, 118*). TNBC remains one of the most difficult to treat, given the lack of precise molecular targets. The most common treatment for triple-negative breast cancer, beyond surgical resection, is chemotherapeutic agents, all of which are less tolerated and have more side effects than targeted therapies that are used in other subtypes of breast cancer (*128*). Despite advancements in breast cancer treatment, TNBCs have the lowest 5-year survival rate as compared to other breast cancer subtypes. As KDM2B appears to be selectively overexpressed in this subtype, characterizing the global transcriptional regulation dependent on KDM2B will be critical to understanding how it impacts oncogenesis in TNBC. Importantly, as most oncogenes impact ribosomes, it will be imperative to characterize the role of KDM2B in ribosomal regulation and, in turn, translational regulation.

1.7. Relevance of the proposed study

Translational control represents a key aspect of cancer development (*129, 130*). Interestingly, the heterogeneity of tumors in regard to response to therapy is in large part due to differential translational regulation (*131, 132*). This further highlights that translational control is a crucial aspect of tumorigenesis. While some aspects of what dictates translational regulation have been described, other players, such as epigenetic regulators like KDM2B, are still poorly understood. Characterizing how KDM2B regulates translation will offer novel insights into general protein translation and how KDM2B may be regulating gene expression in cancer.

1.8. Contributions

Vollter Anastas wrote the introduction. Philip N. Tsichlis edited the introduction.

Chapter 2: Materials and Methods

2.1. Cell Culture

2.1.1. Cell Growth and Media Conditions

MDA-MB-231 cells (ATCC Cat: HTB-26) and HEK293T-LentiX cells (Takara Cat: 632180) were cultured in High Glucose Dulbecco's modified Eagle medium (Sigma Cat: D5796) supplemented with 10% fetal bovine serum (FBS) (Corning Cat: 35-010-CV), penicillin 100u/mL and streptomycin 0.1mg/mL (Sigma Cat: P4333), sodium pyruvate 1mM (Gibco Cat: 11360070), non-essential amino acids (Sigma Cat: M7145), and Plasmocin™ 5 μg/mL (Invivogen Cat: ant-mpp). Cells were subcultured using trypsin (Sigma Cat: T4049) every 2-3 days according depending on confluence avoiding confluency >90%.

2.1.2. Cell Counting

Cell counting was performed using a TC-20 Cell Counter™ (BioRad Cat: 1450102) and trypan blue staining (Sigma Cat: T8154) to determine the percentage of live or dead cells. Briefly, cells were trypsinized, centrifuged, and resuspended in media. Equal volumes of resuspended cells and trypan blue stain were mixed and then pipetted into a cell counter slide (BioRad Cat: 1450011) and then, placed inside the TC-20 Cell Counter™ to determine the cell count.

2.1.3. Cell Line Validation

Genomic DNA (gDNA) was extracted from cells utilizing the DNEasy Blood and Tissue Kit (Qiagen Cat: 69504) following the manufacturer's protocol. gDNA was then submitted to The Ohio State University Comprehensive Cancer Center Genomics Shared Resource (Columbus, OH) for Short Tandem Repeat (STR) profiling (*133*). Cell line STR

profiles were compared to the known standards, and a cut-off of an 80% match rate was used to substantiate that a cell line was validated (*134*). All cell lines were matched at a 95% rate or higher. Cell lines were retested yearly to maintain proper validation.

2.1.4. Mycoplasma Testing of Cell Lines

Mycoplasma testing was performed using a PCR-based assay using the Mycoplasma Detection Kit (ABM Cat: G238). Cells were grown to 50%-70% confluence and left without changing solution for at least 48 hours to ensure the sensitivity of the assay. A sample of the supernatant was obtained from the growth media of the cell as the input for the sample. Mycoplasma testing was performed regularly whenever thawing a new batch of cell lines, prior to preparing parallel sequencing libraries (e.g., RNA-seq, ChIP-seq, Ribo-seq, etc.), and prior to performing laborious/time-consuming experiments (e.g., Polysome profiling). Additionally, cells were also routinely assessed at least twice a year to ensure no contamination had occurred.

2.2. Transient Transfection and Lentiviral Transduction

2.2.1. Transient Transfection

Transfection of plasmids was performed using Lipofectamine 3000™ following the manufacturer's protocol (Thermo Scientific Cat: L3000015). If transfection efficiency was low for any particular plasmid in a cell line, transfection by electroporation was utilized to increase transfection efficiency. SE Cell Line 4D X Kit (Lonza Cat: V4XC-1012) was used for the reagents, and the electroporation was performed in cuvette on a 4D Nucleofecter™ apparatus (Lonza Cat: AAF-1003X) on the pre-determined settings of the cell line as listed by the manufacturer. Constructs used for gene manipulation in cells are listed in Table 5.1.

2.2.2. Lentiviral Production

Lenti-X cells were seeded the day before transfection into a 10cm dish so that they would be 40-60% confluent on the day of transfection. Lipofectamine 3000™ was used for the transfection. Briefly, 2nd generation packaging plasmids (7.2µg of pPax2 (https://www.addgene.org/12260/) and 4.8µg of pMD2.G (https://www.addgene.org/12259/)) were pipetted into a 500µL tube of OptiMEM™ (Thermo Fisher Scientific Cat: 31985070). Additionally, 12µg of the target plasmid, containing the Long Terminal Repeats (LTRs), for lentiviral integration was also added into the tube, after which 48µL of P3000 was added. A separate tube was set up with 500µL tube of OptiMEM™ and 24µL of Lipofectamine™. The tube containing the plasmids and P3000 was then slowly pipetted into the tube containing Lipofectamine™ and incubated at room temperature for 15 minutes. Finally, the entire 1mL mix of Lipofectamine™, plasmid DNA, and P3000 was pipetting over LentiX cells, spreading drops throughout the whole plate.

12-16 hours after transfection, the media was removed from the plate and replaced with fresh media. The plate was left to continue growing and producing the virus for an additional 24 hours. At that point, the viral media was collected and filtered through a 0.45µm filter (Corning Cat: 431220), and fresh media was replaced onto the cells for another 24-hour outgrowth. At the second 24-hour outgrowth, the viral media was collected and filtered through a 0.45µm filter, and bleach was used on the cell plate for 30 minutes to sterilize any remaining virus. Afterward, the plate was properly disposed of in biohazard containers. The viral media collected was either frozen at -80°C

for storage and future use or used the same day on a plate containing the target cell line for the lentiviral construct.

2.2.3. Lentiviral Transduction

Target cell lines were plated to be at 40-50% confluence on the day of viral transduction. Viral media was diluted depending on viral transduction efficiency and supplemented with hexadimethrine bromide (polybrene) (Sigma Cat: 107689) to a final concentration of 8µg/mL to improve transduction efficiency. The diluted viral media containing polybrene was then used to replace the fresh media on the target cell line and left to infect for 24-48 hours.

2.2.4. Antibiotic Selection after Lentiviral Transduction

After viral transduction, cells were split to a low confluence (20-30% confluence) in order to ensure efficient antibiotic selection. A non-transduced cell culture plate of the same cell-line batch was also split in tandem and used to determine when selection had been completed. In MDA-MB-231 cells, 10µg/mL of puromycin (Thermo Fisher Scientific Cat: A1113803) was used for 2-3 days, changing solution every 2 days until the non-transduced cell culture plate was completely dead and 10µg/mL of blastocidin (Thermo Fisher Scientific Cat: A1113903) was used for 5-7 days changing solution every 2 days until the non-transduced cell culture plate was completely dead. After selection, the media was replaced with fresh media without selective antibiotics, and cells were counted and plated for experiments.

2.3. Protein Extraction and Western Blotting

2.3.1. Protein Extraction

Cells were lysed by scrapping with a cell lifter (Corning Cat: 3008) in Radioimmunoprecipitation Assay (RIPA) buffer (25mM Tris-HCl pH 7.6, 150 mM sodium chloride, 1% Nonidet P-40, 0.1% sodium deoxycholate, 0.1% SDS, 5mM EDTA) with Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific Cat: 78440). The protein lysate was then placed in a tube, vortexed, placed on ice, and left shaking for 30 minutes at 4°C to complete the lysis. The protein lysate was then centrifuged at 15,000xg for 15 minutes at 4°C to pellet cellular debris, and the supernatant was collected in a fresh microcentrifuge tube. Protein concentration was determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, 23225).

2.3.2. Western Blotting

Protein lysate was measured and diluted 1:4 in 4X Laemmli Sample Buffer (BioRad Cat: 1610747) with 5mM dithiothreitol (DTT) as the final concentration. The samples were then boiled at 95°C for 10 minutes to complete the protein denaturation.

Protein lysates along with protein ladder (BioRad Cat: 1610374) were loaded onto separate lanes in a 4-15% gradient Mini-Protean TGX™ gel (BioRad Cat: 4561083/4561086). For phospho-gels, SuperSep™ Phos-tag™ Precast Gels (Wako Chemicals Cat: AAL-107) were used according to the manufacturer's protocol. Gels were run in an electrophoresis chamber (BioRad Cat: 1658030) using SDS running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3) for 1-2 hours at a constant 100-140V. Then, the gel was transferred to a PVDF membrane (Thermo Fisher Scientific Cat: 88520) using the transfer apparatus of the electrophoresis chamber (BioRad Cat:

1658030) using Transfer buffer (25 mM Tris, 192 mM glycine, 20% v/v methanol). The transfer was performed in 2-3 hours at a constant 400mA at 4°C with a cold pack in the electrophoresis chamber to maintain cooler temperatures. Larger proteins (>200kDa) were transferred overnight at a constant 150mA at 4°C.

After the transfer, the membrane was obtained and stained with the Revert™ 700 Total Protein stain (Licor Biosciences Cat: 296-11020) following the manufacturer's protocol and imaged on the 700nm channel on an LI-COR Odyssey® Fc Imaging System (Licor Biosciences Cat: Discontinued). Afterward, the membrane was washed in Tris-Buffered-Saline with Tween (TBST) Buffer (20mM Tris-HCl, 150mM NaCl, 0.1% v/v Tween® 20). The membrane was blocked in 5% w/v BioRad Blocking Buffer (BioRad Cat:1706404), a milk-based blocking buffer, in TBST or in 5% w/v bovine serum albumin (BSA) (Fisher Scientific Cat: BP1600-100) in TBST for 1 hour. Afterward, the membrane was cut based on the protein ladder sizes to detect the target protein and incubated with the primary antibody in 5% w/v BSA overnight at 4°C. Following this, the primary antibody was removed, and the membrane was washed 4-5 times in TBST for 5 minutes each time. Primary antibodies used for western blotting are listed in Table 5.3. Then, the membrane was incubated with horseradish peroxidase (HRP)-conjugated secondary antibodies recognizing the species (mouse or rabbit) of the primary antibody in 5% w/v BSA for one hour at room temperature. The membrane was lastly washed 4-5 times in TBST for 5 minutes each time and incubated in either Pierce™ ECL2 Blotting Substrate (Thermo Fisher Scientific Cat: 80196) or SuperSignal™ West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific Cat: 34095) following manufacturer’s protocol. The membrane was finally imaged using the 700nm channel for

the protein ladder and the chemiluminescence for the HRP signal channel on a LI-COR Odyssey® Fc Imaging System.

2.4. RNA sequencing and analysis

2.4.1. RNA Extraction and Sample Quality Control

RNA from cells that were 50-70% confluent containing either a non-hairpin control shRNA or a shRNA targeting *KDM2B* lentiviral integration was collected using the Total RNA Purification Plus Kit (Norgen Biotek Corp. Cat: 48300). The total RNA extracted was run on an Agilent Tapestation utilizing RNA Screentape™ (Agilent Technologies Cat: 5067-5576). RNAs with an RNA Integrity Number (RIN) > 9.5 were submitted for library preparation.

2.4.2. Library Preparation and Sequencing

RNA library preparation and sequencing were performed by the Brigham Young University DNA Sequencing Center (Provo, UT) as contracted out by Genohub Inc. (Austin, TX). Library preparation was performed using a KAPA RNA HyperPrep kit with RiboErase (Roche Cat: KK8560) and underwent paired-end sequencing (2x150bp), obtaining an average of approximately 35 million reads per sample on an Illumina HiSeq Rapid V2 Chip.

2.4.3. RNA Sequencing Analysis

Reads were first inspected for quality, and library preparation adapters were trimmed and removed from the reads using TrimGalore (https://github.com/FelixKrueger/TrimGalore). TrimGalore uses FASTQC (*135, 136*) to visualize sequencing quality and CutAdapt (*137*) to trim adaptor sequences. MultiQC was used to visualize all sequencing quality across all samples (*136*).

RNAdetector was used for the alignment to genome, gene quantification of reads, and differential analysis comparing control and knockdown of *KDM2B* (*138*). Reads were aligned to the *Homo sapiens* genome (GRCh38) using STAR (*139*) and aligned reads were then quantified using featureCounts (*140*). Differential expression analysis was performed in RNAdetector utilizing DESeq2 (*141*), edgeR (*142*), and LIMMA (*143*). The results of all 3 differential expression analysis tools were combined to generate a meta-p-value of results that was congruent across all 3 calculators (DESeq2, edgeR, and LIMMA) using metaSeqR (*144*).

2.4.4. Gene Set Enrichment Analysis

Gene Set Enrichment Analysis (v4.3.2) (GSEA) (*145, 146*) was performed using the reads per million (RPM) table of the 3 replicates of control and 3 replicates of the KD of *KDM2B*. The geometric mean of the 3 control and 3 shKDM2B samples was calculated for each gene loci, and geometric means < 1 were filtered out to remove poorly expressed genes as recommended by GSEA. This resulted in 13250 gene loci remaining, which corresponded to 12538 protein-encoding genes that were used for the analysis. Enrichment analysis was performed spanning the Gene Ontology (GO) Biological Processes (*147*) that had gene lists with >15 genes in each category and with <500 genes in each category as the default setting of GSEA. 10000 permutations were run for each gene set, and an enrichment score was calculated for each permutation. This generated a random distribution of enrichment scores. The actual enrichment score of our data was then compared to this distribution to determine statistical significance. Briefly, statistics were chosen based on the probability that the input ranked list had a better absolute

enrichment score than the 10000 permutations and adjusted for multiple hypothesis testing based on the total number of gene sets considered.

2.4.5. Exon Intron Split Analysis

Quantification of the fold change of each gene based on counts on the exon as compared to the counts on the intron was performed by Exon Intron Split Analysis (EISA) using eisaR (https://github.com/fmicompbio/eisaR) (*148, 149*). Briefly, genome annotations files were generated by the eisaR package from the human genome assembly (GRCh38) that had labeled exons and introns across all protein-encoding gene bodies. The RNAseq reads were aligned to the genome with introns and exons labeled using HISAT2 (*150*) and reads mapping to the exons were counted along with reads mapping to the entire gene body. The reads mapping to solely to the introns were determined by taking the reads mapping to the entire gene body and subtracting the reads mapping to the exons for each protein encoding gene. Finally, the fold change between control and shKDM2B was calculated independently from the exon count file and the intron count file, and the fold changes based on exons and those based on introns were compared to determine if the change in the intron was statistically different from the change in the exon.

2.5. Chromatin Immunoprecipitation Sequencing and Analysis

2.5.1. Lysis and Immunoprecipitation

Samples for Chromatin immunoprecipitation sequencing (ChIP-seq) experiments were prepared using the SimpleChIP Plus Enzymatic chromatin IP kit (Cell Signaling Technology Cat: 9005) according to the manufacturer's protocol. 4 million cells per immunoprecipitation of MDA-MB-231 cells were cross-linked with 1%

formaldehyde (Sigma-Aldrich Cat: F8775) for 10 min at room temperature, for the KDM2B ChIP dual cross linking was performed with an ethylene glycol-bis (succinimidyl succinate) (EGS) (Thermo Fisher Scientific Cat: 21565), in ice-cold PBS to a final concentration of 1.5 mM for 20 min at room temperature in addition to the 10min 1% formaldehyde fixation. Following chromatin digestion with 0.5 µl micrococcal nuclease per immunoprecipitation preparation, the nuclei were completely lysed after sonication with 3 sets of 20s-on 30s-off pulses with 20% amplitude. 2% input DNA that did not undergo immunoprecipitation was used as a control.

2.5.2. Library Preparation and Sequencing

ChIP-Seq libraries were prepared using the DNA Library Prep Kit for Illumina (Cell Signaling Technology Cat: 56795) in combination with the Multiplex Oligos for Illumina (Dual Index Primers) (Cell Signaling Technology Cat: 47538) according to the manufacturer's protocol. The starting input material of the libraries for the PCR amplification and barcoding of the samples was: 5 ng for KDM2B, 1 ng for c-MYC, 50 ng for the histone marks (H3K4me3, H3K27me3, and H2AK119ub) of the ChIP-enriched DNA. Finally, 50 ng of the input DNA as the control. Antibody concentrations and amounts are listed in Table 5.4.

Sequencing was performed by Fulgent Genetics (Temple City, CA) as contracted out by Genohub Inc. (Austin, TX). Sample libraries were pooled and underwent paired-end sequencing (2x150bp), obtaining an average of approximately 60 million reads per sample on an Illumina NovaSeq S4 Chip.

2.5.3. Sequencing Alignment and Analysis

Reads were first inspected for quality, and library preparation adapters were trimmed and removed using CutAdapt (*137*). Reads were then inspected for quality using FASTQC (*135, 136*) and MultiQC (*136*) to visualize all sequencing quality across all samples. Reads were aligned to the *Homo sapiens* genome (GRCh38) using bowtie2 (*151*) and peak calling was performed using MACS2 (*152*). Peaks were then annotated based on their location relative to genomic loci (*153*) and distribution of peaks relative to protein encoding genes were visualized using ChIPseeker (*154*).Visualization was done with fluff to investigate global patterns of binding from ChIP-seq (*155*).

2.6. Assay for Transposase Accessible Chromatin Sequencing and Analysis

2.6.1. Nuclear pelleting, tagmentation reaction, and library preparation

MDA-MB-231 cells were counted to approximately 90,000 cells for each sample, washed once with 1xPBS, and processed using the ATACseq kit (Active Motif, Cat: 53105). Briefly, a lysis buffer was used to break down the plasma membrane, leaving behind nuclei. The nuclei were centrifuged down and then incubated in transposomes to allow for fragmentation and adapter ligation in accessible open chromatin regions (tagmentation). Tagmented DNA was then purified, and PCR was performed, amplifying the tagmented regions and barcoding the individual samples. The resulting libraries were purified using a magnetic bead to isolate DNA and were analyzed on an Agilent Bioanalyzer High Sensitive DNA Chip (Agilent Technologies Cat: 5067-4626) to assess amplicon sizes and DNA concentration.

2.6.2. Sequencing

Sequencing was performed by Fulgent Genetics (Temple City, CA) as contracted out by Genohub Inc. (Austin, TX). Sample libraries were pooled and underwent paired-end sequencing (2x150bp), obtaining an average of approximately 250 million reads per sample on an Illumina NovaSeq S4 Chip.

2.6.3. Sequencing Alignment and Analysis

Reads were first inspected for quality, and library preparation adapters were trimmed and removed using CutAdapt (*137*). Reads were then inspected for quality using FASTQC (*135, 136*). MultiQC was used to visualize all sequencing quality across all samples (*136*). Reads were aligned to the *Homo sapiens* genome (GRCh38) using bowtie2 (*151*) and peak calling was performed using MACS2 (*152*). Peaks were then annotated based on their location relative to genomic loci (*153*) and distribution of peaks relative to protein encoding genes were visualized using ChIPseeker(*154*). Differential peak analysis between control and shKDM2B samples was performed using DiffBind (*156*).

2.7. Breast Cancer Patient database analysis

2.7.1. Data Acquisition

RNA expression information from The Cancer Genome Atlas (TCGA) (*157*) and Genome-Tissue Expression (GTEx) (*158*) database was obtained from the University of Santa Cruz (UCSC) Xena platform (*159*) normalizing the RNA expression values by the upper quartile method (*160, 161*). Subgroup classifications of breast cancer into normal-like, luminal A, luminal B, and basal-like were based on the patients identified as such in the original PAM50 paper (*162*). Protein expression information from breast cancer was

obtained from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) (*163*) from cBioportal (*164, 165*). Correlations between genes using their RNA levels from TCGA or protein levels from CPTAC were calculated based on cBioportal using Spearman correlation coefficients, and statistical significance was determined based on the cBioportal adjusted p-value accounting for multiple hypothesis testing. Additionally, recently published data from Fudan University Shanghai Cancer Center (FUSCC) of the transcriptomes of an expanded patient cohort of triple-negative breast cancer (TNBC) was used to test for clinical correlates to KDM2B (*166*). Fragment per kilobase million (FPKM) count matrices were downloaded from Bio-Med Big Data Center (www.biosino.org) under the node identification OEZ000398. Basal versus non-basal TNBC identification was based on the patient characteristics published by the same group (*166*).

2.7.2. Pre-Ranked GSEA on Correlations

Correlations between KDM2B RNA levels and all expressed genes were downloaded from cBioportal under the total breast cancer database or just the basal-like subgroup as identified(*162*). The same was performed for KDM2B protein levels and all detected protein targets; however, the protein correlation analysis was not done for the basal-like as there were not enough samples in the subgroup alone to accurately rank correlation coefficients. Genes from a list of all detected RNA gene expression or protein expression correlation coefficients with KDM2B were ranked from most positively correlated to most negatively correlated. These pre-ranked lists were used as the input for GSEA to perform pre-ranked enrichment analysis (*145*). Similar to the RNAseq enrichment analysis, 10000 permutations were run for each gene set, and an enrichment

score was calculated for each permutation and compared to the actual enrichment score of our input pre-ranked list. Statistics were determined based on the probability that the input ranked list had a better absolute enrichment score than the 10000 permutations and adjusted for multiple hypothesis testing considering the total number of gene sets considered. The GO Biological term “Ribosomal Biogenesis” was specifically interrogated in the data.

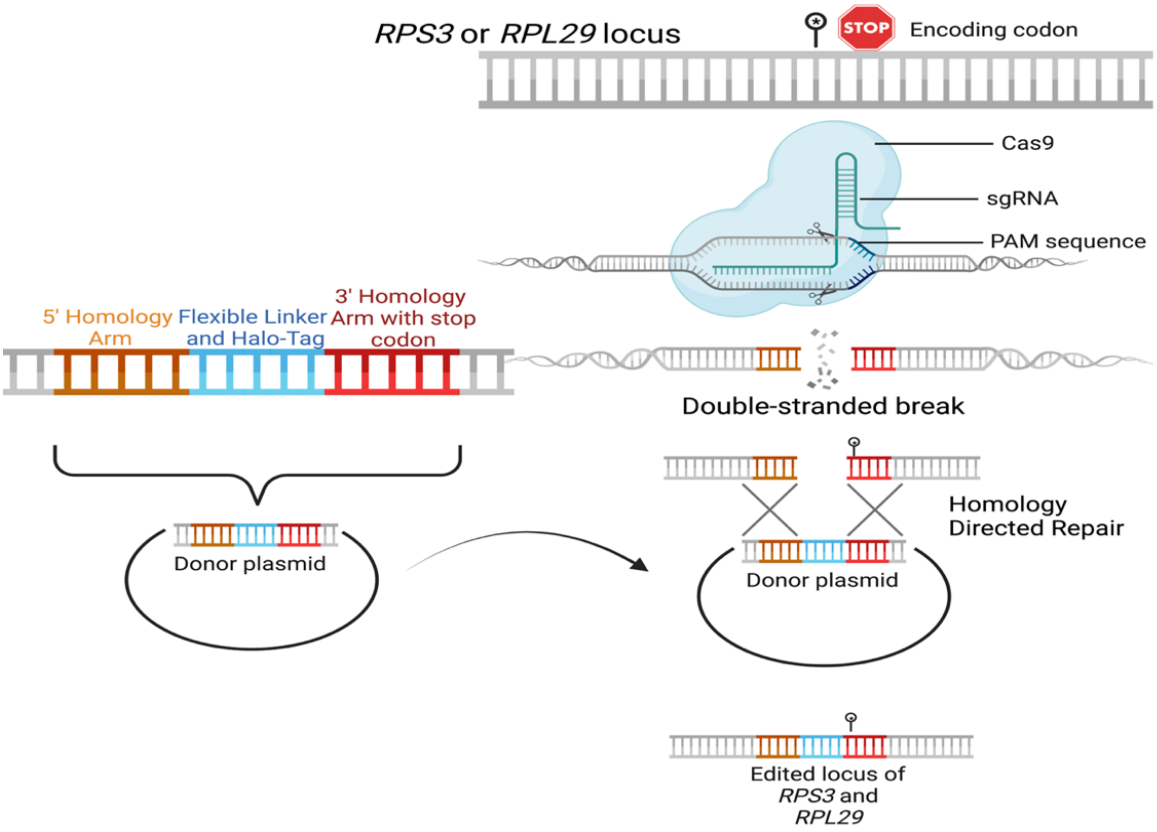


Figure 2.1 - CRISPR Cas9 Editing of Homologous Loci
Diagram depicting strategy of genomic editing of *RPS3* and *RPL29*. sgRNA were designed to target the last coding exon of each loci. A plasmid containing homologous arms matching the 5’ end of the cleavage (as based on the mRNA sequence), the Halo-tag sequence, and the 3’ homologous end of the cleavage was used as a donor plasmid to serve as the template for homology directed repair and edit the individual loci. Figure was made with Biorender (https://biorender.com/).

2.8. CRISPR Cas9 Gene Editing

Editing of the *RPS3* and *RPL29* endogenous loci with HALO-tags was performed as previously described (*167*). Guide RNA (sgRNA) oligos were designed to target the C-terminal end of each locus proximal to the stop codon of the last coding exon. The sgRNA oligos were annealed, phosphorylated, and ligated into an already digested and dephosphorylated vector as described (*168, 169*). The vector chosen contained Cas9 fused to a geminin tag. This ensured that Cas9 would only be expressed in the S or G2 phases of the cell cycle after the DNA had been replicated, increasing the chance for homology-directed repair over non-homologous-end-joining (*170*). Plasmids to serve as the donor plasmid containing homologous arms and HALO-editing were generous gifts from J. Wade Harper (*167*). The plasmids contained homology regions mapping to the region upstream of the stop codon and downstream of the stop codon. In between the regions, the coding region of a serine-glycine rich, flexible linker followed the HALO-tag, and the stop codon is cloned in (Figure 2.1). This plasmid will serve as the donor template for when homology-directed repair occurs to edit each locus and place an in-frame Halo-tag. Constructs used for transfection of Cas9 and the sgRNA are listed in Table 5.1 and of sgRNA design are listed in Table 5.2.

After transfection using Lipofectamine 3000™ of the sgRNA with Cas9 containing plasmid and the donor plasmid, MDA-MB-231 cells were allowed to grow and divide and then were treated with fluorescent Halo-Ligand either JFX-549 (Janelia Research Campus) or JFX-650 (Janelia Research Campus) (*171, 172*) and were investigated under an epifluorescence scope to see if any colonies of fluorescent cells were present, suggesting proper homology-directed repair using the donor plasmid. The

Gene Editing Shared Resource at The Ohio State University (Columbus, OH) performed fluorescence-activated cell sorting (FACS) on a SH800S Cell Sorter (Sony Cat: SH800S), and single-cell clones were plated in a 96-well cell culture plate (Corning, Cat: 3596). Individual clones were analyzed by western blotting to confirm proper editing.

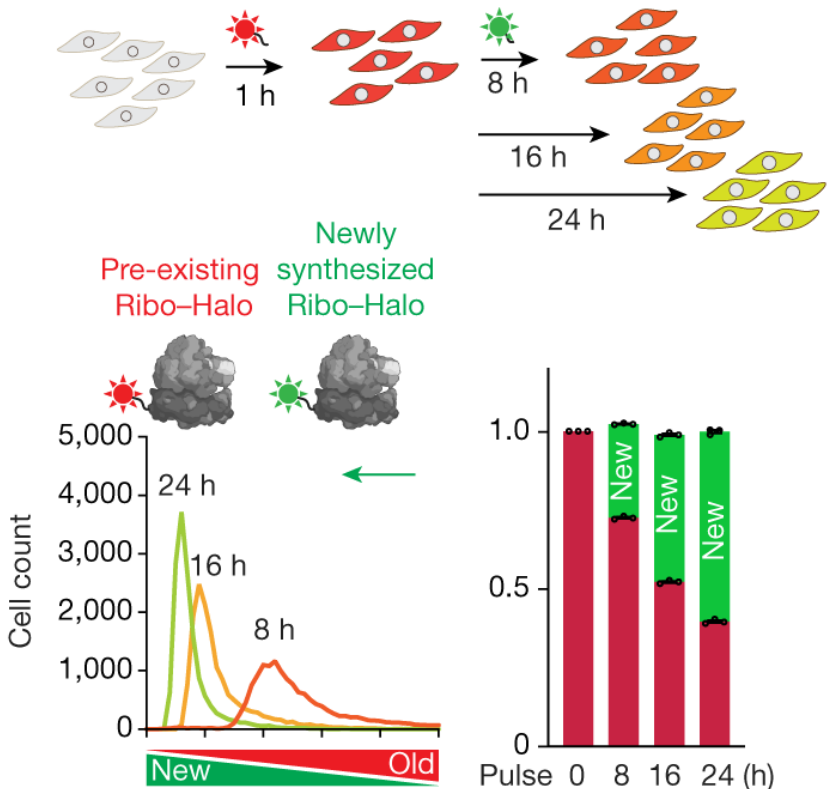


Figure 2.2 - Ribosomal Halo Pulse Chase
Reprinted with permission from An, H., Ordureau, A., Körner, M. et al. Systematic quantitative analysis of ribosome inventory during nutrient stress. Nature 583, 303–309 (2020). https://doi.org/10.1038/s41586-020-2446-y

2.9. Halo-Pulse Chase Labeling and Flow Cytometry Analysis

Edited clones of *RPS3-Halo* or *RPL29-Halo* were transduced with lentivirus containing control or shKDM2B. After transduction and puromycin selection, cells were plated onto a 60mm dish. One replicate plate for each condition was used to determine cell count on the day of staining. Based on the cell count of MDA-MB-231 cells, an

estimated molar amount of RPS3 and RPL29 protein levels was determined based on published molarity calculations per cell on OpenCell (*173*). Based on the estimated molar amount of the endogenous protein, 10x the molar amount of JFX-549 Halo-ligand was added to the cells for 1 hour to saturate the signal of all existing ribosomes currently present in the cell. Afterward, cells were washed and replaced with a different fluorescent halo ligand, JFX-650, to label nascent ribosomal production (*167*) (Figure 2.2). As shKDM2B transduced MDA-MB-231 cells grow slowly (*85*), the pulse of JFX-650 was chased for a longer course of 24 hours to allow a more representative capture of steady-state ribosomal production.

2.10. Polysome Profiling

In order to extract cells at similar confluence, cell counting and plating the day before cell pelleting and freezing were optimized to account for differential growth rates of control and shKDM2B-transduced cells (*174*). It was determined that a ratio of 0.6 (control count/shKDM2B count) with plating 16 hours prior to extraction gave approximately even counts at the point of extraction. Thus, 2.7 million control MDA-MB-231 cells and 4.5 million shKDM2B-transduced MDA-MB-231 cells were plated into 10cm dishes. The following day, one dish from each condition was counted in order to estimate the cell count of replicate plates to be used for the polysome profiling. Afterward, the remaining plates were washed twice in ice-cold 1xPhosphate buffered saline (PBS) (Sigma Cat: D8537) containing 100µg/mL of Cycloheximide (CHX) (Sigma Cat:C4589), keeping the cell plate on the ice at all times. The last wash is aspirated, leaving ~500µL of liquid left, and the cells are scraped gently into a centrifuge tube and centrifuged at 300xg for 5 minutes at 4^{o}C. Any supernatant is quickly pipetted off, and

the remaining cell pellets are flash-frozen by immersion in liquid nitrogen and stored at -80 °C until ready for polysome profiling. 40-45million cells were placed in each flash frozen cell pellet, and this was repeated to obtain 3 biological replicates.

Polysome profiling was performed as previously described (*175*). 10% sucrose solution (10% w/v sucrose, 10mM Tris pH 7.4, 140mM KCl, 10mM $MgCl_2$, 1mM DTT, and 100µg/mL CHX) and 50% sucrose solution (50% w/v sucrose,10mM Tris pH 7.4, 140mM KCl, 10mM $MgCl_2$, 1mM DTT, and 100µg/mL CHX) were made fresh. The 50% sucrose solution was carefully placed as the bottom layer with a syringe, followed by the 10% sucrose solution, carefully layering them, preventing them from mixing. A Gradient Master (Biocomp Cat: 108) was employed to produce a sucrose gradient fraction from 10-50%. The frozen cell pellets were lysed in at least 400µL of polysome lysis buffer (1% Triton-X100v/v 20mM Tris pH 7.4, 140mM KCl, 10mM $MgCl_2$, 1mM DTT, and 100µg/mL CHX) normalizing the volume of lysis for the estimated cell number. The cell pellet was resuspended, vortexed, and placed on ice for 10min to complete the lysis. Afterward, the lysates were centrifuged at 2000xg, and the supernatant was collected. 400µL of lysis buffer was carefully loaded onto each sucrose gradient without disturbing the gradient. Each gradient with lysate was then centrifuged on an ultracentrifuge at 260,000xg for 2 hours with slow acceleration. Brakes were not used for deceleration as this could disturb the gradient; instead, the centrifugation was allowed to decelerate on its own for approximately 1-1.5 hours. The sample is then run from top to bottom (lightest to heaviest fractions) on a gradient fractionator (Biocomp Cat: 152) with absorbance measured at 260nm for RNA species using a Triax 3 wavelength flow cell (Biocomp Cat: FC-3-260-2VIS).

2.11. TMT Proteomics and Analysis

For proteomics analysis, MDA-MB-231 control and shKDM2B cells were trypsinized and counted to be plated at a 0.6 ratio (control/shKDM2B) as described under Polysome Profiling. Thus, 1.5 million cells of control MDA-MB-231 and 2.5 million cells of shKDM2B MDA-MB-231 were plated onto 10cm cell culture dishes in order to have a confluency of 50-60% the following day. Cells were washed 2x in 1xPBS before adding proteomics lysis buffer (8M Urea, 50mM triethylamonium bicarbonate) containing Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific Cat: 78440). The samples were sonicated for 30 seconds at an amplitude of 25% on ice and then quantified using Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, 23225).

The samples were sent to Bioinformatics Solutions Incorporated (Waterloo, ON, Canada), which performed the trypsin digestion, Tandem Mass Tag (TMT) labeling, fractionation, mass spectrometry, and initial analysis. Briefly, 100µg of protein was reduced-alkylated-precipitated and then digested with trypsin overnight. Digested samples were then cleaned by a C18 Stagetip and quantified by Pierce™ Quantitative Peptide Assays & Standards (Thermo Fisher Scientific Cat: 23275). Then, 20µg of each sample was used for TMTpro16plex labeling. Labeling efficiency was determined by running 1µL of each sample on LC-MS/MS and found to have >96% labeling efficiency. The 8 samples (n of 4 control, n of 4 shKDM2B) were then pooled and fractionated by high-pH reverse phase to 44 fractions, which were then combined into 11 fractions. Each fraction was injected twice and analyzed by liquid chromatography–mass spectrometry (LC-MS) on an UltiMate™ 3000 RSLCnano System liquid chromatography system

(Thermo Fisher Scientific Cat: ULTIM3000RSLCNANO) attached to an Orbitrap Fusion™ Lumos™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific Cat: IQLAAEGAAPFADBMBHQ). The mass spectrometry method used was an MS^3 quantification with synchronous precursor selection. Searches of peptide spectra were performed on the Uniport-validated human proteome (*176*) by the group using their in-house software Peaks Studio (v11) (Bioinformatics Solutions Incorporated). The searches were performed using a 15-ppm precursor ion tolerance for peptide mapping. A decoy-fusion method was employed (*177*) to determine if peptide-spectrum matching are true matches or likely false positives. A cut-off of 1% false-discovery-rate based on the was used to filter poor peptide matching. Statistical significance and fold change were calculated by comparing the protein quantification values of control and shKDM2B-transduced samples utilizing Peaks Studio (v11) (Bioinformatics Solutions Incorporated).

2.12. Northern Blotting

2.12.1. Probe Synthesis

Primers were designed to amplify ~200nt regions corresponding to the external transcribed spacers (ETS) and internal transcribed spacers (ITS) of the rRNA genomic loci. Four sets of primers were designed targeting the 5' ETS, ITS1, ITS2, and 3' ETS regions (*178*). The sequences of the oligos used to amplify these regions are listed in Table 5.2. The 3' ETS primer design could not be specific enough for genomic amplification without designing the forward primer to start on the mature 28S rRNA encoding region. Thus, 20nt of the 3' ETS primer is complementary to the mature 28S rRNA. The remaining primers were exclusively in the spacer regions of the rRNA. Genomic DNA (gDNA) was amplified by PCR utilizing Phusion polymerase (New

England Biolabs Cat: M0530S) or Platinum™ SuperFi II polymerase (Thermo Fisher Cat: 12361010). The PCR product was run on a 1% agarose gel to visually inspect the quality and extracted from the gel using the Nucleospin® PCR and gel clean-up kit (Macherey Nagel Cat: 740609). The extracted amplicon underwent another PCR amplification with the same forward primer. However, the reverse primer had an overhang that added a T7 viral RNA polymerase promoter 3'end of the amplicon. The PCR was purified using the Nucleospin® PCR and gel clean-up kit (Macherey Nagel Cat: 740609) and quantified using a NanoDrop™ One/OneC Microvolume UV-Vis Spectrophotometer.

200ng of T7 addended PCR amplicon was used as the input for each RNA probe synthesis using the DIG RNA Labeling Kit (SP6/T7) (Roche Cat: 11175025910) following manufacturer specifications. 10% of the reaction volume was run on an agarose gel to determine if the transcription reaction was efficient. Finally, the dig-11-1UTP synthesized probe was then stored at -80°C until ready to use.

2.12.2. RNA Extraction and Northern Blotting

RNA was extracted using Purelink™ RNA Mini kit (Invitrogen Cat: 12183018A) extracted was run on an Agilent Bioanalyzer utilizing an RNA 6000 Nano Kit™ (Agilent Technologies Cat: 5067-1511) for quality and quantity.

Northern blotting was performed using the NorthernMax™ Kit (Invitrogen Cat: AM1940) following the manufacturer's protocol with buffers and formulations provided in the kit. 1µg of total RNA was mixed with 3 volumes of formaldehyde loading dye with 10µg/mL ethidium bromide added to the mixture. The mixture of each RNA sample was loaded onto a denaturing agarose gel and run in 1xMOPS buffer at 100V for 2 hours. The

gel was briefly imaged under ultraviolet light in order to make sure the RNA had run properly.

After the gel had finished properly running, the transfer setup was prepared as described in the kit protocol. The set-up from bottom to top was multiple dry paper towels, two sheets of dry chromatography paper, three sheets of chromatography paper that were pre-wet with transfer buffer, the BrightStar™-Plus Positively Charged Nylon Membrane (Invitrogen Cat: AM10100), the denaturing RNA gel, three sheets of chromatography paper that were pre-wet with transfer buffer, and finally a bridge of chromatography paper leading to a reservoir of transfer buffer. The RNA was transferred from the gel to the nylon membrane through capillary action and gravity for two hours.

After transfer, proper transfer was confirmed by shining an ultraviolet light and seeing the lingering ethidium bromide stain that had adequately transferred onto the membrane. The RNA was cross-linked by ultraviolet crosslinking using a UV Stratalinker™ 2400 (Agilent Cat: Discontinued). The membrane was then pre-hybridized in 10mL of ULTRAhyb™ Ultrasensitive hybridization buffer at 68°C in a heat-sealable bag for 1 hour. Afterward, an additional 1mL of hybridization buffer containing the dig-UTP labeled antisense probe was incubated at 68°C overnight.

The unbound probe was stripped from the membrane by performing two washes of 2XSSC with 0.1% SDS (300mM NaCl, 30mM sodium citrate pH 7.0, 0.1% w/v SDS) at room temperature for 5 minutes each wash, followed by two washes of 0.1XSSC with 0.1% SDS (15mM NaCl, 1.5mM sodium citrate pH 7.0, 0.1% w/v SDS) at 68°C for 15 minutes each wash.

Afterward, the membrane was washed and prepared for immunological detection using the DIG-Starter Kit (Roche Cat:12-039-672-910). The membrane was washed twice in washing buffer (0.1 M maleic acid pH 7.5, 0.15 M NaCl, 0.3% v/v Tween® 20) for 5 minutes each time, followed by 30 minutes of blocking solution (0.1 M maleic acid, 0.15 M NaCl,1x Blocking Solution) at room temperature. Antibody incubation in anti-digoxigenin conjugated to alkaline phosphatase was performed in the blocking solution containing a 1:15000 dilution of the anti-dig antibody for 1 hour. The membrane was then washed twice in washing buffer for 15 minutes each time before the final wash in detection buffer (0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5). Finally, the membrane was placed in a plastic pouch, and CDP-Star® (Roche Cat: 12041677001) was added dropwise onto the membrane, which was incubated for 5-10 minutes.

The membrane was imaged on the 600nm and chemiluminescence channels of a LI-COR Odyssey® Fc Imaging System, with a 10-minute exposure for each channel to capture the remaining ethidium bromide stain and probe-antibody signal, respectively.

2.13. Biorthogonal Amino Acid Nascent Peptide Labeling

2.13.1. Methionine Starvation, Pulse-chase, Cellular Lysis

Cells grown to 60% confluence in a 10cm dish were starved of methionine, cystine, and glutamine for 1 hour using DMEM without L-methionine, L-cystine and L-glutamine (Sigma Cat: D0422) with 10% FBS similar to what has been previously reported (*179*). After starvation, the media was replaced with DMEM with 10% FBS containing L-cystine, L-glutamine, and 100μM L-homopropargylglycine (L-HPG) (Click Chemistry Tools Cat: 1067) and allowed to incubate for the specified time points and then extracted in 1mL of 0.16% SDS in 1xPBS with Halt™ Protease and Phosphatase

Inhibitor Cocktail (Thermo Fisher Scientific Cat: 78440) with cell scrapping. Protein concentration was determined using the Pierce™ BCA Protein Assay Kit (Thermo Fisher Scientific, 23225), and 2mg of protein was loaded into a fresh tube for each sample, and the solution was filled to 2mL with 1xPBS.

2.13.2. Click-Chemistry Reaction

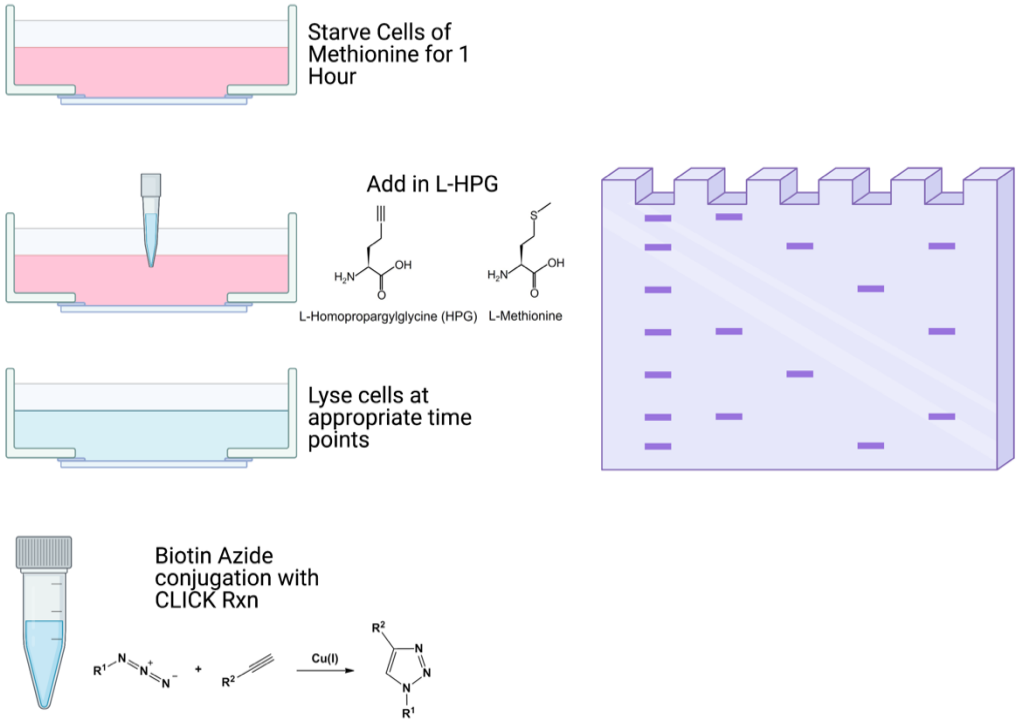


Figure 2.3 - Methodology of biorthogonal amino acid labeling and nascent peptide production.
Diagram describing steps of pulse labeling cells with L-HPG and using click-chemistry to link it to Biotin in order to label nascent peptides. Figure was made using Biorender (https://biorender.com/).

Stock solutions that were 100x concentrations for the click-chemistry reaction were made, resulting in 10mM PEG4 carboxamide-6-Azidohexanyl Biotin (Biotin-Azide) (Click Chemistry Tools Cat: 1265), 10mM Tris[(1-benzyl-1H-1,2,3-triazol-4-yl)methyl]amine (TBTA) (Sigma Cat: 678937), 100mM Tris (2-carboxyethyl) phosphine

(TCEP) (Sigma Cat: C4706), and 100mM Copper (II) Sulfate (Sigma Cat: C1297). TCEP and Copper (II) Sulfate stocks were made fresh the day of the click-chemistry reaction while Biotin-Azide and TBTA were stored at -80 ºC. 20µL of each stock solution (100x Biotin-azide, 100x TBTA, 100x TCEP, and 100x Copper (II) Sulfate) were added to the 2mL of samples collected from L-HPG pulse chase. The samples were incubated in the dark while shaking for 2 hours at room temperature. After two hours, 10mL of pre-chilled acetone (-20 C) was added to each sample, and the samples were placed at -20ºC for the protein to precipitate overnight. The next day, the samples were then centrifuged at 4000xg for 30 minutes at 4ºC, and the acetone was removed. The protein precipitates were allowed to air dry and finally resuspended in 700 µL of 1.2% SDS containing Halt™ Protease and Phosphatase Inhibitor Cocktail (*179*). Lysates were then run on a gel and analyzed via western blotting as described, probing for total protein using Revert™ 700 Total Protein stain (Licor Biosciences Cat: 296-11020) and an anti-biotin antibody to detect biotinylated nascent peptides (Figure 2.3).

2.14. Paired RNA sequencing and Ribosome Protected Fragment sequencing and analysis.

2.14.1. Pellet Freezing and Cell Lysis

Following similar ratios of cell plating as noted in polysome profiling, MDA-MB-231 control and shKDM2B cells were trypsinized and counted to be plated at a 0.6 ratio (control/shKDM2B) with 1.5 million control cells plated, and 2.5 million shKDM2B-transduced cells plated into a 10cm cell culture plate 16 hours before the extraction in order to achieve even confluence of 50-60% the day of collection. The plates were washed twice in ice-cold 1xPhosphate buffered saline (PBS) (Sigma Cat: D8537)

containing 100µg/mL of Cycloheximide (CHX) (Sigma Cat:C4589), keeping the cell plate on ice at all times. The last wash is aspirated, leaving ~500µL of liquid left, and the cells are scraped gently into a centrifuge tube and centrifuged at 300xg for 5 minutes at 4ºC. Any supernatant is quickly pipetted off, and the remaining cell pellets are flash-frozen by immersion in liquid nitrogen and stored at -80 ºC until ready for polysome profiling. 10-12 million cells were placed in each flash frozen cell pellet, and this was repeated to obtain 3 biological replicates.

2.14.2. Paired Ribosome profiling

Library preparation and sequencing were performed by TB-Seq™ (San Francisco, CA). Cells were resuspended in ice cold lysis buffer (20 mM Tris-HCl, pH 7.4, 100 mM NaCl, 5 mM $MgCl_2$ 1% v/v Triton X 100, 1 mM DTT, 20U/mL Turbo DNase I, 0.1 % NP40, 100 µg/ml cycloheximide) and the soluble cytoplasmic fraction was isolated by centrifugation at top speed in a microcentrifuge for 20 min at 4 ºC (*180, 181*). Supernatants were collected and clarified lysates were digested with RNase I for 45 minutes at room temperature. Digestion was stopped with SUPERase-IN™ (Invitrogen Cat: AM2694), and monosomes were purified by size exclusion chromatography on MicroSpin S-400 HR columns (GE Healthcare Cat: GE27-5140-01) as described (*182*). Size selection of footprints with lengths 25-36 nt was performed by electrophoresis on 15% TBE-urea gels. Illumina-ready RIBO-seq libraries were prepared using a SMARTer smRNA-seq kit (TakaraBio Cat: 645031). Library concentrations were measured by qubit fluorometer, and their quality was assessed on an Agilent 2100 bioanalyzer. RIBO-seq libraries were sequenced on an Illumina Novaseq 6000 sequencer, single read, 1x50 cycles, for an average of 90 million reads per sample.

2.14.3. Paired RNA-seq

Library preparation and sequencing were performed by TB-Seq™ (San Francisco, CA). To obtain matched RNA-seq libraries, total RNA was purified from an aliquot of each cell lysate, and rRNA was depleted from the RNA-seq samples using a RIBO-Minus transcriptome isolation kit (Invitrogen Cat: A1083708) following manufacturer instructions. mRNA fragmentation was conducted for 25 min at 94 ºC to generate RNA fragments of sizes similar to those of the ribosome footprints. SMARTer smRNA-seq kit (TakaraBio Cat: 645031) was used to generate Illumina-ready RNA-seq libraries. Library concentrations were measured by qubit fluorometer, and their quality was assessed on an Agilent 2100 bioanalyzer. RNA-seq libraries were sequenced on an Illumina Novaseq 6000 sequencer, single read, 1x50 cycles for an average of 30 million reads per sample.

2.14.4. Identifying a canonical transcriptome of MDA-MB-231 cells

Given the short reads that are produced as a result of ribosome profiling, reads are more likely to map to multiple parts of the genome, confounding results and limiting analysis (*183*). Thus, it became necessary to identify a "canonical" transcriptome to limit where read fragments can align, as they would only be counted if they aligned to the transcripts known to be expressed as described by others (*75, 184*). Definitions vary on what defines a "canonical" transcriptome, with some groups defining it as the transcript isoform with the longest coding region of each gene (*184*) and others defining it as the most abundant transcript isoform of each gene (*75*). As we had already performed an RNAseq that had longer read lengths (2x 150bp), we were able to use that RNAseq to define the most abundant transcript of all the expressed genes in our cell line conditions. RNAdetector (*138*) was used to run SALMON (*185*) and obtain read counts at a

transcript isoform level. Expressed transcripts that were the most abundant isoforms based on SALMON were identified as the canonical transcriptome in our study.

2.14.5. Paired RNA-seq and Ribo-seq Analysis

Reads were first inspected for quality, and library preparation adapters were trimmed from the reads using TrimGalore (https://github.com/FelixKrueger/TrimGalore), which uses FASTQC (*135, 136*) to visualize sequencing quality and CutAdapt (*137*) to trim adaptor sequences. MultiQC was used to visualize all sequencing quality across all samples (*136*). In order to remove rRNA, tRNA, mitochondrial RNA, and mitochondrial rRNA from the Ribo-seq data, an indexed transcriptome containing sequences of all aforementioned RNA species was generated using bowtie2 (*151*); then, the Ribo-seq reads were mapped to that indexed transcriptome using bowtie2 (*151*) and unmapped reads were used as the input for downstream processing representing filtered reads devoid of rRNA, tRNA, mtRNA, and mtrRNA. To ensure ribosome profiling was correctly performed, filtered reads were aligned to the *Homo sapiens* genome (GRCh38) using bowtie2 (*151*) and then analyzed using riboWaltz to determine if the reads displayed proper trinucleotide periodicity. The p-site offsets at each fragment length were analyzed to determine the optimal p-site for each fragment length (*186*). RNA-seq and Ribo-seq reads were then aligned to the *Homo sapiens* genome (GRCh38) using STAR (*139*), and aligned reads were then quantified using featureCounts (140), only quantifying reads that mapped to the coding sequence of the most abundant expressed transcript as characterized by SALMON. The paired RNA-seq and Ribo-seq reads were then analyzed by deltaTE (*187*), which uses DEseq2 (*141*) to calculate the changes in the RNA levels and the changes in the RIBO level. Afterward, the comparison between RNA fold change

and ribosome profiling fold change between control and shKDM2B were analyzed to assign a change in translational efficiency (TE) of each transcript and determine specific translational changes. The TE is based on the fold change calculated in the Ribo-seq relative to the fold change calculated in the RNA-seq. Gene ontology analysis of transcripts with altered translational efficiency was analyzed by Metascape (*188*).

As the most abundant transcripts were already known, information about the sequence and structure of the 5' untranslated region (UTR) was downloaded from the University of California Santa Cruz (UCSC) Genome Browser in their foldUTR5 database as described by others (*75*).

2.14.6. Codon Analysis

As riboWaltz produces codon p-site density data (*186*), the p-site density at each codon was obtained for each sample. Each sample's p-site density on each codon served as the biological replicate to compare p-site density between the control and knockdown of KDM2B.

2.14.7. Ribosomal Heterogeneity Prediction Analysis

Ribosomal heterogeneity (*38*) was analyzed using dripARF (*189*). The 3-D structure of the ribosome, along with information on where specific ribosomal proteins sit, is well characterized. Thus, by taking the reads mapping to the rRNA loci from the Ribo-seq, one can ascertain if there is differential sensitivity to RNAse I at specific ribosomal protein binding locations along the rRNA sequences. Thus, one can predict that a ribosomal protein may be differentially present or missing between the control and knockdown of KDM2B.

2.15. Propidium Iodine staining and Flow Cytometry Analysis for Cell Cycle

Cells were trypsinized and centrifuged at 300xg, and then, the cell pellet was resuspended at 800µL of 1xPBS. While vortexing at a low speed, 2.5mL of 100% -20°C ethanol was added drop by drop until the entire volume was added. The cells were then placed at -20°C overnight to complete fixation.

The cells were then centrifuged at 300xg, the supernatant was removed, and the cells were washed once in 1xPBS and centrifuged again. The wash was removed, and the cells were resuspended in 0.5 mL of FxCycle™ PI/RNase Staining Solution (Invitrogen Cat: F10797). The cells were incubated for 30 minutes to complete staining and then underwent flow cytometry on a BD LSRFortessa™ Cell Analyzer (BD Biosciences). Flow cytometry was analyzed on FlowJo (v10) (https://www.flowjo.com/).

2.16. Statistical software and analysis

Statistical testing and plotting were performed using GraphPad Prism Software (v9.5.1) (https://www.graphpad.com/). The statistical test used for data analysis is labeled in each figure caption when appropriate. Statistical testing and plotting were also performed in RStudio (v2023.03.0+386) (https://posit.co/download/rstudio-desktop/) with similar statistical information displayed in each figure legend when appropriate.

2.17. Contributions

Vollter Anastas served as the lead for protocol development and optimization, writing of the materials and methods, and analysis design. Evangelia Chavdoula collaborated on protocol optimization. Alessandro La Ferlita collaborated on pipeline design and led the bioinformatic analysis. Michael G. Kearse designed the polysome

profiling protocol and aided in optimizing polysome profiling conditions for MDA-MB-231 cells.

Chapter 3: Results

3.1. Overview

Translational control accounts for 40% of protein variance seen in cells, whereas transcriptional control only accounts for 23% of protein variance (*12, 24-27*). Many signaling factors have been described in regulating translation, however, the factory of translation, the ribosome, has been more recently appreciated as participating in the regulation itself (*190*). Initially, it was thought that ribosomes were ubiquitous translators that translated any available mRNA transcripts present in the cell (*38*). This was supported by the fact that exogenous mRNAs, for example, those introduced artificially in experiments or those introduced by viruses, were readily translated into protein by the ribosome. However, recent literature has revealed that ribosomes are keystone pieces in gene regulation, affecting final protein outcome by selectively translating specific mRNAs only under certain conditions (*191*). Ribosomes themselves are able to do this by two major methods: 1. Having heterogeneous compositions where certain ribosomes have specific ribosomal proteins, have altered post-translational modifications, or have certain rRNA modifications (*191*) and 2. Imposing effects on the cell based on their concentration altering which mRNAs they are able to effectively translate with specific mRNAs recruiting ribosomes inefficiently and requiring high concentrations, whereas others translated efficiently in all regardless of concentration conditions (*39*). While these concepts of ribosomal biogenesis are being uncovered, the factors in a cell that establish ribosome heterogeneity or ribosome concentration remain poorly understood.

Efficient ribosomal biogenesis is critical for cancer cell growth and proliferation as the cells need to meet protein production demands in order to divide (*29, 114, 129,*

184, 192). It is during ribosome biogenesis that ribosomal concentration is established, and the final ribosomal abundance in a cell is determined. As mentioned in Chapter 1, ribosomal concentration is postulated to impact translational fidelity to specific transcripts (*29, 39, 40, 42*). Beyond concentration alone, during ribosomal biogenesis, heterogeneity in ribosomal composition through ribosomal proteins, binding partners, and rRNA modifications is established as well. This heterogeneity also impacts translation (*23, 34, 129, 131, 189*). Thus, during cancer development, many proto-oncogenes lead to some level of upregulation and alteration of ribosomal biogenesis as part of their function (*29*). In some cases, oncogenic genes do not impact growth directly, but their changes to the translatome provide a selective advantage for cancer cells. For example, in liver cancer, *MYC* overexpression in a *KRAS*-mutated driven model of liver cancer does not impact cell growth or cell division as compared to the *KRAS* mutation alone. In fact, *MYC* overexpression, despite its status as a transcription factor, minimally alters the transcriptome as compared to the solitary *KRAS* mutation. However, the *MYC* overexpression on top of the *KRAS* mutation promotes the selective translation of certain transcripts, in this case, *PD-L1,* that enables immune cell evasion and enhances cancer metastatic potential (*193*).

KDM2B is a proto-oncogene that is overexpressed in a variety of cancers. Previous work in the lab has uncovered that KDM2B overexpression is able to immortalize embryonic fibroblasts (*89, 194*) and bypass cellular senescence (*90, 126*). Additionally, KDM2B seems to promote more stem cell phenotypes overall (*84, 103, 195*). In line with evidence from non-cancer cells, KDM2B plays a critical role in multiple cancers (*88*), regulating growth (*118, 119*), stem cell self-renewal (*85, 86*), and

epithelial-mesenchymal transitions (*120*). While KDM2B appears to have an oncogenic role and specific genomic loci have been recognized as direct targets of KDM2B, the exact mechanism of how KDM2B functions is poorly understood. Reports vary, with some claiming that KDM2B function is dependent on its histone demethylase activity (*118*), others claiming it is through regulation of CpG island methylation (*103*), and others still claiming it is through the regulation of the polycomb complexes (*93*). Overall, most of the literature on KDM2B focuses on its repressive role as a transcriptional regulator; however, recent literature has suggested that KDM2B also plays a role in active transcription (*86, 103, 196, 197*). Concordant with these findings, KDM2B primarily binds to active gene loci as its DNA binding domain exclusively recognizes unmethylated CpG islands, suggesting it may function as an oncogene by promoting gene expression as opposed to repressing it. Previous work in the lab found that the knockdown of KDM2B led to decreased cell growth (85). However, the mechanism of reduced cell growth still remained unclear. Thus, we sought to characterize the global transcriptome that was dependent on KDM2B and determine which overall genetic programs KDM2B regulates. This study sought to elucidate the direct targets of KDM2B that could promote cell growth and promote oncogenic programs.

Here, we provide evidence that KDM2B is expressed at higher levels in breast cancer relative to non-cancerous breast tissue, especially in the basal-like breast cancer subtype. Furthermore, we demonstrate that the knockdown of *KDM2B* in a basal-like breast cancer cell line (MDA-MB-231) leads to potent downregulation of ribosomal biogenesis genes and a critical upstream transcriptional regulator, *MYC*. Interestingly, KDM2B displays a positive correlation with the identified downregulated genes in our

RNA-seq and also a strong positive correlation with the overall ribosomal biogenesis genes in patient samples of basal-like breast cancer. This highlights how the RNA-seq performed in cell lines may be representative of KDM2B's regulation in primary tumors from patient samples.

KDM2B demonstrates high co-occupancy with MYC at the proximal promotor elements of multiple genes, most notably ribosomal biogenesis genes, demonstrating that KDM2B likely has a direct effect on the expression of ribosomal genes. Knockdown of *KDM2B* reduces overall ribosomal abundance in the cells and, in turn, decreases the rate of global translation in the cells. The knockdown of *KDM2B* appeared to lower the levels of the 40S ribosomal subunit more potently than the 60S ribosomal subunit. Analysis of rRNA processing intermediates revealed that loss of *KDM2B* led to the accumulation of 18S rRNA processing intermediates with less notable effects on 28S rRNA processing, matching the previous finding of more potent repression of the 40S ribosomal subunit as opposed to the 60S ribosomal subunit. Finally, Ribo-seq analysis revealed multiple transcripts with differential translation efficiencies after the knockdown of KDM2B, with cell cycle regulators being particularly affected and more likely translationally repressed after the knockdown. In summary, KDM2B is an upstream transcriptional regulator of ribosomal biogenesis that impacts rRNA processing and ribosomal abundance. This, in turn, leads to both global and transcript-specific translational changes that impact cell growth and cell-cycle control.

3.2. Rationale

Previous work in the lab and work by others demonstrated that KDM2B is critical for cancer stemness and cell growth (*84-86, 119, 120, 123, 125, 126*). However, which

transcripts, globally, are prone to KDM2B regulation and how these targets promote breast cancer growth remains poorly characterized. Additionally, in light of conflicting evidence in regards to KDM2B's oncogenic function (*112-114*), it is imperative to perform a transcriptome study in order to provide insight into KDM2B's overall regulatory network and characterize how it can impact tumor growth, positively or negatively. As previous work in the lab had characterized some functions of KDM2B in breast cancer cell lines (*85*), the same cancer type was chosen as the model system to study KDM2B. Early investigation detailed that KDM2B was selectively overexpressed in basal-like breast cancer as opposed to other subtypes (see Section 3.3.1). A basal-like breast cancer cell line, MDA-MB-231, was chosen as a model for the study. MDA-MB-231 cells are a well-studied cell line, have a transcriptome signature extremely similar to that of patient tumors (*198*), have a chromatin signature in line with other basal-like breast cancer cell lines, and have one of the highest metastatic and growth potentials (*199*), allowing potential metastatic phenotypes to be interrogated. All of these reasons make the cell line a good model for studying gene regulation by KDM2B in aggressive basal-like breast cancer.

3.3. Results

3.3.1. KDM2B is expressed in higher levels in basal-like breast cancer.

KDM2B is expressed at higher levels in breast cancer as compared to normal breast tissue from GTEx (*158*). Additionally, KDM2B is more expressed in cancer tissue than in normal breast tissue, according to the data from TCGA (*157*) (Figure 3.1A). Subgroup analysis based on the PAM50 (*162, 200*) data demonstrated that its highest expression is observed in basal-like breast cancer (Figure 3.1B). Paired samples of

normal breast tissue and tumor breast tissue arising from the same patient were analyzed for RNA levels of KDM2B.

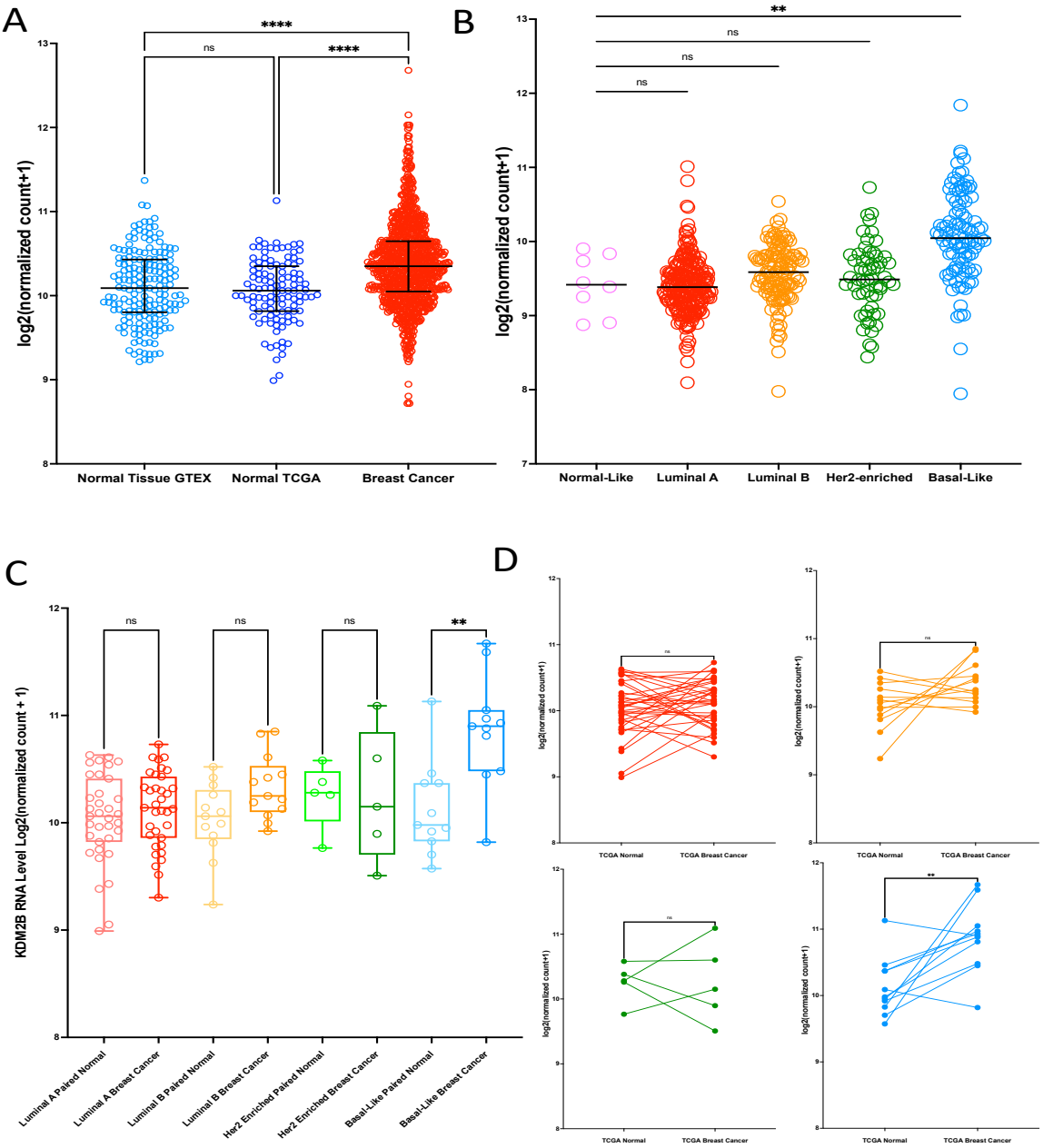


Figure 3.1 - KDM2B is highly expressed in basal-like breast cancer.
A. Comparison of KDM2B RNA levels between normal breast tissue from GTEx or TCGA and tumor breast samples. (Kruskal-Wallis test followed by a Dunn's multiple comparisons test). B. PAM50 subgroup comparisons all within tumor samples of KDM2B RNA levels (Kruskal-Wallis test followed by a Dunn's multiple comparisons test). C. KDM2B RNA levels in paired tumor and normal breast cancer samples in the same patient across the PAM50 subtypes. D. Bipartite graphs of paired comparing levels of KDM2B in tumor to the pair normal breast tissue (Wilcoxon signed-rank test) **p-value <

Interestingly, levels of KDM2B between tumor and normal breast tissue did not appear to be different in Luminal A, Luminal B, and Her2-enriched subtypes. However, paired analysis within the basal subgroup demonstrated that KDM2B was significantly increased in the tumor samples as compared to the normal tissue (Figure 3.1C, D). This indicates that the upregulation of KDM2B appears to occur mainly in basal-like breast cancer and that the increased levels are in tumor tissue as opposed to just higher in patients predisposed to basal-like breast cancer. This suggests that KDM2B has a pro-oncogenic function in basal-like breast cancer.

3.3.2. KDM2B knockdown impacts the expression of genes involved in ribosomal biogenesis and translation.

To identify the genes regulated by KDM2B in basal-like breast cancer, we employed RNA-seq, comparing the transcriptomes of control and shKDM2B MDA-MB-231 cells. Gene Set Enrichment Analysis (GSEA) (*145*) of the data, investigating the entirety of the expressed genes, demonstrated significant positive enrichment of the gene ontology (GO) categories “Growth factor response” and “BMP signaling,” and significant negative enrichment of the GO category “Ribosomal biogenesis” and “Translation” (Figure 3.2A). Ribosomal biogenesis had the largest magnitude of an absolute normalized enrichment score (Figure 3.2A, B), suggesting that the knockdown of KDM2B had the largest concerted effect on the expression of genes involved in ribosomal biogenesis. Reactome gene sets (*201*) were used to investigate KDM2B-dependent specific enrichment in a different database. Reactome enrichment analysis revealed a negative enrichment of genes involved in translation, similar to what was represented in the Gene Ontology analysis (Figure 3.2A, C). Given that cancer cells are

highly dependent on increased protein translation (*29*), this observation supports the role of KDM2B in oncogenesis. To visually show the role of KDM2B in the regulation of individual genes involved in ribosomal biogenesis, genes in the ribosome biogenesis GO group meeting an adjusted p-value threshold of < 0.05 were plotted in a heat map (Figure 3.2C) with the most significant changes being downregulated after the knockdown of KDM2B. In order to confirm that these targets were also downregulated at the protein level and not just the transcript level, TMT proteomics was employed to quantify the changes in the protein levels of ribosomal biogenesis targets. The proteins that corresponded to the significantly altered transcripts based on the RNA-seq were plotted to observe differences in protein levels (Figure 3.2D). Some protein targets were not detected in the TMT proteomics, and others were not significantly altered at the protein level despite the change in the RNA level. However, overall, protein levels of the detected proteins involved in ribosomal biogenesis were downregulated, mimicking the RNA-seq analysis that was performed (Figure 3.2D). This demonstrates that the majority of targets regulated at the RNA level by KDM2B had corresponding depletions in their protein level. Finally, as an orthogonal approach to confirm protein depletion, the expression of a subset of genes encoding ribosomal proteins (RPS28) or factors involved in the transcription and processing of ribosomal RNA (TCOF1, UTP14A, FBL, and NAT10) was also examined by western blotting. The results (Figure 3.2E) validated the change in the expression of ribosomal biogenesis factors at the protein level across 2 different shRNA constructs.

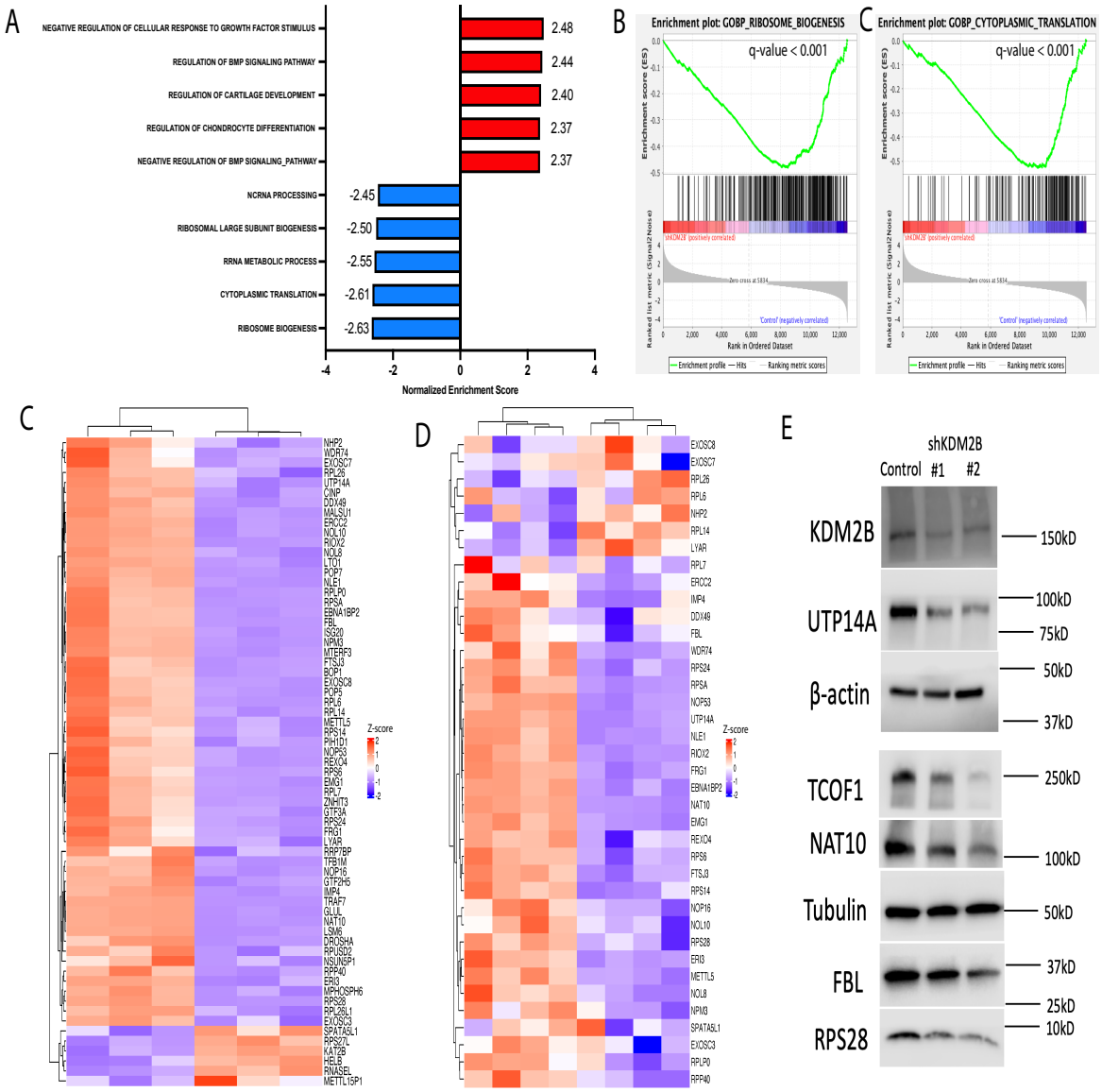


Figure 3.2 - Knockdown of KDM2B globally downregulates ribosomal biogenesis. A. Top 5 positively enriched and top 5 negatively enriched gene ontology terms after knockdown of KDM2B with normalized enrichment scored next to the bar graphs as calculated by GSEA. B. Ribosomal Biogenesis gene set enrichment plot with q-value listed accounting for multiple hypothesis testing. C. Cytoplasmic translation gene set enrichment plot with q-value listed accounting for multiple hypothesis testing. D. All significantly differentially expressed genes based on RNA-seq analysis comparing control and knockdown of *KDM2B* displayed on a heatmap. Heatmap made with ComplexHeatmap R package (https://github.com/jokergoo/ComplexHeatmap). Unbiased clustering of samples and genes were performed by the package. Signal intensities normalized for each gene based on Z-scores calculated form the geometric mean of each gene. D. Heatmap made with ComplexHeatmap of TMT proteomics intensities of the significant ribosomal biogenesis targets based on the RNAseq (C). Signal intensities normalized for each protein based on Z-scores calculated form the geometric mean of each protein. E. Western blot analysis of the comparing levels of downstream targets for selected ribosomal biogenesis factors.

Previous literature demonstrated that KDM2B regulated the expression of certain miRNA(*85, 126*) and this, in turn, impacted gene expression. Differential transcription would cause a change in intron levels and exon levels of a transcript as there would be decreased levels of the pre-mRNA; however, post-transcriptional regulation would lead only lead to changes in the exon levels as the effect of miRNA would only be on the mature mRNA transcript (*148, 149*). RNA fold change calculations between control and shKDM2B-transduced cells based on intron quantification and exon quantification were compared. Overall, intron and exon quantifications led to comparable results and correlated well with one another (Figure 3.3). Ribosomal biogenesis targets displayed high concordance between exon and intron quantification, suggesting they were regulated transcriptionally as opposed to post-transcriptionally (Figure 3.3).

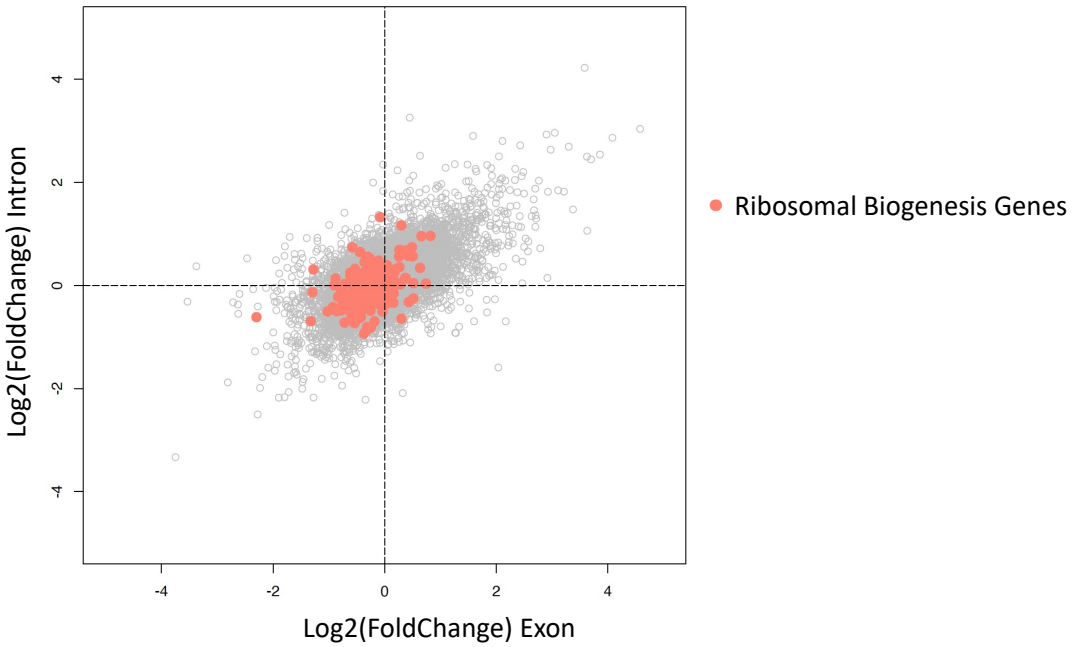


Figure 3.3 - KDM2B primarily regulates ribosomal biogenesis factors transcriptionally. Fold changes were calculated based changes in the intron level and changes in the exon levels of the transcripts detected in the RNA-seq performed. Fold changes were plotted to determine which transcripts were concordant and which changes were discordant.

3.3.3. KDM2B demonstrated high co-occupancy with MYC at target gene promoters.

As ribosomal biogenesis targets appeared to be transcriptionally regulated, we postulated that they were direct targets of KDM2B. KDM2B is a lysine-specific demethylase that binds to unmethylated CpG islands, mainly at the transcriptional start site of most target genes. It has been associated with both activating transcription and repressing transcription at target genes (*84, 86, 102*). Performing chromatin immunoprecipitation followed by sequencing (ChIP-seq) on the endogenous KDM2B, H3K4me3, H2AK119ub, H3K27me3, and MYC allowed us to investigate the binding of KDM2B relative to the chromatin landscape. H3K4me3 denotes actively transcribed genes, whereas H2AK119ub and H3K27me3 denote repressed genes. Notably, H2AK119ub is placed by the Polycomb Repressor Complex 1 (PRC1) which KDM2B functions as the DNA targeting component of the non-canonical PRC1.1 (*93, 102, 115*). In the non-canonical PRC1.1, H2AK119ub is placed first, which then recruits components of the Polycomb Repressor Complex 2 (PRC2), which then places H3K27me3 leading to chromatin compaction and gene repression(*100, 101*). Interestingly, KDM2B displays very little co-occupancy with H2AK119ub and H3K27me3 but displays high enrichment in active genes with high colocalization with H3K4me3 (Figure 3.4A). Additionally, KDM2B displayed high co-occupancy with MYC, a known regulator of ribosomal biogenesis (Figure 3.4A). Specifically, at target ribosomal biogenesis genes identified from the RNAseq, KDM2B is bound at their promoter along with MYC (Figure 3.4B, C, D, E). The ChIP-seq data demonstrated that

KDM2B is directly binding to these target genes and suggested that the regulation of those targets is direct.

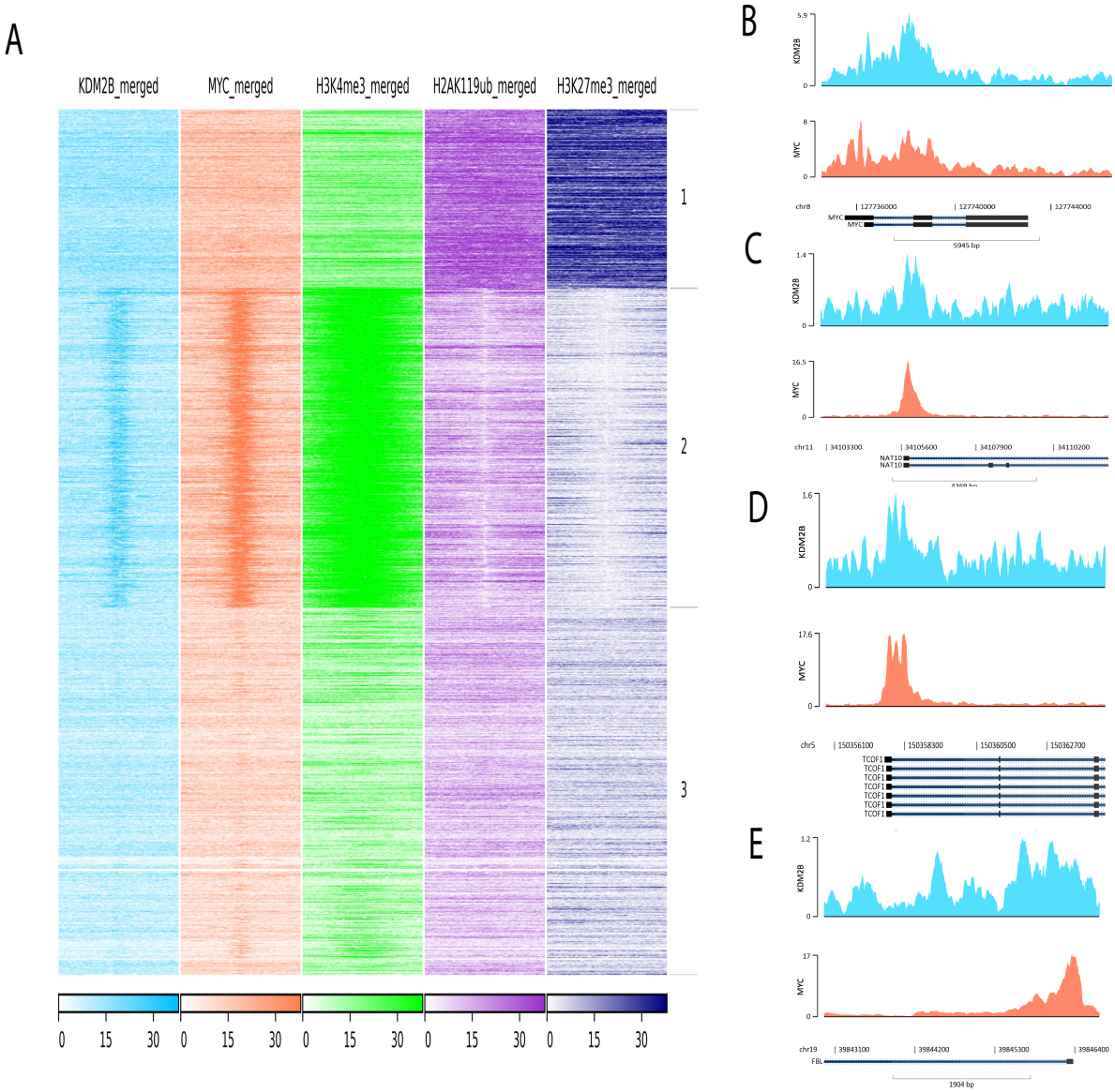


Figure 3.4 - KDM2B binds to active genes in the TSS along with MYC.
A. Heatmap representation of signal intensity at each transcript centered transcriptional start site (TSS) of each transcript expanding 5kb on each side. ChIPseq alignment of KDM2B and MYC intensity at the B. MYC promoter C. NAT10 Promoter D. TCOF1 promoter and E. FBL Promoter.

KDM2B, as part of the PRC1.1, acts in tandem with the PRC2 to repress chromatin. Previous data from the lab showed that KDM2B also regulates the levels of PRC2 components (*85*). Thus, we next asked if KDM2B was regulating gene expression through its action on PRC2 components. EZH2 is the main catalytic component of the PRC2 complex; therefore, we knocked down EZH2 to determine if the PRC2 is necessary for the expression of ribosomal biogenesis targets. In line with the ChIP-seq, where KDM2B has low co-occupancy with repressive marks (Figure 3.4A) knockdown of KDM2B reliably decreases protein levels of upstream ribosomal biogenesis factors; however, knockdown of EZH2 has no impact on their levels (Figure **3.5**). This demonstrates that the PRC2 is not critical for the expression of ribosomal factors. Therefore, KDM2B's regulation of the PRC2 components or KDM2B's function in recruiting the PRC2 is not essential for the expression of ribosomal biogenesis genes.

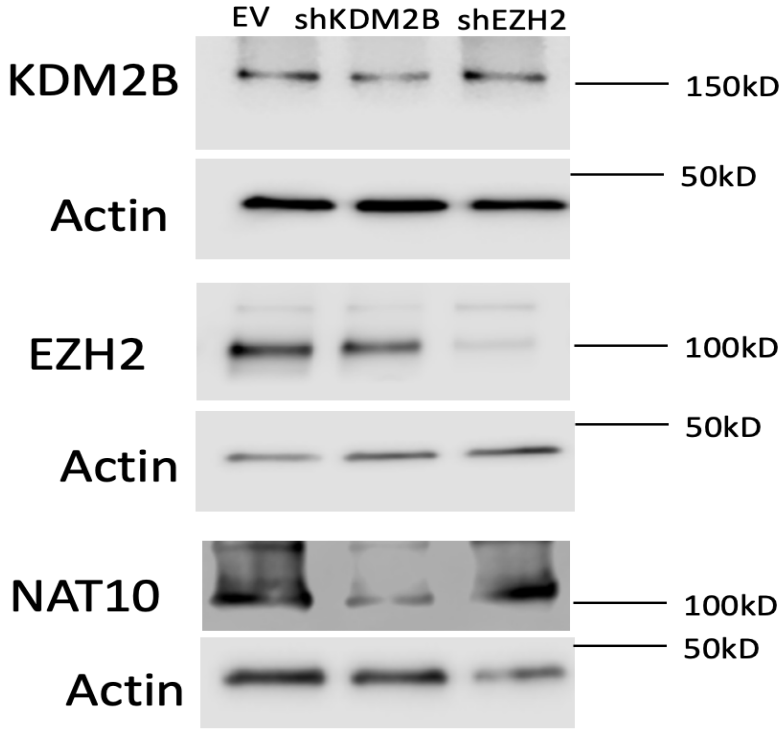


Figure 3.5 - Regulation of ribosomal biogenesis does not depend on the PRC2. Western blot analysis investigating the expression of NAT10 after knockdown of KDM2B or EZH2. Experiment was completed in biological triplicate.

Interestingly, previous reports had identified that KDM2B was able to affect the SWI/SNF complex and recruit it to target genes, leading to increases in chromatin accessibility at those loci (*84, 97, 197*). To investigate if there were changes in the chromatin architecture following the knockdown of KDM2B that could explain the downregulation of genes involved in ribosomal biogenesis, we utilized Assay for Transposase-Accessible Chromatin followed by sequencing (ATAC-seq). MYC had a modest decrease in accessibility in its promotor (log2(FC) = 0.6) (Figure 3.6A) however, the most dramatic changes appeared to be in gene loci that were poorly expressed or not expressed (Figure 3.6B, C). Some individual ribosomal biogenesis displayed decreased accessibility at their promotors, but the majority of the changes were modest.

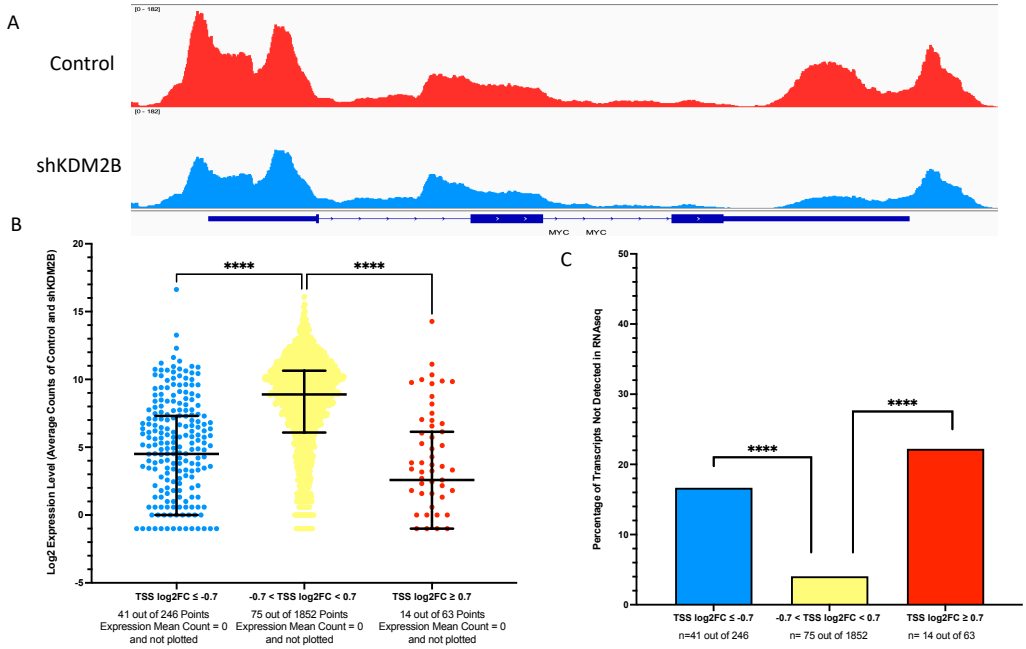


Figure 3.6 - Chromatin accessibility changes after knockdown of KDM2B
A. Chromatin accessibility signal normalized to total reads from merged 3 biological replicates of control and knockdown of KDM2B on the MYC genomic locus. B. Basal expression levels of transcripts categorized based on changes of chromatin accessibility at the transcriptional start site (TSS) of loss of chromatin accessibility (blue), no change (yellow), or gain of chromatin accessibility. Statistics were calculated with Kruskal-Wallis test followed by a Dunn's multiple comparisons test. C. Percentage of transcripts in each accessibility category of which there was no RNA expression detected. Statistics were calculated with a Fisher's exact test. ****p-value < 0.0001

3.3.4. KDM2B expression correlates with the expression of genes encoding ribosomal proteins and ribosomal biogenesis factors in basal-like breast cancer.

We next analyzed publicly available breast cancer (BRCA) datasets to determine whether KDM2B RNA levels correlate with the RNA levels of ribosomal biogenesis factors regulated by KDM2B (Figure 3.2). Our analysis of BRCA samples containing RNA-seq data in The Cancer Genome Atlas (TCGA) (*202*) revealed that KDM2B expression correlates positively with the expression of regulators of ribosomal biogenesis, which we identified as targets of KDM2B in our RNAseq study (*TCOF1, NAT10, FBL,* and *MYC*) (Figure 3.7A, B, C,D).

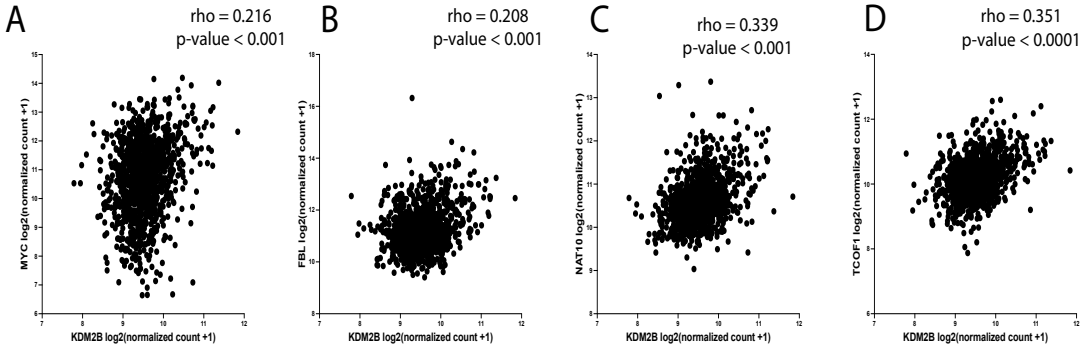


Figure 3.7 - *KDM2B* RNA levels correlates with RNA levels of upstream ribosomal biogenesis factors in breast cancer patients
Correlation plots generated from RNA expression data of TCGA breast cancer patients between *KDM2B* and A. *MYC*, B. *FBL*, C. *NAT10*, D. *TCOF1.* Correlations and statistics were calculated using spearman correlation test on Graphpad Prism.

KDM2B and its targets were expressed at higher levels in basal-like breast cancer (Figure 3.8A), raising the possibility that the observed correlations in the entire BRCA database (Figure 3.7A, B, C, D) could be due to confounding with the basal-like cancer subtype just having a higher coordinate expression of these genes. To address this

possibility, we examined the KDM2B correlations in basal-like BRCA, and we confirmed that the correlations hold true, even when we limit the analysis to a single subtype (Figure 3.8B, C, D, E). In order to assess more globally whether KDM2B correlated with overall ribosomal biogenesis factors, additional analyses were employed to investigate the entire subset of genes involved in ribosomal biogenesis. In these analyses, we ranked all of the transcripts detected in patient breast cancer data from most positively correlated to most negatively correlated with KDM2B, and we performed a pre-ranked GSEA. This revealed significant positive enrichment in genes involved in ribosomal biogenesis both in the entire BRCA cohort (Figure 3.8F) and the basal-like subgroup (Figure 3.8G). This result mirrors our RNA-seq data in the MDA-MB-231 cell line. This also holds true when investigating correlations at the protein level (Figure **3**.**8**H). Overall, these observations demonstrated that levels of KDM2B positively correlate globally with ribosomal biogenesis genes in patient breast cancer samples, mimicking the experimental data in the cell line.

A newer cohort of triple-negative-breast cancers that contains both basal-like and non-basal-like tumors was recently published from Fudan University Shanghai Cancer Center (FUSCC), providing an opportunity to test the strength of the KDM2B correlations in a larger set of tumors (*166*). Analysis of the entire cohort demonstrates similar findings to the results we obtained from the analysis of the TCGA dataset (Figure 3.8I). Notably, the correlations were stronger in the basal-like than in the non-basal-like subset of these tumors, once again highlighting the role of KDM2B in basal-like breast cancer.

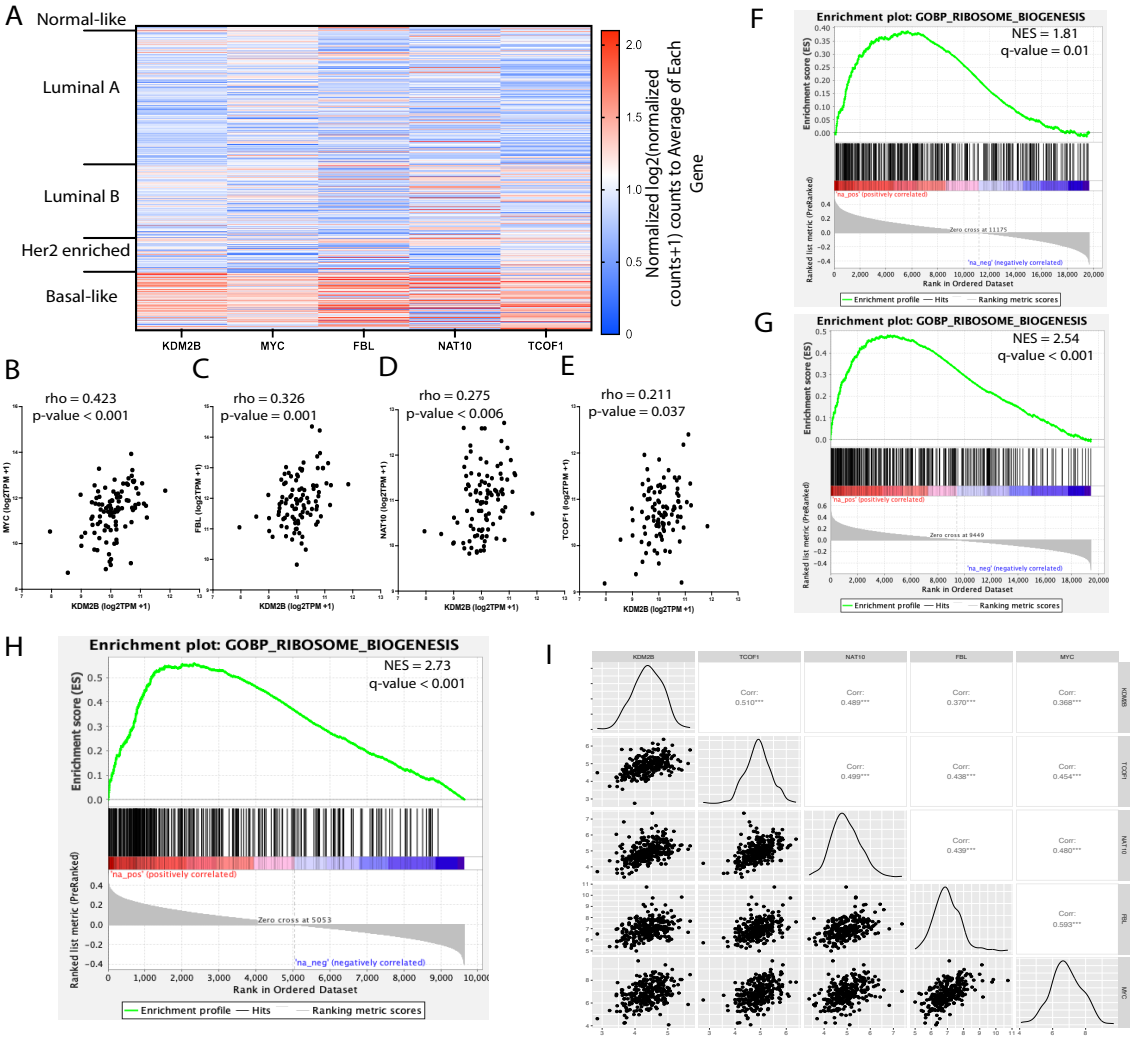


Figure 3.8 - KDM2B globally correlates with ribosomal biogenesis genes in basal-like breast cancer.

A. Heatmap of RNA expression of *KDM2B*, *MYC*, *FBL*, *NAT10*, and *TCOF1* grouped by PAM50 subtypes of breast cancer. Heatmap made on Graphpad Prism. Correlation plots generated from RNA expression data of TCGA basal- like breast cancer patients between *KDM2B* and B. *MYC*, C. *FBL*, D. *NAT10*, E. *TCOF1.* Correlations and statistics were calculated using spearman correlation test on Graphpad Prism. Pre-ranked GSEA ranking *KDM2B* RNA-level spearman calculated correlations with RNA levels of other transcripts from most positively correlated to most negatively correlated in F. All TCGA breast cancer and G. TCGA Basal-like breast cancer. H. Pre-ranked GSEA ranking KDM2B protein level spearman calculated correlations with other proteins from TCGA CPTAC database from most positively correlated to most negatively correlated. I. FUSCC triple-negative database displaying correlations between the different upstream ribosomal biogenesis genes and *KDM2B.* Correlation plots displayed on left, spearman rho-values and statistics displayed on right. Values displayed are log2(FPKM+1) as calculated by the group. Plot made with GGally R package (https://www.rdocumentation.org/packages) ***p-value <0.001.

3.3.5. The knockdown of KDM2B decreases ribosome numbers by lowering the rate of ribosomal biogenesis.

To determine whether the decrease in the expression of ribosomal proteins and ribosomal biogenesis factors reflects a reduction in ribosome numbers, we employed polysome profiling to determine the relative abundance of ribosomal subunits in control and shKDM2B-transduced MDA-MB-231 cells. Sucrose density centrifugation of equal amounts of cytoplasmic lysates (Figure 3.9A) from three control and three shKDM2B cultures revealed a more prominent free ribonucleoprotein peak in the shKDM2B lysates (Figure 3.9B) suggesting that the knockdown of KDM2B results in an increase in the abundance of free RNA that is not bound to ribosomes, a finding consistent with a low ribosome number. Focusing on the ribosomal subunits and the polysomes in the same density gradients revealed a significant decrease in the 40S subunits, a smaller decrease in the 60S subunits, a significant decrease in the 80S ribosomes, and a significant decrease in the polysomes of shKDM2B cells (Figure 3.9C).

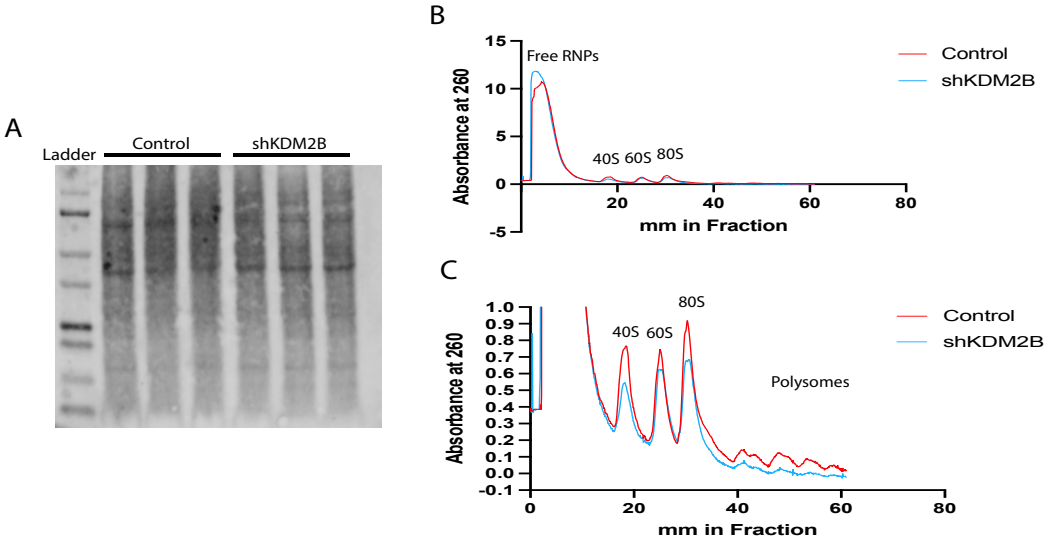


Figure 3.9 - Knockdown of KDM2B reduces ribosomal abundance.
A. Revert total protein stain of cytoplasmic lysates loaded for polysome profiling. Polysome profiling absorbance characterization showing B. Entire signal across the gradient or C. Just monosome and polysome fractions. Performed in triplicate with mean signal across all 3 replicates for each condition displayed.

The preceding data suggested a decrease in the abundance of ribosomes in steady state shKDM2B cells. Given that the expression of ribosomal proteins and ribosome biogenesis factors is downregulated in shKDM2B cells, we hypothesized that the decrease in the number of ribosomes is due to a reduction in the rate of ribosomal biogenesis. To address this hypothesis, we took advantage of the fact that free ribosomal proteins have a very short half-life, but they become stabilized by association with the ribosomes, whose half-life is exceedingly long (*203-205*). Given the rapid degradation of unbound ribosomal proteins, the accumulation of such proteins over time is proportional to the rate of ribosome assembly. Therefore, a change in their abundance over time provides a surrogate measure of the rate of ribosomal biogenesis. To monitor newly synthesized ribosomal proteins, we first tagged the endogenous RPL29 and RPS3 proteins with a C-terminal Halo tag as described (*167*), using CRISPR-Cas9-based gene editing in MDA-MB-231 cells. RPL29 and RPS3-edited, control and shKDM2B transduced cells were then treated to saturation for two hours with the Halo-tag ligand JFX549 (excitation 549nm/emission 570nm) (*171, 172*), which like all Halo-tag ligands binds the Halo-tag covalently and irreversibly. Following treatment and washing to remove the excess unbound ligand, non-labeled newly synthesized ribosomes were labeled with a second Halo-tag-ligand JFX650 (excitation 650nm/emission 667nm) (*171, 172*). To this end, the cells were pulsed with JFX650 for 24 hours, and immediately after, the relative number of newly synthesized ribosomes was determined by tracking the intensity of JFX650 fluorescence in single cells by flow cytometry. The results were plotted across 3 clones of MDA-MB-231 cells with edited RPL29-Halo and 3 clones of MDA-MB-231 cells with edited RPS3-Halo. The data demonstrated that the relative

JFX650 fluorescence was consistently decreased in all shKDM2B cells when looking at *RPS3*-edited cell lines (Figure 3.10 A, B) or *RPL29*-edited cell lines (Figure 3.10 C, D).

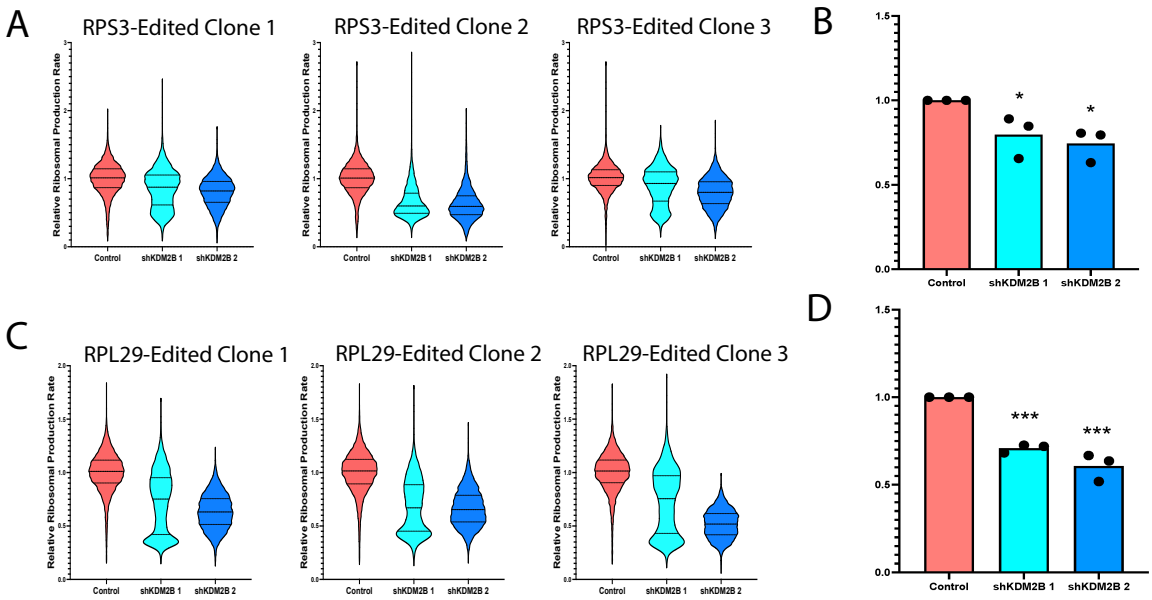


Figure 3.10 - KDM2B reduces the rate of ribosomal biogenesis.
Flow cytometry of individual cell signals displaying the ratio of fluorescence of newly labeled JFX650 signal to initially saturated JFX549 signal and then normalized to the median of control edited cell lines across 3 different clones of A. *RPS3*-edited cell lines and C. *RPL29*-edited cell lines. Median ratios for each clone and each condition were obtained as biological replicates and summarized for B. *RPS3*-edited cell lines and D. *RPL29*-edited cell lines. Statistics were calculated using a One-way ANOVA followed up a with Dunnett's multiple comparisons test on Graphpad Prism. *p-value < 0.05, ***p-value <0.001

3.3.6. The KDM2B knockdown disproportionally inhibits the biogenesis of the 40S ribosomal subunit.

The ribosome profiling (Figure 3.9C) suggested that in a steady-state, normally growing cells, there may be an imbalance in the abundance of the 40S and the 60S ribosomal subunits as the 40S subunit was more impacted than the 60S subunit. We reasoned that if there is an imbalance in the abundance of ribosomal subunits, there would also be an imbalance in the abundance of proteins associated with these subunits.

This question was addressed by interrogating our Tandem mass tag (TMT) mass spectrometry to examine the proteome of control and shKDM2B cells. The volcano plot (Figure 3.11) shows the change in protein abundance in shKDM2B relative to control cells. The results confirmed that the abundance of 40S proteins (Figure 3.12A) is decreased disproportionally relative to the abundance of 60S proteins (Figure 3.12B), supporting the hypothesis that the knockdown of KDM2B disproportionally inhibits the biogenesis of the 40S ribosomal subunit.

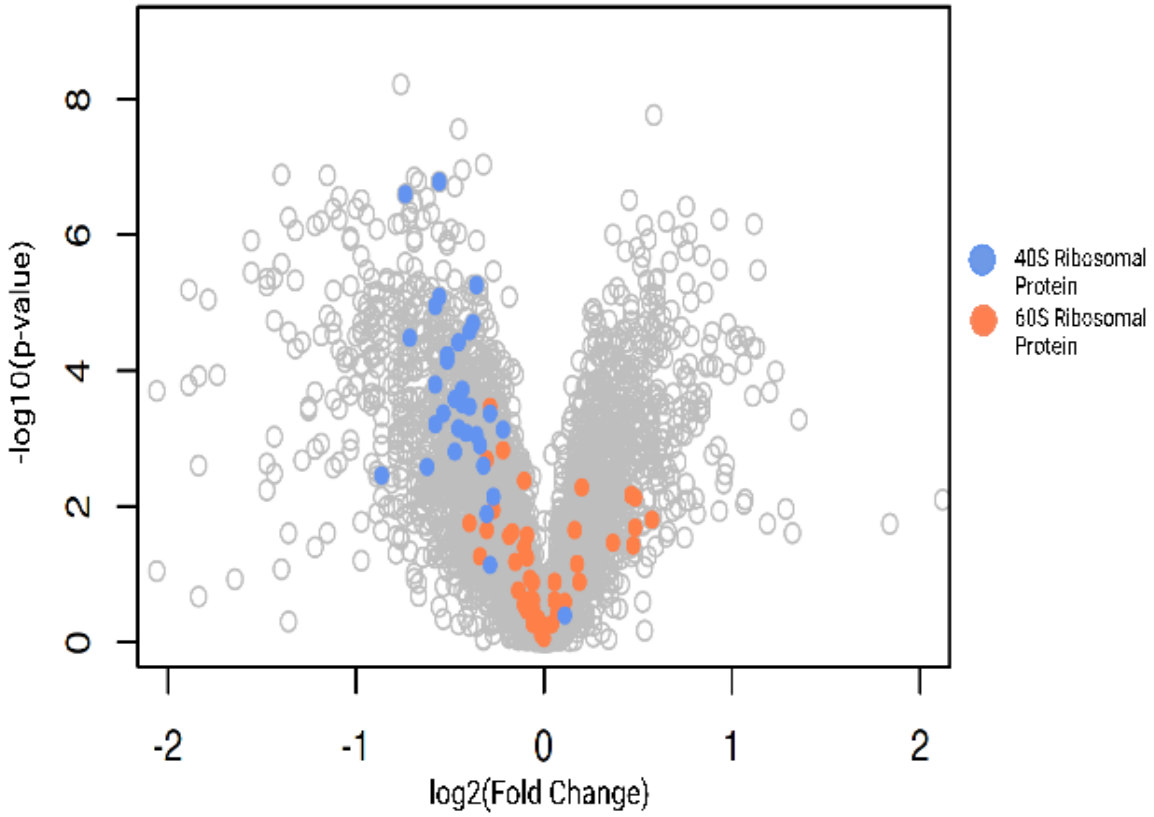


Figure 3.11 - KDM2B alters the proteome impacting ribosomal subunit abundance. Volcano plot of all detected proteins in the TMT proteomics analysis based on log2(fold change) and -log10(p-value). 40S and 60S ribosomal subunits are labeled as listed.

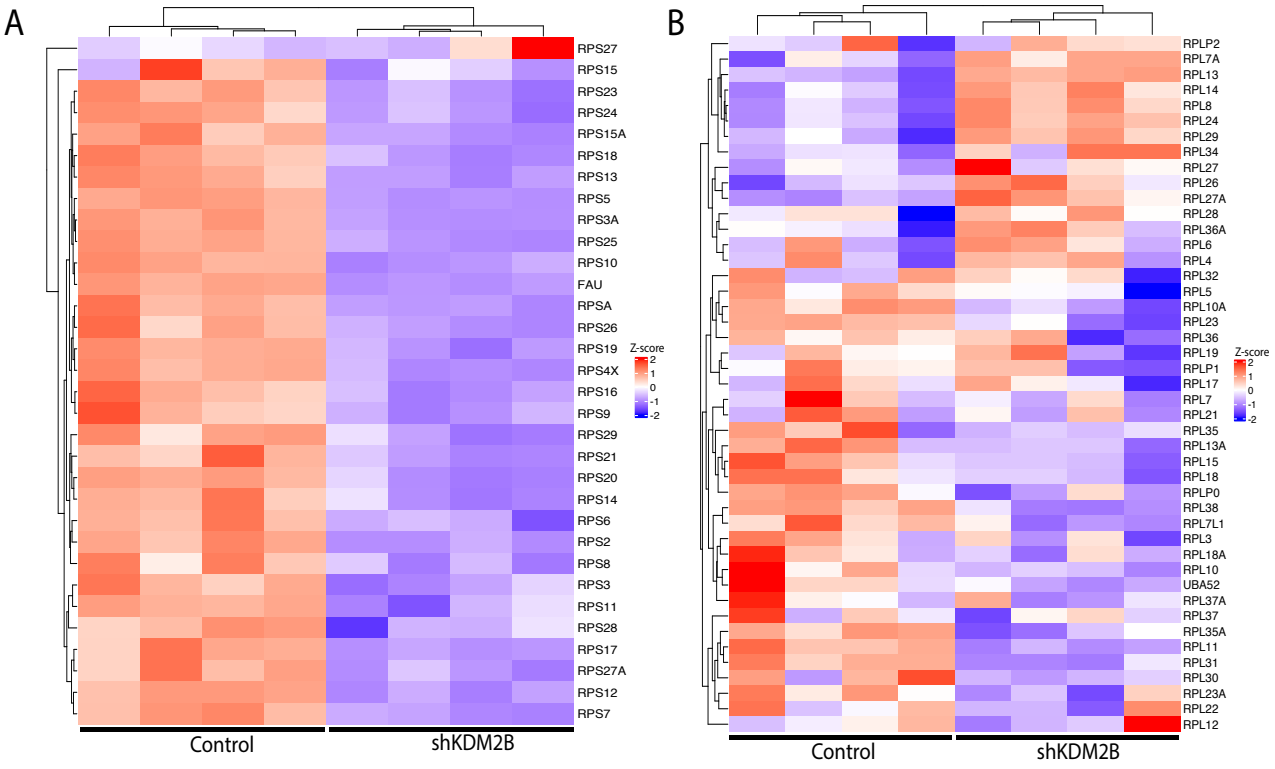


Figure 3.12 - KDM2B more potently repressed small ribosomal proteins as opposed to large. Proteins detected by TMT proteomics comparing control and knockdown of *KDM2B* displayed on a heatmap for A. 40S ribosomal proteins and B. 60S ribosomal proteins Heatmap made with ComplexHeatmap R package (https://github.com/jokergoo/ComplexHeatmap). Unbiased clustering of samples and genes were performed by the package. Signal intensities normalized for each protein based on Z-scores calculated form the geometric mean of each protein.

3.3.7. The knockdown of KDM2B interferes with pre-ribosomal RNA processing, primarily the processing leading to the mature 18S ribosomal RNA.

Ribosomal RNAs are transcribed from ribosomal DNA arrays. In humans, these rDNA arrays are located in the short arms of the acrocentric chromosomes 13, 14, 15, 21, and 22. Each unit of these arrays is transcribed as a single 47S RNA, which is then processed in the nucleolus to generate the mature 28S, 18S, and 5.8S ribosomal RNAs. As the mature rRNA species come from a single transcription event, we hypothesized that the depletion of 40S relative to the 60S ribosome was due to an imbalance in

processing. The 47S RNA contains, in addition to the sequences of the mature ribosomal RNAs, two External Transcribed Spacers (5' and 3' ETS) and two Internal Transcribed Spacers (ITS1 and ITS2). The spacers are removed during processing, and they are not represented in the mature ribosomal RNAs. Therefore, spacer sequences can be used as probes for Northern blots to investigate ribosomal RNA processing.

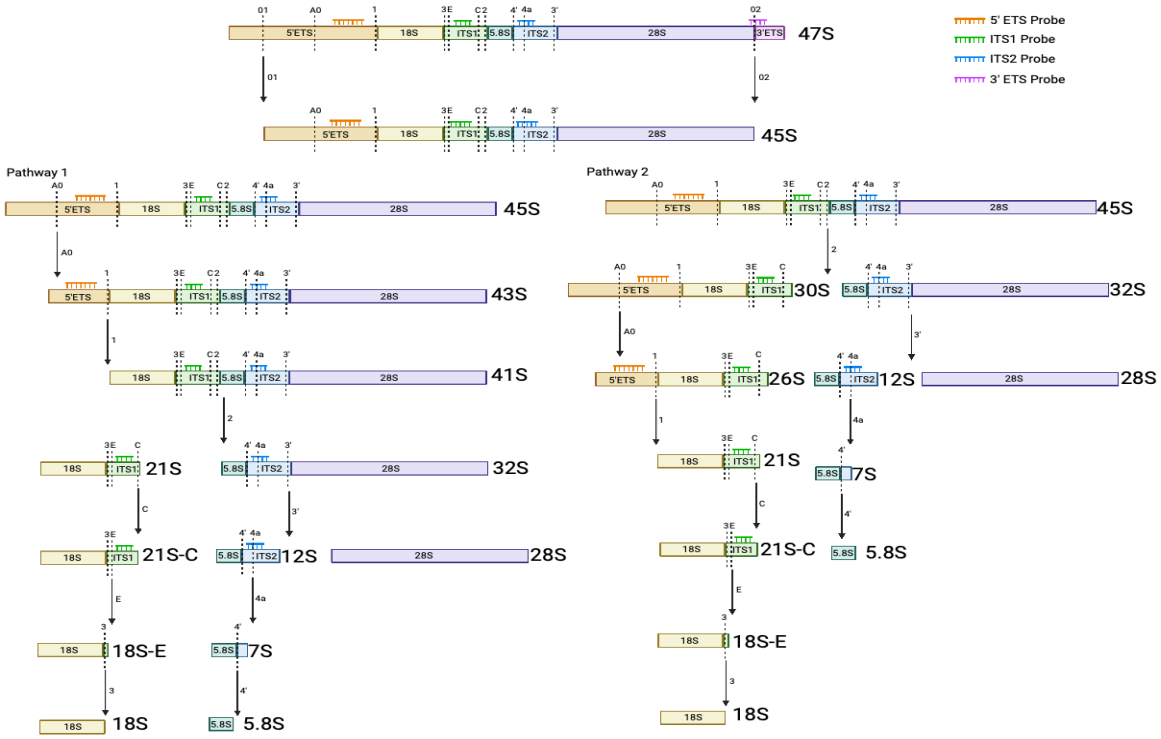


Figure 3.13 - Overview of rRNA processing steps and intermediates.
Diagram displaying the different pathways in humans of rRNA processing and the different cleavage events that occur. Additionally, probe design and what intermediates they would be able to detect via Northern blotting are also shown here.

The first rRNA processing step consists of two cleavage events, which remove the 3' ETS and a short fragment of the 5' ETS from the 47S RNA, giving rise to the 45S RNA. The 45S RNA is then cleaved at a site between ITS1 and 5.8S, giving rise to the 30S and 32S RNAs, which are precursors of the mature 18S and 28S rRNAs, respectively. The 30S RNA undergoes two cleavage events, which remove the 5' ETS in two steps, giving rise to the 26S and 21S precursors of the mature 18S rRNA.

Alternatively, these two cleavage events may occur in the 45S RNA prior to the ITS1/5.8S cleavage, giving rise to 43S and 41S RNAs. These and all the remaining rRNA processing steps are shown in (Figure 3.13).

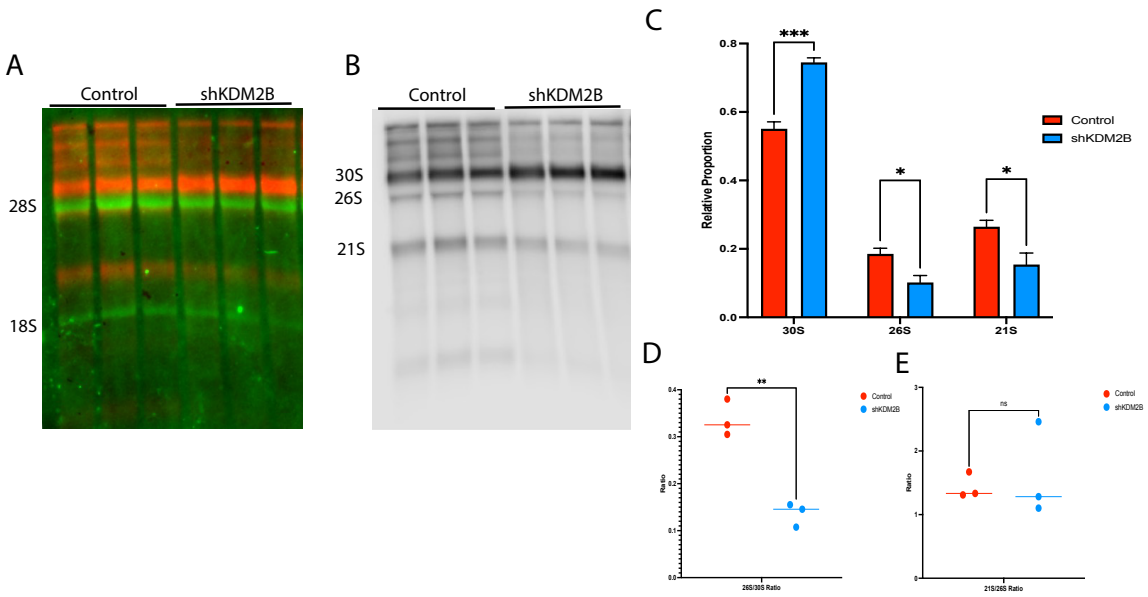


Figure 3.14 - ITS1 Northern blot probe after knockdown of KDM2B.
A. Northern blot signal from lingering ethidium bromide stain (green at 600nm) and chemiluminescence probe signal (red at chemiluminescence signal). B. Northern blot signal from antibody recognition of dig-probe hybridization. C. Quantification of relative proportion of rRNA intermediates with ratio of D. 26S/30S and E. 21S/30S. Statistics were calculated using multiple unpaired t-test correcting for multiple hypothesis testing by the Holm-Šídák method on Graphpad Prism; Error bars represent the standard deviation from the mean *p-value < 0.05, **p-value <0.01, ***p-value <0.001, ns not significant.

To determine whether KDM2B regulates rRNA processing, we first probed a Northern blot of total RNA from three control and three shKDM2B cultures of MDA-MB-231 cells with the ITS1 probe. This probe can theoretically detect the 47S/45S, 43S, 41S, and 30S RNAs, as well as the 30S-derived precursors of the mature 18S RNA. The results showed that whereas the 30S RNA was more abundant in the shKDM2B cells, the abundance of the 26S and 21S RNA species was not proportionally increased and, in fact,

was decreased (Figure 3.14A, B, C). Therefore, although the knockdown of KDM2B does not affect the ITS1/5.8S cleavage of the 45S RNA that gives rise to the 30S and 32S RNAs, it regulates the first 5' ETS cleavage, which gives rise to the 26S RNA. Interestingly, whereas the ratio of the abundance of the 26S and 30S RNAs was significantly higher in the control cells (Figure 3.14C, D), the ratio of the abundance of the 26S and 21S RNAs was not significantly different in the control and shKDM2B cells (Figure 3.14C, E), suggesting that knockdown of KDM2B impacts the first cleavage of the 5' ETS (A_0), but not the second cleavage (A_1), which gives rise to the 21S RNA.

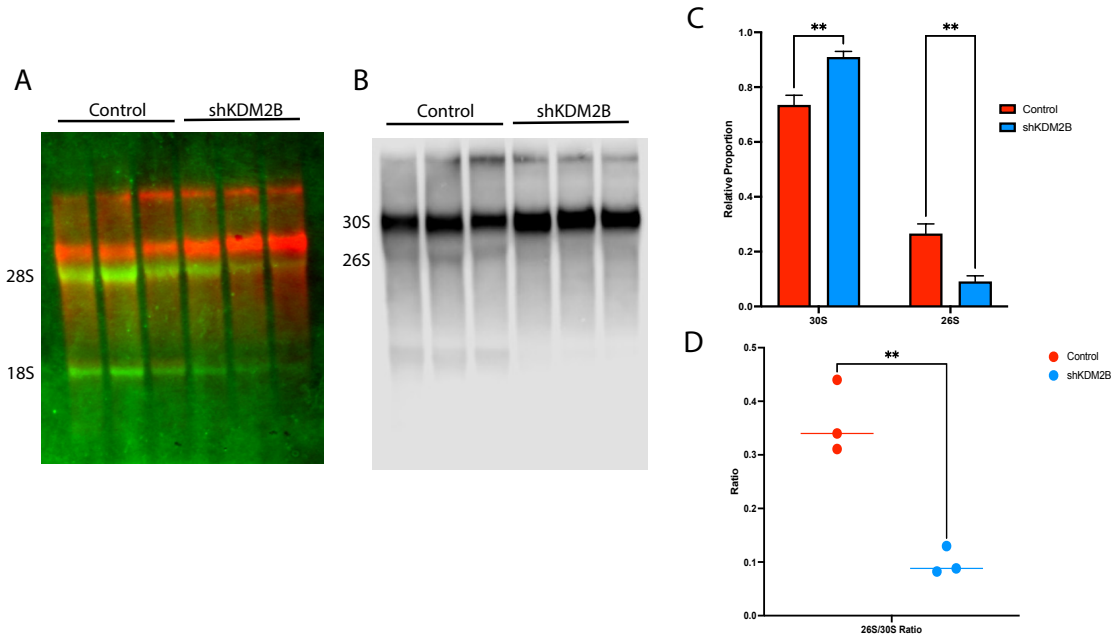


Figure 3.15 - 5'ETS Northern blot probe after knockdown of KDM2B.
A. Northern blot signal from lingering ethidium bromide stain (green at 600nm) and chemiluminescence probe signal (red at chemiluminescence signal). B. Northern blot signal from antibody recognition of dig-probe hybridization. C. Quantification of relative proportion of rRNA intermediates with ratio of D. 26S/30S. Statistics were calculated using multiple unpaired t-test correcting for multiple hypothesis testing by the Holm-Šídák method on Graphpad Prism; Error bars represent the standard deviation from the mean **p-value <0.01.

A second experiment, probing the same RNAs with the 5' ETS probe, confirmed the slightly higher abundance of the 30S RNA species and the defective processing of

30S to the 26S RNA species (Figure 3.15A, B, C, D). In this experiment, we did not detect the 21S RNA because it could not be detected with the 5' ETS probe.

The experiments in Figure 3.14 and Figure 3.15 showed that the most abundant RNA species detected with the ITS1 and 5' ETS probe was the 30S RNA. In the ITS1 probe, two significantly less abundant species migrating slower than the 30S RNA are likely the 43S and 41S RNAs. Their low abundance relative to the 30S RNA suggests that the ITS1/5.8S site is the most commonly first cleaved 45S RNA site in MDA-MB-231 cells as in pathway 2 as opposed to pathway 1 (Figure 3.13). Given that KDM2B does not affect the initial processing of the 47S/45S RNA, we conclude that KDM2B does not significantly alter the transcription of the rDNA arrays.

The preceding experiments confirmed that KDM2B promotes the processing of the 30S, which is the precursor of the 18S rRNA. To determine whether the ribosomal RNA processing leading to the mature 28S rRNA is also affected after the knockdown of KDM2B, we probed the same cell RNAs with the ITS2 probe, which should detect the 32S RNA and its downstream product 12S. The 32S RNA is produced by the same single 45S RNA cleavage event, which produces the 30S RNA. Given that this cleavage is not regulated by KDM2B, we would expect that the abundance of 32S in the shKDM2B cells would be equal to its abundance in the control cells. However, this is not the case, as the 32S RNA is slightly less abundant in the shKDM2B cells (Figure 3.16A, B). Although the absolute signal appears to be decreased, the relative abundance of 32S RNA is not statistically significant (Figure 3.16C). Even still, the processing of the 32S to the 12S RNA is affected by the loss of KDM2B, as demonstrated by the comparison of the 12S to

32S ratio in the control and shKDM2B cells, which was significantly decreased following knockdown (Figure 3.16D).

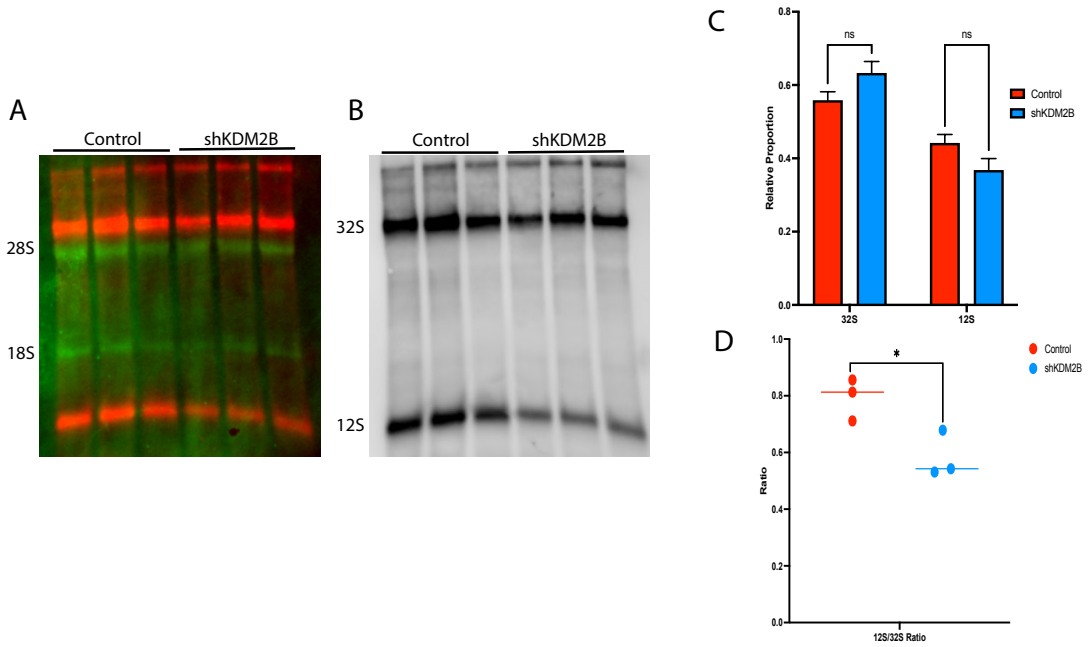


Figure 3.16 – ITS2 Northern blot probe after knockdown of KDM2B.
A. Northern blot signal from lingering ethidium bromide stain (green at 600nm) and chemiluminescence probe signal (red at chemiluminescence signal). B. Northern blot signal from antibody recognition of dig-probe hybridization. C. Quantification of relative proportion of rRNA intermediates with ratio of D. 12S/32S. Statistics were calculated using multiple unpaired t-test correcting for multiple hypothesis testing by the Holm-Šídák method on Graphpad Prism; Error bars represent the standard deviation from the mean *p-value < 0.05, ns not significant.

This observation suggested that the 32S RNA may be relatively unstable as the 30S from the ITS1 and 5' ETS probes (Figure 3.13, Figure 3.14) seemed to occur without problem, but there was slightly less 32S RNA following the loss of KDM2B. The instability of the 32S RNA is also supported by the Northern blot in Figure 3.17, which was probed with an RNA probe containing 3'ETS sequences, along with 20 bases from the 3' end of the mature 28S RNA. This shows that in shKDM2B cells, the 32S RNA is detected as a smear and not as a distinct band (Figure 3.17A, B). Additionally, comparing

the precursor 47S RNA to the overall signal where the 32S RNA suggested that early processing of the 47S to 32S is not affected by the loss of KDM2B (Figure 3.17C, D).

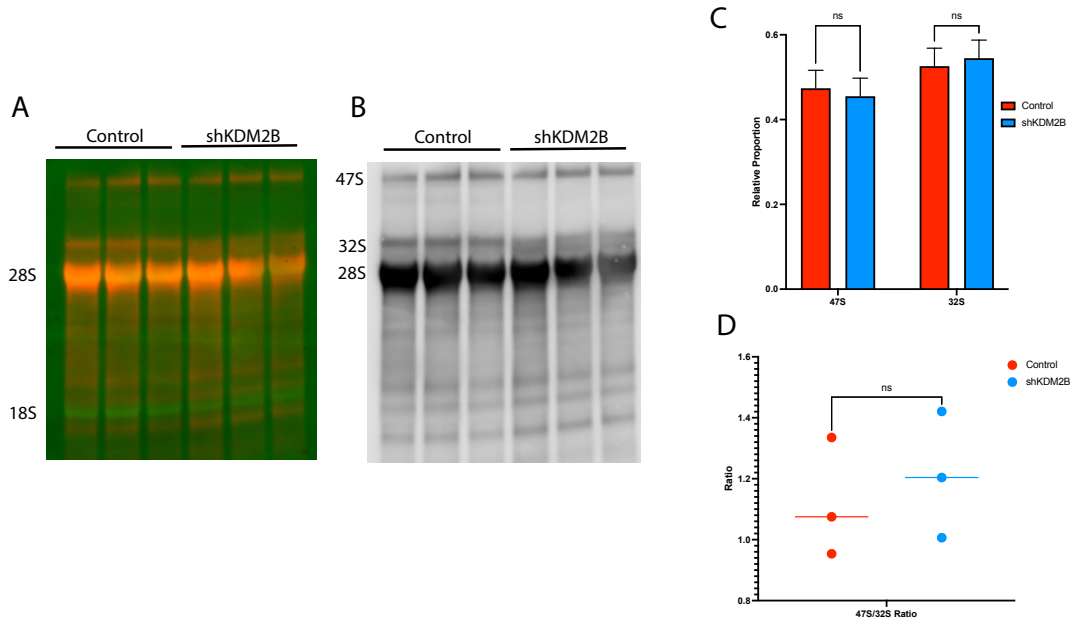


Figure 3.17 - 3'ETS Northern blot probe after knockdown of KDM2B.
A. Northern blot signal from lingering ethidium bromide stain (green at 600nm) and chemiluminescence probe signal (red at chemiluminescence signal). B. Northern blot signal from antibody recognition of dig-probe hybridization. C. Quantification of relative proportion of rRNA intermediates with ratio of D. 47S/32S. Statistics were calculated using multiple unpaired t-test correcting for multiple hypothesis testing by the Holm-Šídák method on Graphpad Prism; Error bars represent the standard deviation from the mean. ns not significant

The loss of KDM2B results in a defect that primarily affects the processing, leading to the 18S rRNA (Figure 3.14, Figure 3.15). This is in agreement with the ribosome profiling and quantitative proteomics data, which suggested that the loss of KDM2B gives rise to an imbalance between the 40S and the 60S ribosomal subunits. The potential destabilization of the 32S RNA (Figure 3.17A, B) may be a compensatory cellular mechanism aiming to minimize the imbalance. This agrees with recent studies showing that rRNA not yet incorporated into the assembled pre-ribosomal units may

undergo PAPD5-dependent polyadenylation, followed by 3'to 5' degradation by the exosome (*206*).

3.3.8. The knockdown of KDM2B inhibits global mRNA translation.

The observed defect in the biogenesis of the 40S ribosomal subunit, which is known to be required for translational initiation, suggested that the knockdown of KDM2B may inhibit the global rate of mRNA translation. This was also supported by the ribosome profiling experiment in Figure 3.9, which showed that the knockdown of KDM2B results in a decrease of the signal in the polysome fraction and a higher fraction of RNAs unbound from ribosomes. To address this hypothesis, we pulsed methionine-starved control and shKDM2B-transduced MDA-MB-231 cells with L-Homopropargylglycine (L-HPG), an amino-acid analog of methionine, which has been modified with an alkyne group. Nascent proteins labeled with the modified amino acid underwent a "click" chemistry to allow for the addition of Biotin-azide in order to track the nascent peptide production (Figure 2.3). The results showed that the knockdown of KDM2B decreases global translation dramatically. While the nascent translation was detected in the control cells by 7.5 minutes, it was first robustly detected in shKDM2B cells at 30 minutes (Figure 3.18A, B) and its rate lagged behind that of the control cells for the entire 3-hour period of monitoring (Figure 3.18C, D). Quantification of the data in A revealed that the knockdown of KDM2B induced at least a 60% decrease in the rate of translation at all time points (Figure 3.18A, D).

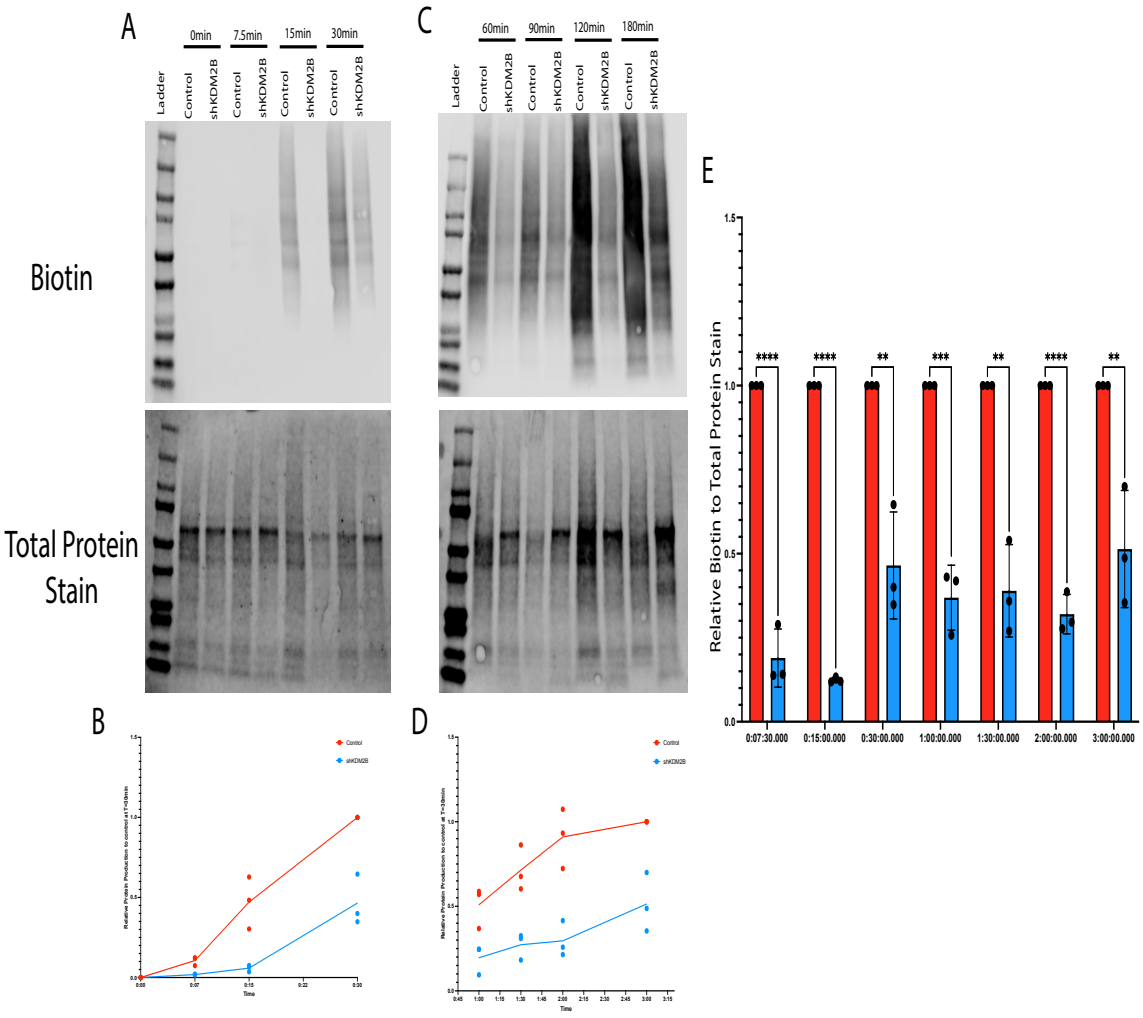


Figure 3.18 - KDM2B reduces global translational rates
A. Representative blot of biotin signal from antibody labeling nascent peptides from early timepoints with the corresponding Total Revert Protein Stain and quantified in across all 3 replicates in B. Normalized for the signal of the control in the last timepoint within each gel. C. Representative blot of biotin signal from antibody labeling nascent peptides from later timepoints with the corresponding Total Revert Protein Stain and quantified in across all 3 replicates in D. Normalized for the signal of the control in the last timepoint within each gel. E. Quantification of signal of shKDM2B at each time point relative to the control within each replicate. Statistics were calculated using multiple unpaired t-test correcting for multiple hypothesis testing by the Holm-Šídák method on Graphpad Prism; Error bars represent the standard deviation from the mean. **p-value <0.01, ***p-value <0.001, ****p-value <0.0001

3.3.9. The inhibition of the global translation rate in the shKDM2B cells is independent of the Integrated Stress and Ribotoxic Stress Responses

The decrease in the rate of the global mRNA translation in shKDM2B cells is likely to be caused by the low abundance of ribosomes. Additional factors that may contribute to the translational repression include the integrated stress response (ISR) and ribotoxic stress response (RSR), both of which can be activated by ribosome collisions occurring as a result of cellular stress (*207*). One of the factors inducing such stress is amino acid starvation, which may occur in shKDM2B-transduced cells because of defects in the biosynthesis of multiple non-essential amino acids (*208*).

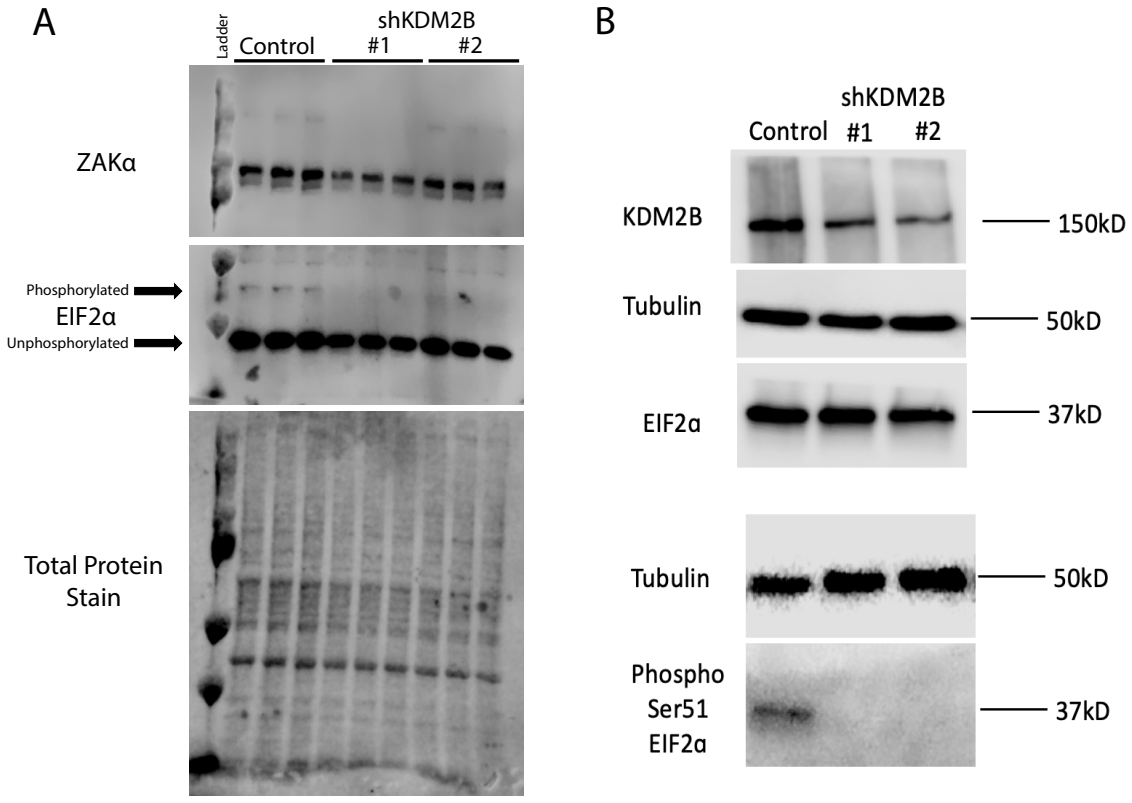


Figure 3.19 - KDM2B does not hyperactivate the ISR or RSR
A. SuperSep™ Phos-tag™ gel western blot with Total revert protein stain (Bottom) and probed with EIF2α (Middle) and ZAKα (Top) in 3 replicates of control or knockdown of KDM2B across two different shRNA constructs. Phosphorylation was labeled when it was able to be detected. B. Western blot of lysates extract from control or knockdown of KDM2B probing for total EIF2α or a phospho-Ser51 EIF2α specific antibody with corresponding controls.

Ribosome collisions induce phosphorylation and activation of the sentinel MAP3 Kinase ZAKα, which phosphorylates and activates the p38MAPK and JNK pathways, inducing apoptosis. ZAKα also activates GCN2, which phosphorylates eIF2α, a critical component of the Integrated Stress Response, at Ser51 (*130, 209-211*). Phosphorylation of eIF2α at this site increases its affinity to eIF2B and interferes with the eIF2B guanine nucleotide exchange activity. This inhibits the formation of the eIF2-GTP-tRNA (Met) ternary complex and translational initiation. To address this question, we analyzed control and shKDM2B-transduced MDA-MB-231 cell lysates for ZAKα phosphorylation (*207*). In addition, we probed the same lysates with antibodies that detect eIF2α phosphorylated at Ser51 or total eIF2α. These experiments failed to detect ZAKα phosphorylation (Figure 3.19A). Moreover, they showed that contrary to our initial expectation, shKDM2B transduction is associated with reduced phosphorylation of eIF2α (Figure 3.19A, B). Thus, we conclude that the shKDM2B-induced translational repression is independent of the activation of stress-induced ISR and RSR.

3.3.10. KDM2B knockdown dysregulates transcript specific translation.

To determine whether the global impairment of translation, observed following the knockdown of KDM2B, affects all mRNA transcripts equally or if some are more affected than others, we performed ribosome profiling in normally growing control and shKDM2B-transduced MDA-MB-231 cells. Deep sequencing of cDNA libraries constructed from ribosome-protected RNA fragments (RIBO-seq) (*212*) revealed that most signals map to mRNA coding regions and display clear periodicity, indicating that they accurately detect ribosomal footprints. In addition, the data showed a roughly even distribution of ribosome-protected fragments across mRNA coding regions, genome-

wide. Given that defects in translational elongation and translational termination result in a high density of ribosomes in the 5' end or the 3'end of mRNA coding regions, respectively, this finding suggests a defect in translational initiation, as it would be expected from the observation that the loss of KDM2B impairs the biogenesis of the 40S ribosomal subunit primarily (Figure **3.11**).

After quality control checks, the data was then interrogated for differential translation. Fold changes were calculated for the paired RIBO-seq and RNA-seq data. Analysis of the normalized fold change of ribosome-protected fragments detected by RIBO-seq with the normalized fold-change of mRNA reads in a paired RNA-seq

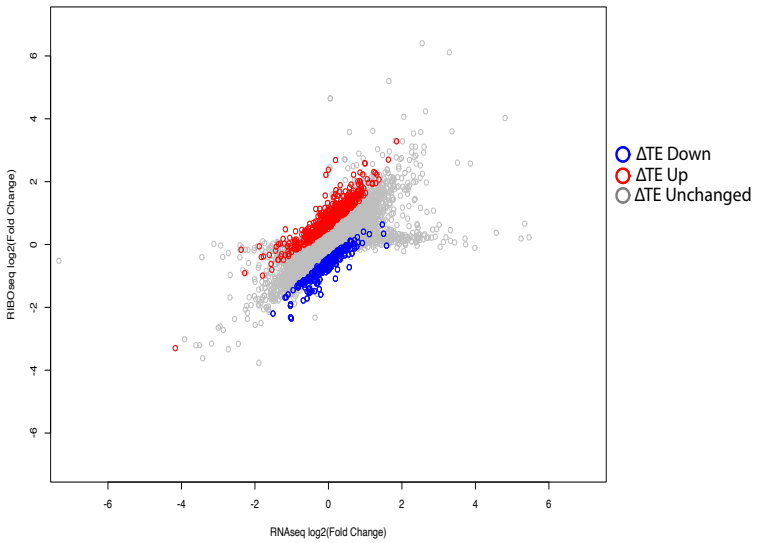


Figure 3.20- KDM2B impacts translational efficiency of multiple transcripts. Log2(Fold Change) comparing control and shKDM2B calculated for the paired RNA-seq (X-axis) and Ribo-seq (Y-axis) plotting all detected transcripts. Transcripts were labeled in red if they had increased translational efficiency and blue if they had decreased translational efficiency.

demonstrated high concordance (Figure 3.20). However, some select transcripts displayed discordance and highlighted uneven translational output when controlled for change in the RNA. Many of these mRNA transcripts were represented by a disproportionately high or disproportionately low number of ribosome-protected fragments, suggesting that the effects of the KDM2B knockdown on translational efficiency varied between transcripts (Figure 3.20).

We next sought out what mechanisms could explain the changes in translation efficiency of certain transcripts but not others. Differences in transcript specific translation could be the result of codon bias. Support for this hypothesis was provided by our observation that the knockdown of KDM2B resulted in uneven changes in the expression of tRNA amino acyltransferases and amino acid metabolism (*208*). Comparison of p-site abundance of ribosome-protected fragments at different amino acid codons in control and shKDM2B cells, however, failed to reveal statistically significant KDM2B-dependent changes (Figure 3.21A, B). We conclude that the effects of the KDM2B knockdown on transcript-specific translation are not due to bias in codon usage.

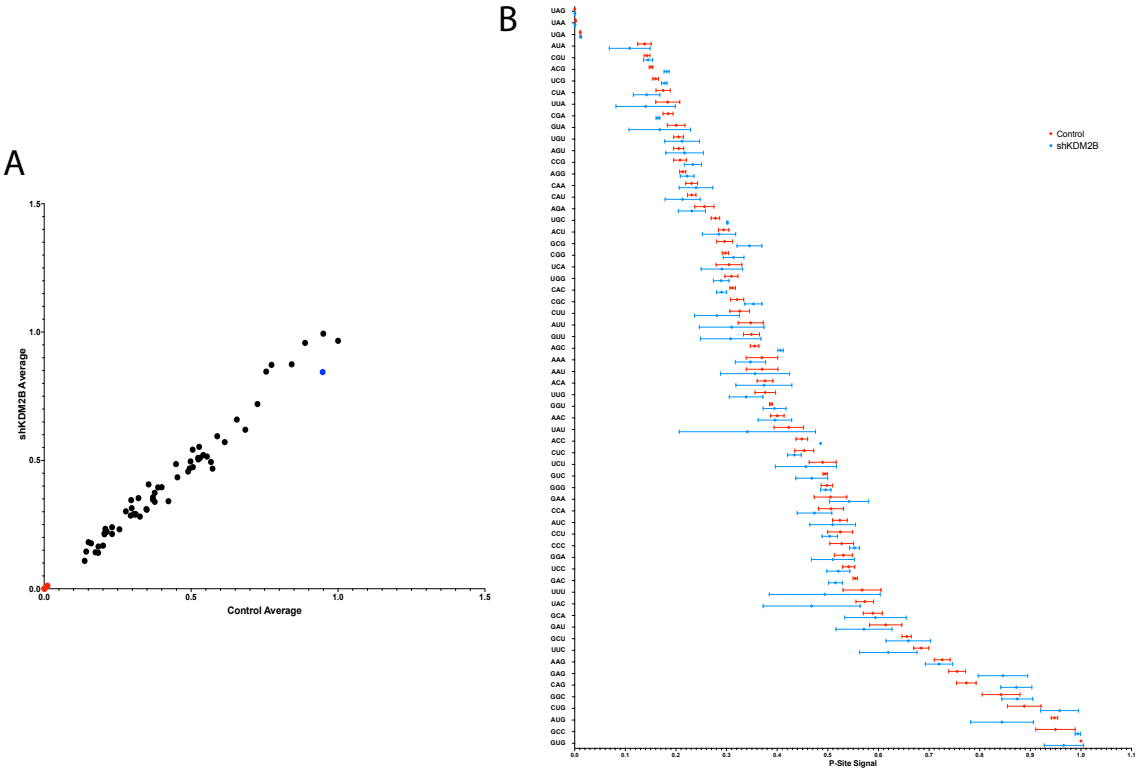


Figure 3.21 - Knockdown of KDM2B does not alter codon usage
A. P-site signal from each codon plotted onto a dot plot. Blue; start codon AUG. Red; codons UAA, UGA, and UAG. B. Individual codon p-site signal plotted with the mean and standard deviation. Statistics were calculated with multiple unpaired t-tests correcting for multiple hypothesis testing with Bonferroni-Dunn method.

The efficiency of mRNA translation is known to depend on structural motifs that are located primarily in its 5'UTR (*75, 192, 213*). We therefore questioned whether the structure and base composition of the mRNA 5' UTR influences the effects of KDM2B on translational efficiency. This question was addressed with the analyses in Figure 3.22, which show that transcripts with shorter, unstructured, AU-rich 5' UTRs exhibit decreased translational efficiency in shKDM2B-transduced cells (Figure 3.22A, B, C). In contrast, the translational efficiency of transcripts with longer, structured, GC-rich 5' UTRs remains unchanged or is increased (Figure 3.22A, D, E). Given that global translation is downregulated in these cells, this observation may also mean that the decrease in the translational efficiency of mRNAs with long, complex 5'UTRs is less pronounced than the decrease in the translational efficiency of mRNAs with less complex 5' UTRs. Stated differently, the translational initiation of mRNAs with complex 5' UTRs is protected because it is independent of the knockdown KDM2B. Additionally, as open reading frame (ORF) length is also known to play a role in translational efficiency, we investigated the length of the coding sequences (CDS) of each transcript. We found that transcripts with longer CDS were more likely to be minimally affected or translationally protected by loss of KDM2B. In contrast, those with shorter CDS were more likely to be translationally repressed and more potently affected by loss of KDM2B (Figure 3.22F).

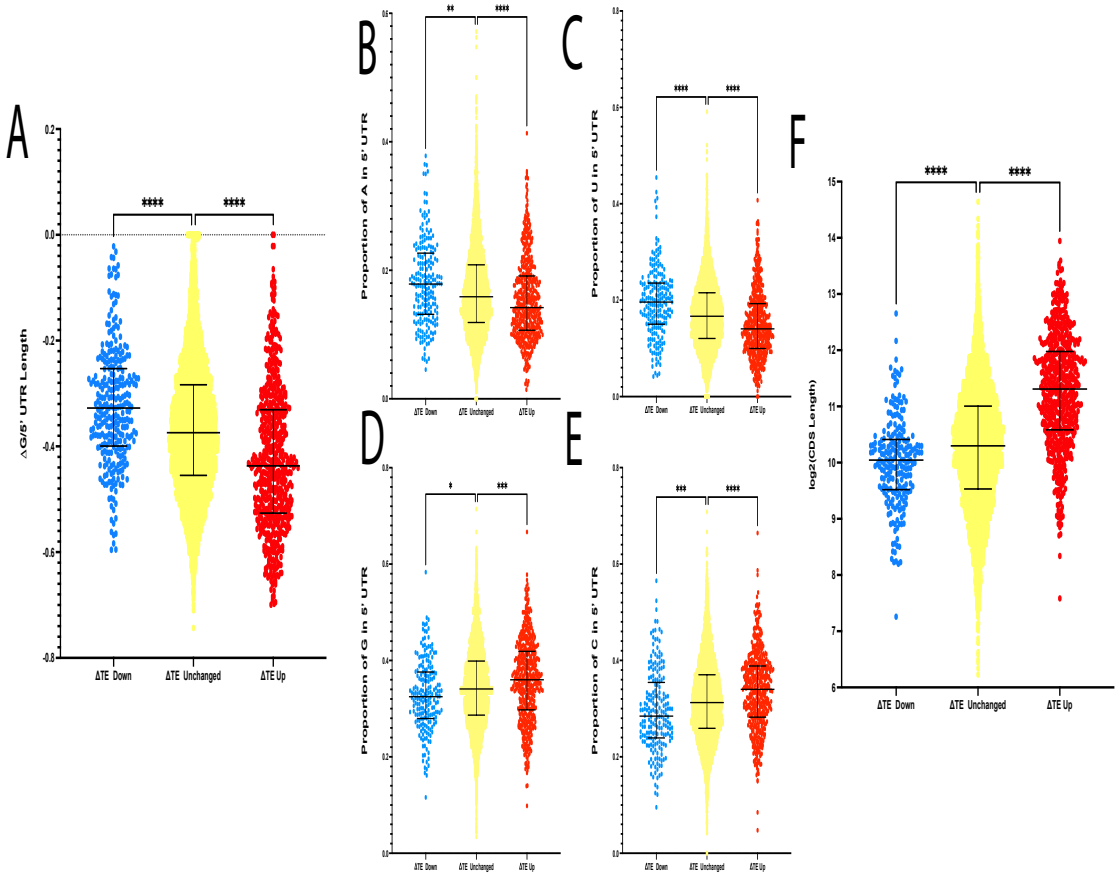


Figure 3.22 - KDM2B translationally regulated transcripts have differences in 5' UTR composition and CDS length.
Transcripts characterized as significantly decreased translational efficiency in blue (log2(FC)<-0.5), increased translational efficiency (log2(FC)>0.5), or unchanged were plotted for A. 5' UTR structuredness, B. 5' UTR adenine composition, C. 5' UTR adenine composition, D. 5' UTR guanosine composition, E. 5' UTR cytosine composition, and E. CDS length. Statistics were calculated with Kruskal-Wallis test followed by a Dunn's multiple comparisons test. *p-value < 0.05, **p-value <0.01, ***p-value <0.001

Investigating the translationally regulated transcripts with an enrichment analysis, we find significant gene ontologies in transcripts displaying decreased translational efficiency after the knockdown of KDM2B (Figure 3.23A) and transcripts displaying increased translational efficiency transcripts after the knockdown of KDM2B (Figure 3.23B). As global translation is decreased following knockdown (Figure 3.18), we focused on transcripts with decreased translational efficiency as they represent transcripts

in which the protein steady state would be affected more potently in the context of overall translational repression. Genes involved in the G2/M phases of the cycle, such as CDK1 and CCNB1, demonstrated significant decreases in translational efficacy from the paired RNAseq/RIBOseq and were confirmed to be decreased at the protein level (Figure 3.23C). We next hypothesized that decreased translation and steady-state protein levels of these transcripts would coincide with a reduced percentage of cells in the S/G2 phases of the cell cycle. Using propidium iodide staining and flow cytometry, we found that the knockdown of KDM2B led to a significant increase in the proportion of cells in the G1 phase of the cell cycle with a significant decrease of cells in the S and G2 phases of the cell cycle (Figure 3.23D).

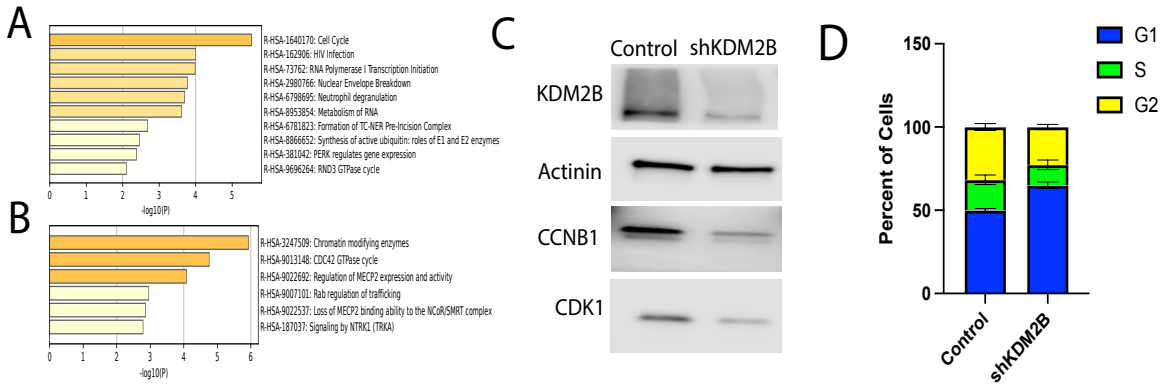


Figure 3.23 - KDM2B alters the translation of cell-cycle related transcripts. Metascape enrichment analysis of transcripts that were A. translationally downregulated or B. translationally upregulated relative to all transcripts detected by the Ribo-seq. C. Western blot of lysates extract from control or knockdown of KDM2B CDK1 and CCNB1 with corresponding controls. D. Propidium iodine quantification of cell cycle as percentages of G1, S, or G2 phases of the cell cycle. 3 replicates were used for each biological condition. Mean and standard deviation are presented for each cell cycle phase.

3.3.11. KDM2B does not alter the overall ribosomal composition.

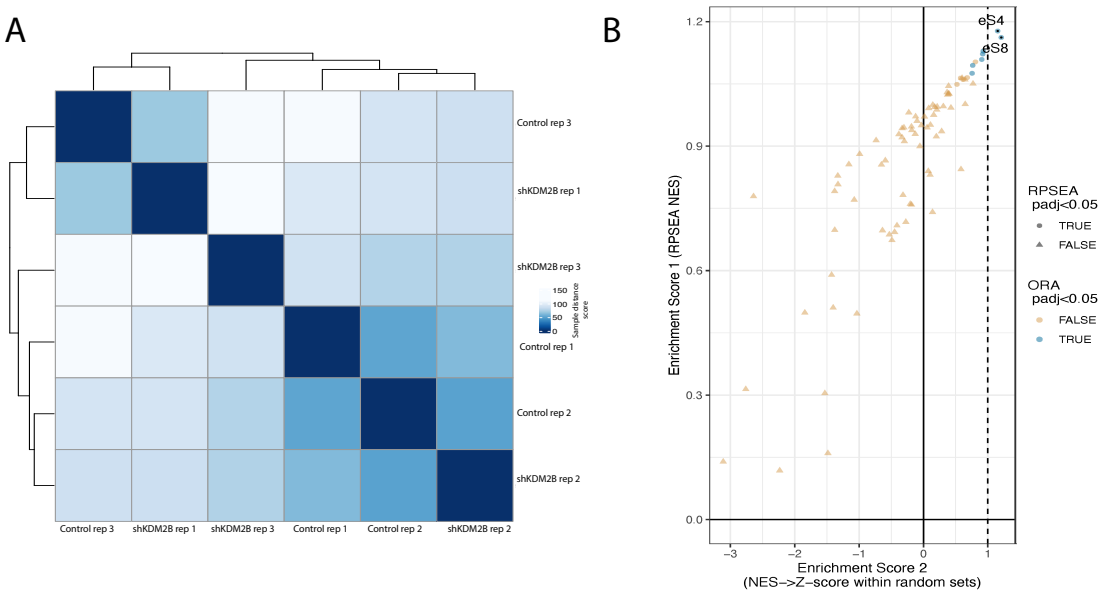


Figure 3.24 - KDM2B does not significantly alter ribosomal composition. Ribosomal heterogeneity analysis by dripARF performed on rRNA reads from Ribo-seq demonstrated by A. Matrix where individual samples undergo hierarchal clustered based on calculated distance between samples by principal component analysis. B. Mapping of enrichment scores comparing rRNA between control and KDM2B for individual ribosomal proteins across two enrichment score calculations.

Finally, while KDM2B potently impacts ribosomal abundance (Figure 3.9), it may also be impacting ribosomal composition by specifically altering the levels of some subunits over others, leading to ribosomes lacking specific ribosomal proteins (*23, 34, 129, 131, 189*). While no ribosomal protein appeared to be an outlier as compared to the other members of the same ribosomal subunit (40S or 60S) (Figure 3.11, Figure 3.12), we could not exclude the effect of ribosomal heterogeneity in differential translation. Ribosomal heterogeneity analysis was performed based on patterns of rRNA fragments generated during RNAse I treatment to generate ribosome-protected fragments for Ribo-seq (*189*). This analysis revealed that knockdown of KDM2B does not have a unique ribosomal heterogeneity signature compared to the control. Further supporting this is that samples did not display hierarchal clustering into their respective groups (Control and

shKDM2B), demonstrating that there are minimal differences in ribosomal heterogeneity (Figure 3.24A). While RPS4X and RPS8 are just over the enrichment threshold and deemed statistically significantly altered between control and shKDM2B samples (Figure 3.24A), they do not appear to be the most altered based on proteomics analysis (Figure 3.11, Figure 3.12A). Given that previously reported heterogeneity analysis had a much larger magnitude of change in enrichment scores when determining true heterogeneity (*189*), it is likely these two proteins are not representative of real heterogeneity between control and knockdown of KDM2B. Given that knockdown of KDM2B leads to a defect in 18S rRNA processing of the small ribosomal subunit and that these proteins are both small ribosomal subunit proteins, the signal most likely is from rRNA availability due to a processing defect (Figure 3.14) as opposed to true heterogeneity.

3.4. Acknowledgements

We want to thank Michael G. Kearse at The Ohio State University for his time and help with teaching polysome profiling. We would also like to thank Giovanni Nigita for his feedback on the analysis of the RNA-seq and Ribo-seq data. We also thank all members, past and present, of the Tsichlis laboratory for their helpful feedback on the project.

3.5. Contributions

Vollter Anastas served as the lead for formal analysis, investigation, software, visualization, and preparation of the manuscript, as well as project conceptualization. Evangelia Chavdoula supported planning and performing experiments, taking the lead on performing the ChIP-sequencing library preparation and optimization, working collaboratively to perform the ATAC-sequencing, and helping prepare the cell culture for

multiple experiments. Alessandro La Ferlita collaborated on the analysis concept design for the RNA-seq, Ribo-seq, ChIP-seq, and ATAC-seq and performed the bioinformatics analysis on said projects, generating multiple datasets in which to interrogate. Michael G. Kearse performed and aided in polysome profiling and offered feedback on multiple abstracts submitted as a result of this project. Philip N. Tsichlis supervised the project, supported project conceptualization, and performed manuscript reviewing and editing.

Chapter 4: Discussion

4.1. Summary and Implication

In summary, KDM2B appears to act as an upstream transcriptional regulator of ribosomal biogenesis. Knockdown of *KDM2B* reduces ribosome production and ribosomal proteins globally. This, in turn, leads to a decreased abundance of ribosomal proteins and a corresponding decrease in translation. Curiously, the 40S ribosomal subunit appears to be more dramatically affected relative to the 60S ribosomal subunit. As rRNA components of the 40S and 60S ribosomal subunits are transcribed from a single entity, which is then processed, this led us to hypothesize that this was due to a defect in processing. Although we noted defects in both 18S and 28S rRNA processing, the primary defect appears to be in the 18S rRNA, specifically the 30S to 26S rRNA processing. Finally, in order to characterize which transcripts were more sensitive to the global translation repression versus which ones were more resistant, we performed Ribo-seq to characterize changes in translational efficiency across all detected protein-encoding transcripts. From this analysis, we identified that transcripts with long and structured 5' UTRs are more likely to be resistant to global translational repression induced by the knockdown of KDM2B. In contrast, transcripts with short and unstructured 5' UTRs are more likely to be sensitive to loss of KDM2B. Categorically, transcripts that were more sensitive (decreased translational efficiency) were more likely to be cell-cycle regulators, particularly those involved in later-stage transitions in the cell cycle (S/G2/M). Our results suggest that KDM2B promotes an oncogenic program altering ribosome biogenesis, rRNA processing, ribosomal abundance, protein output, and cell growth (Figure 4.1).

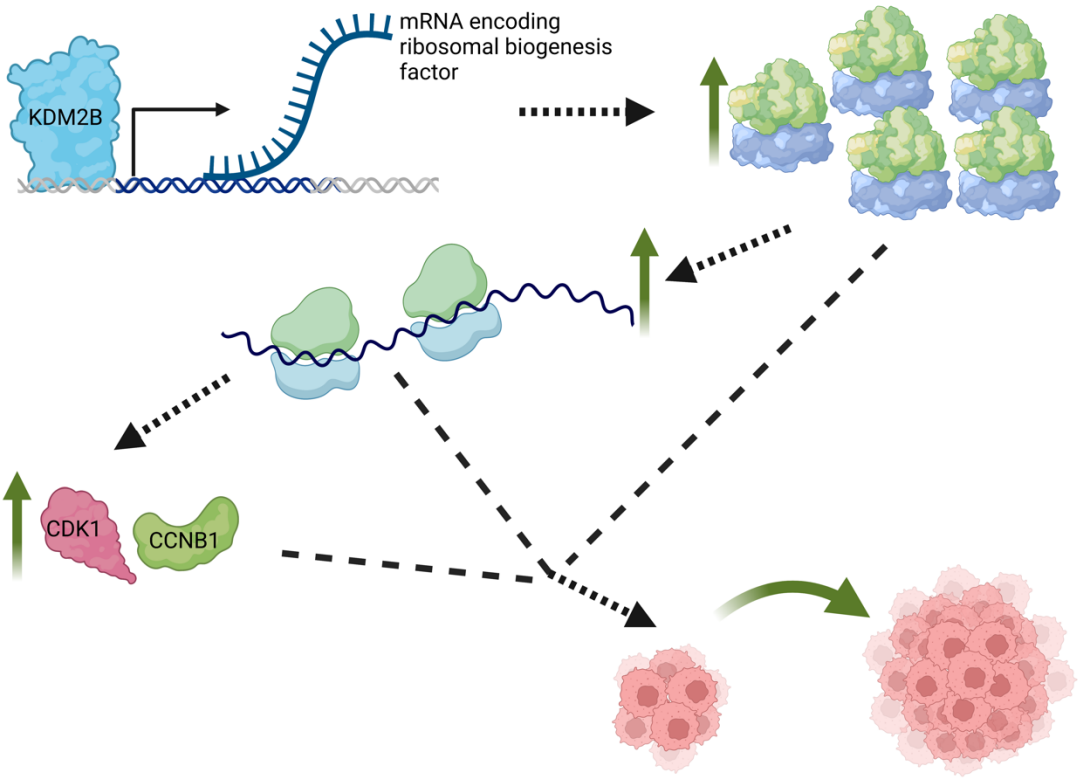


Figure 4.1 - Overall model of KDM2B regulation of ribosomal biogenesis.
KDM2B directly regulates ribosomal biogenesis gene transcripts by binding directly binding to their promoter. This in turn leads to upregulation of key ribosomal biogenesis genes that promote a higher ribosomal abundance and thus protein production. These changes also lead to changes in translation efficiency of key cell cycle regulates (CDK1, CCNB1). All of these effects (increased ribosomal abundance, increased protein translation, increased levels of cell-cycle promoting factors) all promote increased cancer cell growth.

Intriguingly, basal-like breast cancer is characterized by *TP53* mutations and *c-MYC* copy number increases (*166, 214*), which contribute to ribosomal dysregulation through loss of the nucleolar stress response via mutations of *TP53* and increased ribosomal production via *c-MYC* increased expression (see Section 1.3). In line with this, basal-like breast cancer cells are characterized by having cellular architecture with very prominent nucleoli relative to the nucleus and very scant cytoplasm (*215, 216*). It is in

this cancer subtype that we identified that KDM2B is overexpressed. Expectedly, the basal-like breast cancer subgroup displays a transcriptome signature of high ribosomal biogenesis and increased levels of *c-MYC* (*162, 166*). Within this subgroup, KDM2B displays a significant positive correlation between its levels and those of ribosomal biogenesis factors both at the protein and RNA level (Figure 3.8), suggesting that our experimental data in MDA-MB-231 cell lines might be representative of KDM2B's transcriptional regulation in patient breast cancer, particularly, in basal-like breast cancer (*214*).

4.2. Mechanism of the regulation by KDM2B

KDM2B is a complex multi-domain protein that has multiple reported modalities of gene expression regulation. It has been reported to function in histone demethylation targeting H3K4me3 (*112*), H3K79me (*91, 217*), and H3K36me1/2 (*89, 123, 218*). Additionally, it has been reported to alter chromatin status in the cell by recruiting members of the SWI/SNF complex, leading to open chromatin (*94, 197*) and also by binding to CpG islands and preventing the activity of DNA methyltransferases, thus rendering the loci more transcriptionally active (*94, 103*). While all these are potential functions of KDM2B, its best-described function is that through the PRC1.1 complex, acting as the DNA binding agent recruiting the complex to target loci (*86, 92-94*) (See Section 1.4).

The majority of the early literature on KDM2B primarily focused on its catalytic activity as a histone demethylase (*89, 91, 112, 123, 217, 218*); however, recent literature highlights that KDM2B has very poor histone demethylase function with KDM2B requiring very high concentrations to even minimally demethylate its canonical target

H3K36me1/2 (*219*). Notably, KDM2B's paralog, KDM2A, maintains its ability to demethylate H3K36me1/2 and is able to canonically bind to the acidic patch of the histone protein as recognition of the mark, which KDM2B is not able to do efficiently, (*219*). This seems to be modulated by the N-termini of each protein where the KDM2B N-terminus placed onto KDM2A renders KDM2A ineffectual as a demethylase, and the KDM2A N-terminus placed onto KDM2B grants it a gain-of-function in its demethylase activity (*219*). This illustrates that KDM2B has intrinsic demethylase activity but that the N-terminus renders it null. It is plausible that KDM2B demethylase activity is context-dependent and that under specific cellular conditions, KDM2B might be able to act as a demethylase; however, KDM2B is overall a poorly expressed protein. In HEK293 cells, KDM2B has a concentration of 8.1nM, whereas its catalytically active paralog, KDM2A, is expressed at 120nM. Given that KDM2B requires high concentrations to have some activity (*219*), cellular conditions would need to be heavily skewed to benefit KDM2B's functional capacity to exhibit demethylase activity (*173*). In line with this recent structural and biochemical insight into KDM2B, multiple groups have failed to identify global changes in H3K36me1/2, H3K79me, or H3K4me3 after gain or loss of KDM2B (*93, 220*). In regard to our data, H3K36me1/2, H3K79me, or H3K4me3 are active marks; therefore, if KDM2B's central regulatory role was demethylation of these active marks, its target loci would be genes that it repressed, and consequently those genes would become upregulated when KDM2B was knocked down. Ribosomal biogenesis genes, on the other hand, are downregulated after loss of KDM2B and KDM2B directly bound to the promotors of ribosomal genes. Thus, even if KDM2B contained demethylase activity, it would not explain the effect we noted on ribosomal biogenesis genes.

Changes in chromatin status offered a promising explanation as to how KDM2B was potentially promoting gene expression (*94, 197*). However, we failed to detect any large changes in chromatin status. It is unlikely that ribosomal genes would display completely closed chromatin states following the knockdown of KDM2B as they are essential genes. However, we hypothesized upstream factors might have their expression regulated by changes in accessibility induced by KDM2B. While we detected MYC as a potential target, the most significant changes we detected were in poorly expressed genes (Figure 3.6). This is in line with other literature where inhibition of the SWI/SNF complex initially led to a global decrease in accessibility across all genes as one would expect, but targets that are well expressed are able to compensate through alternative mechanisms, whereas poorly expressed targets remain at closed chromatin state without compensation (*221*). It may be that in the acute phases of KDM2B knockdown, there is global accessibility loss that then partially recovers, but given the nature of shRNA and knockdown experiments, we are not able to capture acute phase effects.

MYC represented a promising target where KDM2B had strong binding based on the ChIP-seq (Figure 3.4B), and MYC's promotor displayed a loss of accessibility based on the ATAC-seq (Figure 3.6A). Thus, we postulated that KDM2B's main regulatory function was to impact MYC, and MYC then exerted positive downstream effects on ribosomal biogenesis targets. Experiments to evaluate the sufficiency of MYC alone, in the absence of KDM2B, to rescue ribosomal biogenesis were set up; however, MYC overexpression was able to overcome the shRNA-induced knockdown of KDM2B and increase KDM2B expression to levels similar to control. Thus, we were not able to accurately test the sufficiency of MYC to drive ribosomal biogenesis without KDM2B.

Additionally, similarly, as the knockdown of KDM2B reduces the level of MYC, the knockdown of MYC also reduces the levels of KDM2B (data not shown), suggesting that these two co-regulate one another, making it technically challenging to interrogate an epistatic relationship between MYC and KDM2B.

Another possible explanation for KDM2B's regulation of ribosomal genes was the PRC1 complex. The PRC1, in breast cancer, is oncogenic and promotes cancer growth; however, the oncogenic targets of the PRC are not the same in different subtypes of breast cancer (*98*). This suggests that the regulation of the PRC is context dependent. RING1B, the common factor to all PRC1 complexes, demonstrated that MYC target genes such as ribosomal biogenesis factors were selectively downregulated in the triple-negative breast cancer (TNBC) cell line, MDA-MB-231 but not in the luminal breast cancer cell line, T47D (*98*). Additionally, the group identified that while the canonical PRC1 binding partners (PCGF2/4) appear to strongly bind to RING1B in both ER+ and TNBC cell lines, RING1B additionally binds to noncanonical PRC1 components, specifically the PRC1.1 components (KDM2B, PCGF1) in the TNBC cell line (*98*). This suggests that the reason the knockdown of RING1B affects different genes in different cell lines is that it participates in different complexes. Specifically, it is more likely to participate in the PRC1.1 complex along with KDM2B than in the PRC1.2 or PRC1.4 complex in MDA-MB-231. This is likely why MYC target genes are repressed following RING1B knockdown in MDA-MB-231 cell lines but not in the T47D cell line. Based on our data in MDA-MB-231, the knockdown of KDM2B leads to the downregulation of MYC targets, as did the knockdown of RING1B in their data, making it more plausible that RING1B and KDM2B are acting as part of the PRC1.1 complex to lead to the

regulation of ribosomal biogenesis factors. As KDM2B is the necessary genomic targeting component of the PRC1.1 and our data has shown that KDM2B directly binds to ribosomal biogenesis gene targets in MDA-MB-231 (Figure 3.4), this suggests that ribosomal biogenesis is regulated through the PRC1.1 via direct binding from KDM2B. However, targets being PRC1.1 or PRC1.2/1.4 exclusive have not been elucidated.

Finally, KDM2B mainly binds to unmethylated CpG islands, which tend to be enriched in housekeeping genes (*222*). We were able to detect a similar binding pattern for KDM2B, suggesting it may play a role in the maintenance of housekeeping gene expression like ribosomal biogenesis targets. Previous reports have demonstrated that KDM2B protects CpG islands at genes from DNA methylation, thus preventing target genes from being repressed (*103, 104*). It is unlikely that CpG islands near housekeeping genes will be methylated in immortalized cell lines; however, this mechanism offers another hypothesis about how KDM2B might be maintaining active transcription at these sites.

4.3. KDM2B as an oncogene

We have shown that KDM2B in basal-like breast cancer promotes cell growth by maintaining the expression of ribosomal biogenesis factors and increasing translational output. One question that remains is whether or not KDM2B can be classified as an oncogene. It certainly has pro-survival and growth functions, but whether it is sufficient to drive cancer remains unknown. Defining an oncogene as a "gene that is capable of transforming cells in culture or inducing cancer in animals" (*223*) makes it challenging to classify KDM2B as an oncogene with the current presented data. However, KDM2B's initial discovery, where it was named NDY1 for Not-Dead-Yet 1, showed that KDM2B

was a common integration site of the Moloney Murine Leukemia Virus (MoMuLV) provirus integration in MoMuLV-induced rat T-cell lymphomas and that it was able to immortalize fibroblasts through overexpression (*89*). This suggests that KDM2B is sufficient to drive certain cancer types and to induce a critical transforming aspect of cancer biology, loss of replicative senescence. This line of data suggests that KDM2B could be classified as an oncogene and that it is capable of transforming cells to lose senescence. However, pan-cancer analysis found that only 3% of cancer samples displayed genomic alterations in KDM2B (*224*). While the majority of these alterations were amplifications suggesting that KDM2B was oncogenic, the fact that genomic alterations were rare makes it unlikely for KDM2B to be an initial driver of cancer. It is more likely that KDM2B serves as a pro-survival and growth phenotype and is important, at least in basal-like breast cancer, for general growth. Therefore, its upregulation in that subgroup serves more as a means to an end to increase ribosomal biogenesis further. Thus, KDM2B is likely not an oncogene capable of driving cancer itself but more of a key factor that is required for efficient cancer growth. However, given the initial discovery of KDM2B/NDY1 (*89*), we cannot completely exclude the possibility of oncogenicity without transformation assays or mouse models to definitively label KDM2B as an oncogene or not.

4.4. Alterations of ribosomal abundance and processing

Ribosome assembly is a complex energy-demanding process, which is initiated in the nucleolus with the RNA Polymerase I mediated synthesis of the 47S pre-ribosomal RNA (*28*). Recruitment of a large number of ribosomal proteins (RPs) and ribosome biogenesis factors (RBFs) to the nascent pre-rRNA gives rise to the 90S early pre-

ribosomal particles. Concurrently, the pre-rRNA undergoes cleavage at multiple sites, which gives rise to the precursors of the 18S, 5.8S, and 28S ribosomal RNAs, which are then processed to the mature rRNA forms in the 40S and 60S processomes (SSU and LSU respectively). Recent cryo-EM studies revealed some of the complex structural changes occurring as the SSU processome transitions from a stage prior to the pre-ribosomal RNA A1 cleavage (pre-A1) to the stage that follows it (post-A1) (*225, 226*). Based on these studies, several critical molecules that may contribute to the shKDM2B-induced processing defects could be identified. One of them is the small nucleolar RNA (snoRNA) U3, which base-pairs with the 5' ETS of the pre-ribosomal RNA and the 18S rRNA and guides the formation of the central pseudoknot, an essential conserved feature of the 40S processome. Another critical snoRNA is the U17 RNA, which binds UTP23, a putative endonuclease that may cleave pre-rRNA at the A0 site and may contribute to the A1 and the A2a pre-rRNA cleavage. In addition to the endonucleases, UTP23 and UTP24, other proteins of significant interest include UTP4 and WDR75, which bind the evolutionarily conserved motif (ECM) at the extreme 5' end of the pre-rRNA after the A' cleavage, UTP14A and RRP9, which bind the U3 snoRNA, the chaperones HEATR1 and UTP20, and the ribosomal proteins RPS24, RPS19 and RPS4X. Many of these proteins involved in the 30S processing into the 26S are downregulated (Figure 3.2, Figure 3.11), including some ribosomal proteins that are directly involved in the processing. While one keystone gene regulatory event may trigger inefficient processing and lead to ribosomal imbalance, it is more likely that a long-lived intermediate (30S pre-rRNA intermediate) (Figure 3.14, Figure 3.15) requires multiple factors in its processing regulated by KDM2B. This is also demonstrated in a comprehensive analysis where individual

nucleolar factors were globally investigated for defects in processing (*227*). Knockdown of 625 individual nucleolar factors identified by mass spec and their impact on rRNA processing based on northern blot probing revealed that 286 factors led to at least a >1.5 absolute fold change in an rRNA intermediate (*227*). Out of these factors, a total of 72 (~25%) had a similar phenotype as noted by knockdown of KDM2B with accumulation of the 30S rRNA. This suggests that many factors participate in 30S processing, and thus, it is unlikely that the downregulation of a single factor is responsible for the phenotype noticed.

4.5. Decreased levels of eIF2α

The reason for the decreased phosphorylation of eIF2α in shKDM2B cells is currently unknown. One likely possibility is that cancer cells have higher amounts of proteotoxic stress due to enhanced mRNA translation, which induces proteotoxic ER stress and the activation of PERK, an eIF2α kinase. When global mRNA translation is repressed following the knockdown of KDM2B, this proteotoxic stress may be relieved, leading to decreased PERK activation and eIF2α phosphorylation. PERK is a type I ER membrane protein, which consists of luminal, transmembrane, and cytoplasmic domains. Under physiological conditions, glucose-regulated protein 78 (GRP78), a member of the HSP70 heat shock protein family, binds the PERK luminal domain and maintains PERK in an inactive state. During ER stress, GRP78 is released to bind misfolded proteins accumulating in the ER, and the GRP78-free PERK undergoes dimerization, activation, and autophosphorylation (*228*). Given that loss of KDM2B repressed translation, there is likely less misfolded protein in the context of decreased translational demand. Notably, work by others in the lab demonstrated while bortezomib treatment was able to induce

eIF2α phosphorylation in control cells, it was unable to do so in shKDM2B-transduced cells (*208*). This suggests that blocking the proteasome is insufficient to induce cellular stress after the knockdown of KDM2B. This is likely because there is less translation globally and, therefore, slower accumulation of proteotoxic stress even in the presence of bortezomib.

4.6. Ribosomes and Cell Cycle

Our RIBOseq analysis revealed that the translation of genes involved in later phases (S/G2/M) had translational efficiency after the knockdown of KDM2B. This coincided with a higher proportion of cells in the G1 phase of the cell cycle. Recently, a proteomic characterization comparing phases of the cell cycle was performed (*229*). The group excluded known cell cycle regulators from their analysis to characterize what processes were occurring outside of known cell cycle progression. The data demonstrated that ribosomal biogenesis appears to mainly occur in the G1 phase of the cell cycle relative to other phases (*229*). Naturally, for a cell to undergo symmetric division, it would have to produce roughly twice the number of ribosomes that it had at the start of G1 by the time it was undergoing mitosis in order to maintain proper ribosomal abundance. This steady increase of ribosomal amount, based on proteomic analysis (*229*), appears to mainly occur during G1. In addition, it seems that knocking down individual ribosomal subunits (*230, 231*) leads to global decreases in overall ribosomal abundance. Thus, it is not unexpected that the knockdown of individual ribosomal proteins also leads to cells spending a higher proportion of their time in the G1 phase of the cell cycle (*230, 231*). Overall, this suggests that ribosome biogenesis occurs mainly in G1 and that decreases in ribosomal proteins or ribosomal abundance leads to stalling in the G1 phase

of the cell cycle. This implies the existence of crosstalk (*232*), where cells are able to discern the number of ribosomes present in a period of time and respond by extending or shortening G1 to the right amount and maintaining ribosomal homeostasis during divisions. Recent evidence from single-cell Ribo-sequencing data demonstrated that during mitosis, there are changes in the translational efficiency of transcripts along the continuum of the cell cycle (*16*). Thus, we postulate that ribosomal abundance changes alter the translational efficiency of cell-cycle related transcripts and, therefore, alter their final protein output. This, in turn, leads to changes in the timing of the cell cycle and provides an elegant mechanism for the cell cycle to be sensitive to ribosomal abundance. Our data on the knockdown of KDM2B mimics the data of others (*230, 231*) that knocked down individual ribosomal proteins. Both experiments led to decreased ribosomal abundance, albeit through different mechanisms, but converged on identifying cell-cycle regulators as translationally affected and cells stalling on the G1 phase on the cell-cycle. This suggests that our cell cycle translational efficiency changes and G1 accumulation of cells following the knockdown of KDM2B is due to decreased ribosomal abundance.

Further evidence of this is provided by the fact that the translational efficiency of many transcripts is altered during phases of the cell cycle, with shorter transcripts and GC% poor being more translated during early phases of the cell cycle as opposed to later phases (*233*). We notice a downregulation of such transcripts following the knockdown of KDM2B despite the cells accumulating more in the G1 phase of the cell cycle. This suggests that simply being in G1 does not increase the translation of these transcripts. More likely, these transcripts are required to have high translational efficiency during G1

in order to progress to later stages. One hypothesis is that ribosomal abundance promotes the translation of these transcripts, allowing the cell to progress to later stages. The short, GC% poor 5' UTR may be sensitive to ribosomal levels where high ribosomal abundance is required to translate such transcripts efficiently, thus providing an intrinsic sensor of ribosomal abundance to progress into the cell cycle. We provide evidence for this with the fact that the knockdown of KDM2B reduces the translation of GC% poor, short transcripts in the context of decreased ribosomal abundance. Additionally, this highlights that the accumulation of cells in the G1, following knockdown of KDM2B, is not the primary defect but rather a secondary defect as a consequence of alterations in ribosomal abundance leading to a decreased translational efficiency of GC% poor, short transcripts.

4.7. How can ribosomal concentration affect transcript-specific translation?

The finding that KDM2B reduced the translational efficiency of short transcripts over long transcripts is counterintuitive. We would have expected that the mRNAs most sensitive to translational inhibition would be the mRNAs with long complex 5'UTRs that can interfere with ribosome scanning mechanisms and translational initiation. The fact that the opposite is observed suggests that the translational efficiency of mRNAs with complex 5' UTRs may depend on mechanisms that are unaffected or minimally affected by the loss of KDM2B.

Notably, loss of ribosomal proteins induces a similar effect, whereupon decreased ribosomal abundance due to mutations or loss of ribosomal proteins led to decreased translational efficiency of short transcripts as opposed to longer ones (*41, 42*). This data mainly focused on ribosomal diseases such as Diamond-Blackfan Anemia. Groups investigating these pathologies characterized how mutations or knockdown of ribosomal

proteins impacted translation selectivity (*41, 42*). Similar to our data, most findings converged on decreased translation of cell-cycle regulators along with decreased translation of short, unstructured transcripts. This suggests that our Ribo-seq findings were most likely due to decreased ribosomal abundance.

While this data overlap is promising, it still begs the question of why shorter transcripts appear to be more sensitive to ribosomal depletion. Interestingly, shorter transcripts tend to have a higher ribosomal density per length of transcript and have much higher basal translation efficiency as compared to longer transcripts (*41, 212*). This means that shorter transcripts are able to recruit ribosomes very efficiently. One explanation for this phenomenon is that shorter transcripts are translated via what is known as the closed-loop model of translation. The transcripts form a closed loop, with the starting 5' UTR of the transcript coming into contact with its terminal 3' UTR via RNA binding proteins, which mainly include translation factors. This connection between the two ends forms a closed loop by which the main coding sequence is translated (*234*). Thus, ribosomes that are terminating near the 3'UTR are close to the initiation point at the 5' UTR again and are able to re-initiate onto the same transcript. This creates an environment where shorter transcripts are less limited by elongation time and are able to re-initiate faster and exhibit increased basal translational efficiency. In contrast, longer transcripts have longer elongation times and, thus, are less able to reinitiate efficiently and, therefore, exhibit decreased basal translational efficiency. Mathematical prediction modeling of this concept was able to accurately model data in yeast and humans when accounting for high re-initiation rates (*235*). Firstly, the group commented, as others have described, that there was a negative correlation between coding sequence length and

ribosome density in both yeast and humans (Figure 4.2). Additionally, they noted that longer transcripts had a decreased effective initiation rate (Figure 4.2). The group predicted that re-initiation was the reason for this discrepancy and that ribosomes on shorter transcripts were much more able to re-initiate than ribosomes on longer transcripts. The work demonstrated that high re-initiation is able to predict length-dependent translational efficiency (Figure 4.2). In their model, a 0 % re-initiation rate would mean that longer and shorter transcripts have even ribosomal densities during translation, and even a 50% re-initiation rate would offer minimal difference to explain the ribosomal density difference between short and long transcripts. However, a re-initiation efficiency of 100% was able to better predict the differences in ribosomal density between short and long transcripts (Figure 4.2). In regard to KDM2B, decreased ribosomal availability, particularly for the 40S ribosomal subunit, may limit this process and decrease the re-initiation rate on shorter transcripts. This would create a situation where shorter transcripts appear relatively downregulated translationally as they more readily lose ribosomal density. In contrast, longer transcripts, even if global ribosomes are decreased, would appear relatively upregulated translationally as they are less dependent on re-initiation.

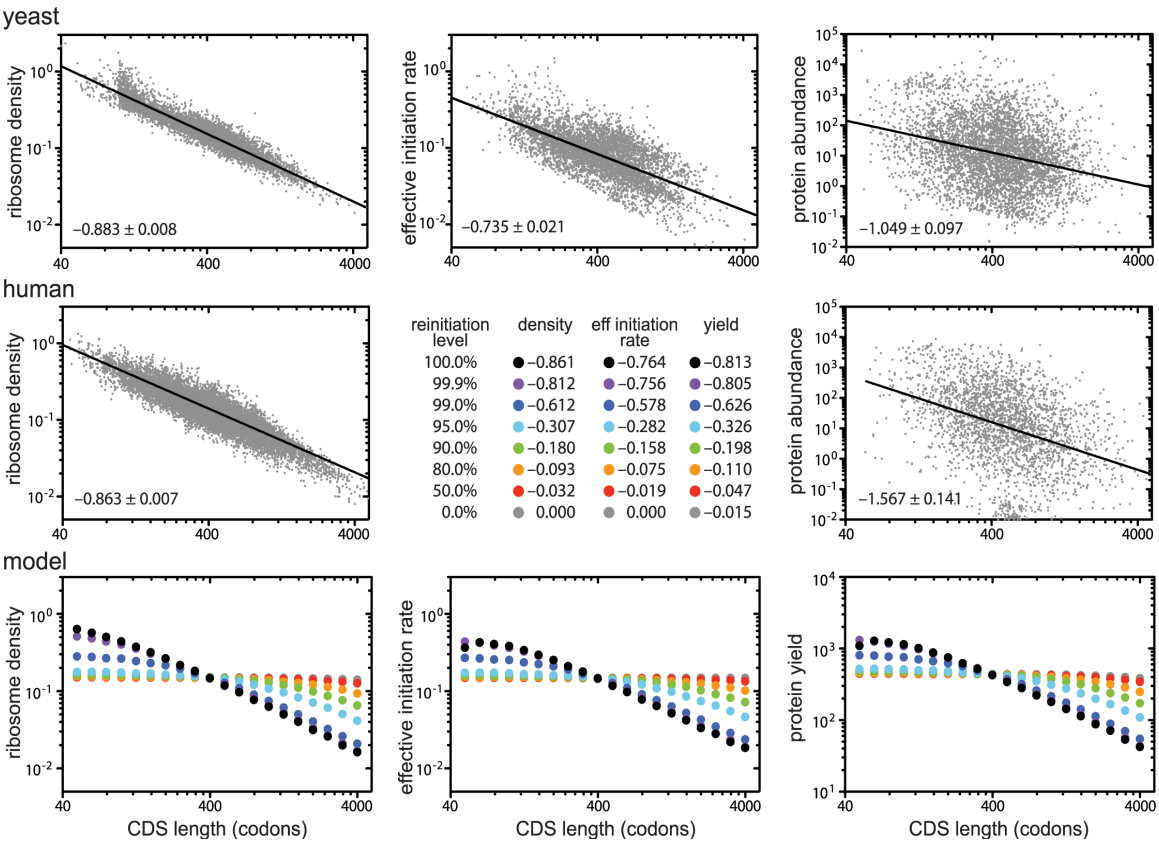


Figure 4.2 - High Re-initiation rates are able to predict ribosomal density on short and long transcripts.
Reprinted with permission from Rogers DW, Böttcher MA, Traulsen A, Greig D (2017) Ribosome re-initiation can explain length-dependent translation of messenger RNA. PLoS Comput Biol 13(6): e1005592. https://doi.org/10.1371/journal.pcbi.1005592

Interestingly, the mathematical model assumes that transcripts are equally able to form closed loops. Thus, the only difference between short and long transcripts in the mathematical model is the transit time of ribosomes during elongation (*235*). However, more recent biological data suggested that short transcripts are more likely to form closed-loops than long transcripts (*236-238*); therefore, not only do longer transcripts benefit less from closed-loop translation as compared to shorter transcripts, but they also are less able to even enter closed-loop models further limiting their ability to re-initiate ribosomes (*236-238*). Additionally, work from the Gilbert lab identified that mutants of RACK1, which associates with the 40S ribosome, lead to decreased translation on strong

closed-loop forming transcripts as opposed to longer transcripts that are less likely to form closed-loops (*236-238*). This work demonstrated that mutations of a protein associated with the 40S ribosomal subunit selectively impacted the translation of closed-looped forming transcripts. Interestingly, the knockdown of KDM2B reduces the levels of RACK1 at the RNA and protein levels (Figure 3.2). Thus, the translational phenotype noted may be dependent on RACK1. It is also plausible that RACK1 aids in the function of the 40S ribosome, and the phenotype of which transcripts are affected is due to the relative decrease of functional 40S ribosomal subunits.

In order to form a closed-loop for translation, the initiation factors in the 5' UTR need to bind to Poly-A Binding Protein (PABP) in the 3'UTR. It had been thought that cap-dependent translation is exclusively under the control of the eIF4F complex, a trimeric complex composed of the scaffold protein eIF4G, the cap-binding protein eIF4E, and the RNA helicase eIF4A. More recently, however, this dogma was challenged, with studies showing that at least 20% of capped mRNAs utilize, instead of the major eIF4G isoform (eIF4G1), the eIF4G1 homolog eIF4G2 (also known as DAP5, NAT1, or p97), which lacks eIF4E binding domain and utilizes eIF3d to initiate cap-dependent translation (*239*). Interestingly, eIF4G2 also lacks a PABP binding domain, suggesting that it does not participate in closed-loop translation (*240*). The interaction of the eIF3d-eIF4G2 complex with the target mRNAs is stabilized by binding to the complex stem-loop structures of post-transcriptional control elements (PCEs) in their 5' UTRs. Other proteins that may replace eIF4G1 in cap-dependent translation of mRNAs with complex 5' UTRs include the translation initiation factor eIF3 (eIF3a, eIF3b and eIF3g) (*241*), the helicase DHX9/RHA (*242*), the CBP80/20 (NCBP1/NCBP2)-dependent translation

initiation factor (CTIF) (*243*), and others (*244, 245*). These proteins may be recruited to their target mRNAs by eIF3d, as is the case for eIF4G2 (*239*), or other cap-binding proteins, as is the case for DHX9/RHA, which is recruited by the NCBP3/NCBP1 CBP complex (*246*). These alternative mechanisms of the translation may also help explain the bias of specific transcripts after the loss of ribosomal proteins, as many alternative complexes lack the ability to bind to PABPs and thus fail to form a closed-loop model, which tends to be more readily formed in short mRNAs (*236-238, 240*).

4.8. Future Directions

Major questions still remain regarding how KDM2B is able to regulate gene expression beyond its traditionally thought repressive function. Given that KDM2B binds to the promoters of these loci, it is likely to function as a direct transcriptional regulator of ribosomal genes. Thus, it is unlikely that the mechanism of regulation is that KDM2B is repressing a negative regulator of ribosomal biogenesis and more likely participating in the activation of transcription on these genes. Future work should characterize which domains of KDM2B are critical for regulating gene expression of ribosomal targets and if specific isoforms of KDM2B are more or less involved in this regulation. Given that recent literature has called into question the demethylase activity of KDM2B (*219*), future work should also characterize histone marks to see if they are altered following the knockdown of KDM2B.

KDM2B's co-occupancy with MYC on the chromatin of target genes raises the question of whether their function depends on one another or if there is an epistatic relationship between KDM2B and MYC. Knockdown of KDM2B or MYC reduced the expression of the other, and rescue of MYC onto the knockdown of KDM2B was able to

restore KDM2B's level despite the presence of the shRNA. Developing a knock-out of KDM2B would be able to overcome this issue; however, given that KDM2B regulates key processes for cellular growth and survival, it will be difficult to effectively generate knock-out clones of KDM2B that are able to survive unless compensatory mechanisms are present in the cell. In fact, over time, we have noted that the shRNA, despite being constitutive, becomes less effective as the cells pass through more passages, suggesting a selective pressure to express KDM2B in these cells. Therefore, cells might not survive following the knockout of KDM2B. Other tools, such as dCas9-KRAB, where a dead-nuclease Cas9 fused to KRAB is targeted to a specific locus and epigenetically represses the locus, may be able to prevent the co-regulation of one gene to the other and determine the sufficiency of KDM2B or MYC alone in impacting ribosomal biogenesis genes. This tool can be used while being under the control of an inducible promoter and, thus, is not reliant on cells surviving knockouts. However, as has been the case with many genetic programs regulating core cellular functions, upstream regulators are often co-regulated in order to bring about proper gene expression of all necessary components (See Section 1.2) (*234*). Thus, the regulation of KDM2B and MYC may not be epistatic, but rather, both are important for the transcription of ribosomal biogenesis genes.

The disproportionate effect between the 40S and 60S subunits was characterized to be a specific defect in processing in the 30S pre-ribosomal intermediate. As proteomics data and RNA-seq data exist on processing defects following the knockdown of nucleolar factors, those datasets should be interrogated to generate a list of factors that are impacted by KDM2B and participate in the processing of the 30S to 26S based on already known effectors (*227*). These factors should be used in knockdown experiments to demonstrate

they are necessary for 30S processing in this specific cell line model and interrogate if they phenocopy the knockdown of KDM2B in terms of rRNA processing. Then, the factors that phenocopy the knockdown of KDM2B should be used in rescue experiments to see if they can rescue the 30S to 26S processing defect after the knockdown of KDM2B. One major complication is that multiple ribosomal subunit proteins participate in the processing of rRNA and also display decreased transcript levels after the knockdown of KDM2B. Thus, it becomes challenging to tease apart if the effect is due to a ribosomal protein that ends up being rate-limiting or an upstream ribosome biogenesis factor (*227*). Stated differently, are the ribosomal proteins downregulated because processing is impacted, or are they direct participants in the process and a primary cause of the rRNA processing defect? While rescue experiments can be trialed, traditional rescue experiments following the knockdown of KDM2B would be challenging as many of these factors are co-regulated. For example, assuming that KDM2B regulates a "keystone" factor in the processing and loss of KDM2B leads to that keystone factor becoming rate limiting. Eventually, the other factors involved in the process will also decrease their expression level as they are co-regulated (*61*). Thus, by the time one rescues the initial keystone perturbation, other factors now have become rate-limiting, and no rescue is noted. Regardless, experiments demonstrating the necessity of the factors will solidify which of these factors, with their levels impacted by KDM2B, play a role in the processing defect seen after the knockdown of KDM2B. As data already exists (*227*) on many nucleolar factors, one can interrogate our existing transcriptomic or proteomic data on which factors are high-yield to investigate.

Transcript-specific translational regulation is a key aspect of gene expression control. KDM2B impacts many upstream factors, making it challenging to ascertain one factor that is modulating the translational efficiency changes of any given transcript. Alternative translational initiation mechanisms provide an attractive hypothesis, where some transcripts are dependent on more canonical mechanisms via eIF4G1 as opposed to eiF4G2 or other factors. Multiple groups have found that loss of ribosomal protein factors leads to decreased translation of short, strong closed loop forming transcripts (*41, 42, 212, 236-238*), which mimics our data following the loss of KDM2B. This lends credence to our hypothesis that the translational changes noted by KDM2B are dependent on decreased ribosomal abundance. However, it is currently impossible to exclusively alter the number of ribosomes in a cell and ascertain how ribosomal number exclusively impacts translation. However, eIF4G1 and eIF4G2 represent two factors that participate in initiation, with eIF4G1 being able to form closed loops due to PABP binding domains and eIF4G2 being unable to do so (*240*). Given the pattern of short transcripts being closed-loop dependent and translationally downregulated by loss of KDM2B, these factors provide an avenue to investigate if initiation factor selection predicts translationally sensitivity to loss of KDM2B. This can be done by performing RNA-immunoprecipitation-sequencing to characterize which transcripts are bound to eIF4G1 or eIF4G2 in control and knockdown of KDM2B. We would identify four groups of transcripts: ones that exclusively use eIF4G1, use eIF4G2, switch from eIF4G1 to eIF4G2 or switch from eIF4G2 to eIF4G1 following the knockdown of KDM2B. These categories can then be compared to our KDM2B Ribo-seq data to identify if different alternative translation mechanisms and dynamics predict translation changes by KDM2B.

This would characterize transcript specific translational initiation mechanisms and, therefore, provide more specific insight into translational regulation by KDM2B.

4.9. Contributions

Vollter Anastas wrote the discussion. Philip N. Tsichlis edited the discussion.

Chapter 5: Appendix

Table 5.1 - Table of Constructs

Construct	Source	Usage
pLKO.1-Control	https://www.addgene.org/10879/	Control plasmid with non-targeting shRNA
pLKO.1-shKDM2B #1	Millipore Sigma pre-designed shRNA TRCN0000238779	shRNA lentiviral plasmid targeting *KDM2B*
pLKO.1-shKDM2B #2	Millipore Sigma pre-designed shRNA TRCN0000118437	shRNA lentiviral plasmid targeting *KDM2B*
pLKO.1- shEZH2	Millipore Sigma pre-designed shRNA TRCN0000018365	shRNA lentiviral plasmid targeting *EZH2*
pPAX2	https://www.addgene.org/12260/	2nd generation lentiviral packaging plasmid
pMD2.G	https://www.addgene.org/12259/	2nd generation lentiviral packaging plasmid
pSMART RPS3-HALO-FLAG Homology Cloning Vector	Gift from J. Wade Harper (*167*)	Donor plasmid for the editing of RPS3 endogenous locus via CRISPR Cas9
pSMART RPL29-HALO Homology Cloning Vector	Gift from J. Wade Harper (*167*)	Donor plasmid for the editing of RPL29 endogenous locus via CRISPR Cas9
pSpCas9-2A-Puro	https://www.addgene.org/62988/	Vector backbone to clone sgRNA into a transient puromycin selectable vector for knockouts or testing gRNA efficiency
pSpCas9-hGem	https://www.addgene.org/71707/	Vector backbone to clone sgRNA into a transient transfection vector where Cas9 is fused to the degradation tag of Geminin

Table 5.2 - Table of DNA Oligos

Oligonucleotides	Name	Oligonucleotide sequence (5' to 3')
5' ETS Cloning and T7 Addition from gDNA	5' ETS Forward Primer	CCCGTGGTCTCTCGTCTTCT
	5' ETS Reverse Primer	CTGTCCGGGAGGGACCAC
	5' ETS Reverse Primer T7 addition	TAATACGACTCACTATAGGGctgtccg ggagggaccac*
ITS1 Cloning and T7 Addition from gDNA	ITS1 Forward Primer	GCGAGAGCCGAGAACTC
	ITS1 Reverse Primer	GAGGAGGGCACCGAGACC
	ITS1 Reverse Primer T7 addition	TAATACGACTCACTATAGGGgaggag ggcaccgagacc*
ITS2 Cloning and T7 Addition from gDNA	ITS2 Forward Primer	GTCCCCCTAAGCGCAGAC
	ITS2 Reverse Primer	GGCTCTCTCTTTCCCTCTCC
	ITS2 Reverse Primer T7 addition	TAATACGACTCACTATAGGGggctctc tctttccctctcc*
3' ETS Cloning and T7 Addition from gDNA	3' ETS Forward Primer	CTCGACACAAGGGTTTGTCC
	3' ETS Reverse Primer	AGGCGGGAACCGAAGAAG
	3' ETS Reverse Primer T7 addition	TAATACGACTCACTATAGGGaggcgg gaaccgaagaag*
RPS3 sgRNA Oligos	Strand Targeting Oligo	caccGACATACCTGTTATGCTGTG**
	Complementary Oligo	aaacCACAGCATAACAGGTATGTC**
RPL29 sgRNA Oligos	Strand Targeting Oligo	caccGAGATATCTCTGCCAACATG**
	Complementary Oligo	aaacCATGTTGGCAGAGATATCTC**

*T7 addition is capitalized, whereas the annealing segment matching to the reverse primer sequence is lowercase.
**sgRNA targeting sequence is capitalized, whereas the lower-case additions (CACC on the targeting oligo or AAAC on the complementary oligo are used for cloning purposes as described in Section 2.8.

Table 5.3 - Table of Antibodies for Western Blots

Antibody	Catalog Number	Concentration used
anti-Biotin-HRP	Cell Signaling Technology Cat: 7075S	1:2000
anti-CCNB1	Cell Signaling Technology Cat: 12231S	1:1000
anti-cdc2 (CDK1)	Cell Signaling Technology Cat: 9116S	1:1000
anti-EIF2α	Cell Signaling Technology Cat: 9722S	1:1000
anti-EZH2	Cell Signaling Technology Cat: 5246S	1:1000
anti-FBL	Cell Signaling Technology Cat: 5246S	1:1000
anti-JHDM1B (KDM2B)	Millipore Cat: 09-864	1:1500
anti-NAT10	Proteintech Cat: 13365-1-AP	1:1000
anti-Phospho-EIF2α Ser51	Cell Signaling Technology Cat: 9721S	1:1000
anti-RPL29	Proteintech Cat: 15799-1-AP	1:1000
anti-RPS28	Thermo Scientific Cat: PA5-45721	1:1000
anti-RPS3	Cell Signaling Technology Cat: 9538S	1:1000
anti-TCOF1	Proteintech Cat: 11003-1-AP	1:1000
anti-UTP14A	Proteintech Cat: 11474-1-AP	1:1000
ZAKα	Thermo Scientific Cat: A301-993A	1:1000
α-Actinin	Cell Signaling Technology Cat: 6487S	1:2000
α-Tubulin	Sigma Cat: T5168	1:3000
β-actin	Cell Signaling Technology Cat: 4970S	1:3000

Table 5.4 - Table of Antibodies for ChIP-seq

Antibody	Catalog Number	Concentration used
anti-JHDM1B ChIP+ (KDM2B)	Millipore Cat: 17-10264	1:100
anti-Myc	Cell Signaling Technology Cat: 18583S	1:100
anti-H3K4me3	Cell Signaling Technology Cat: 9751S	1:50
anti-H3K27me3	Cell Signaling Technology Cat: 9733S	1:50
anti-H2AK119ub	Cell Signaling Technology Cat: 8240S	1:100

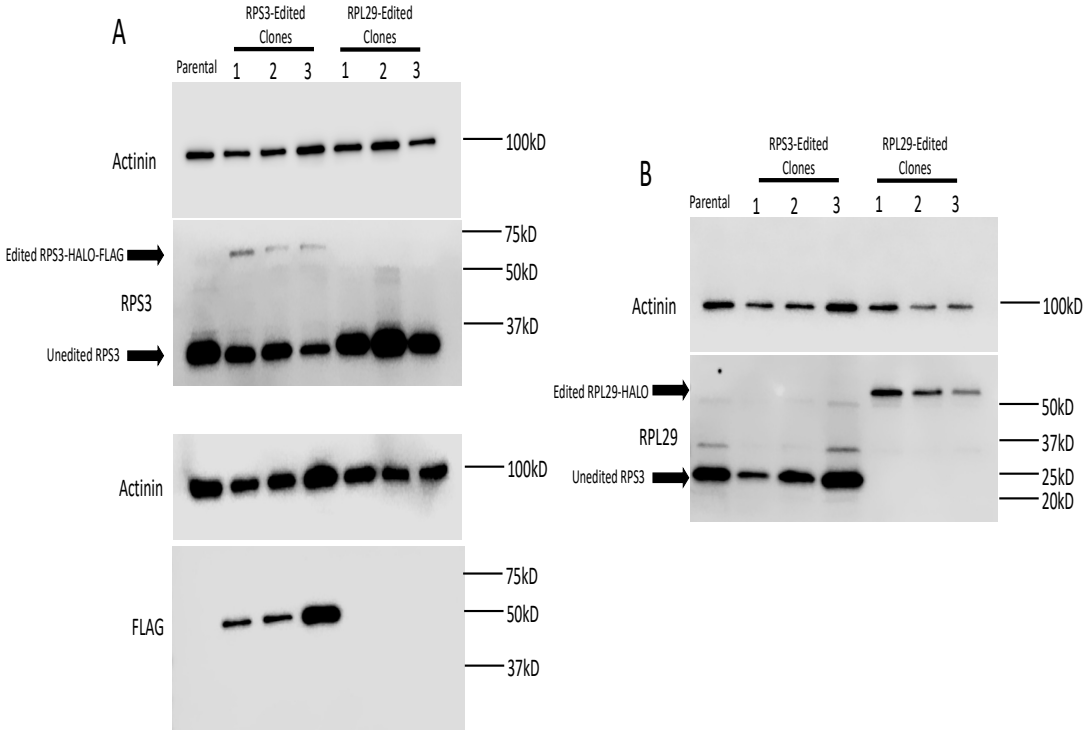


Figure 5.1- Confirmation of Editing of *RPS3* and *RPL29* loci
Western blot analysis probing for endogenous ribosomal proteins to detect size shift for both A. RPS3 and B. RPL29. RPS3 homology design also contained a A.FLAG tag allowing for FLAG detection of proper editing.

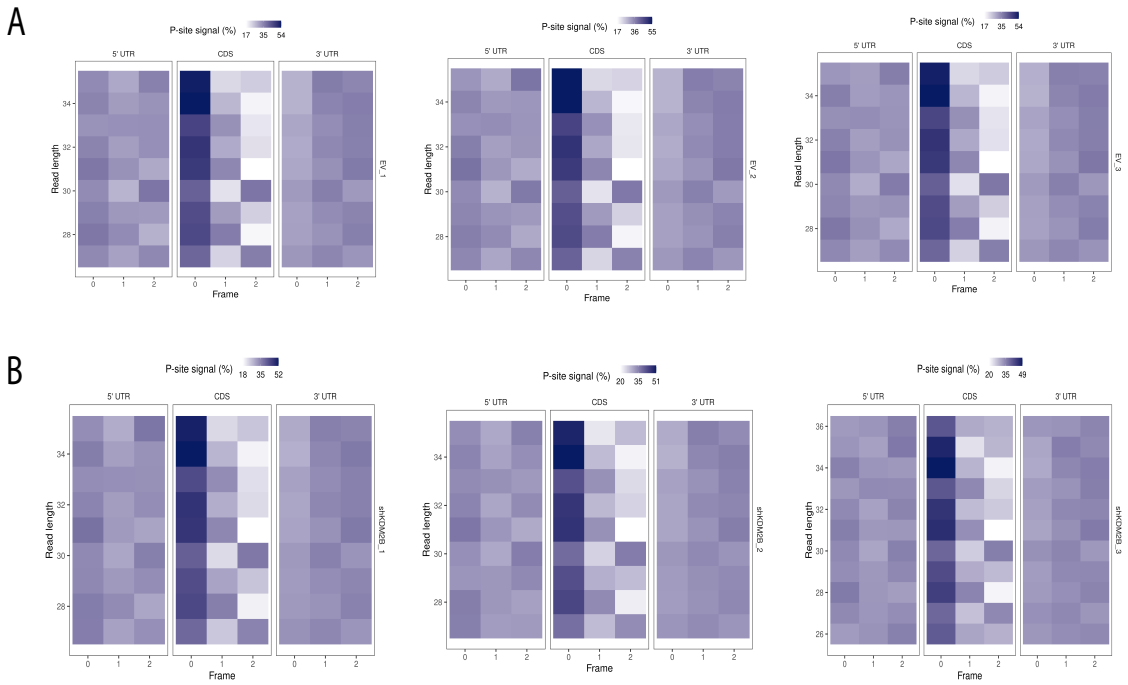


Figure 5.2 - P-site density of Ribo-seq across all samples.
P-site % at each frame (0,1,2) in the 5'UTR, CDS, and 3'UTR for A. 3 replicates of the Control and B. 3 replicates of the shKDM2B.

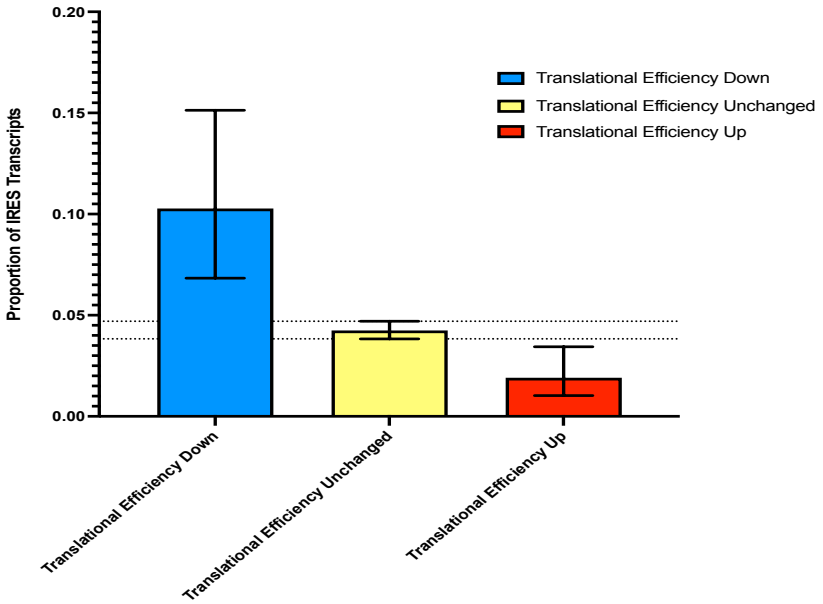


Figure 5.3 Proportion of IRES-containing transcripts in different translational categories
Proportion of transcripts that contain experimentally validated IRES elements from a curated database (*1*). Error bars represent the 95% confidence interval of each proportion with dotted lines denoting the 95% confidence interval of the translationally unchanged transcripts.

Chapter 6: Bibliography

1. J. Zhao, Y. Li, C. Wang, H. Zhang, H. Zhang, B. Jiang, X. Guo, X. Song, IRESbase: A Comprehensive Database of Experimentally Validated Internal Ribosome Entry Sites. *Genomics Proteomics Bioinformatics* **18**, 129-139 (2020); published online EpubApr (10.1016/j.gpb.2020.03.001).
2. V. Iadevaia, X. Wang, Z. Yao, L. J. Foster, C. G. Proud, Evaluation of mTOR-regulated mRNA translation. *Methods Mol Biol* **821**, 171-185 (2012)10.1007/978-1-61779-430-8_10).
3. T. V. Pestova, V. G. Kolupaeva, I. B. Lomakin, E. V. Pilipenko, I. N. Shatsky, V. I. Agol, C. U. T. Hellen, Molecular mechanisms of translation initiation in eukaryotes. *Proceedings of the National Academy of Sciences* **98**, 7029-7036 (2001)10.1073/pnas.111145798).
4. C. Mayer, I. Grummt, Ribosome biogenesis and cell growth: mTOR coordinates transcription by all three classes of nuclear RNA polymerases. *Oncogene* **25**, 6384-6391 (2006); published online Epub2006/10/01 (10.1038/sj.onc.1209883).
5. S. M. Kooistra, K. Helin, Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Biol* **13**, 297-311 (2012); published online EpubApr 4 (10.1038/nrm3327).
6. A. Sachs, R. Davis, The poly(A)-binding protein is required for translation initiation and poly(A) tail shortening. *Mol Biol Rep* **14**, 73 (1990)10.1007/bf00360421).
7. S. Xue, M. Barna, Cis-regulatory RNA elements that regulate specialized ribosome activity. *RNA Biol* **12**, 1083-1087 (2015)10.1080/15476286.2015.1085149).
8. P. G. Donlin-Asp, W. Rossoll, G. J. Bassell, Spatially and temporally regulating translation via mRNA-binding proteins in cellular and neuronal function. *FEBS Lett* **591**, 1508-1525 (2017); published online EpubJun (10.1002/1873-3468.12621).
9. M. Alqinyah, S. B. Hooks, Regulating the regulators: Epigenetic, transcriptional, and post-translational regulation of RGS proteins. *Cell Signal* **42**, 77-87 (2018); published online EpubJan (10.1016/j.cellsig.2017.10.007).
10. S. J. Serrano-Gomez, M. Maziveyi, S. K. Alahari, Regulation of epithelial-mesenchymal transition through epigenetic and post-translational modifications. *Mol Cancer* **15**, 18 (2016); published online EpubFeb 24 (10.1186/s12943-016-0502-x).
11. Y. Liu, A. Beyer, R. Aebersold, On the Dependency of Cellular Protein Levels on mRNA Abundance. *Cell* **165**, 535-550 (2016); published online EpubApr 21 (10.1016/j.cell.2016.03.014).
12. F. Edfors, F. Danielsson, B. M. Hallstrom, L. Kall, E. Lundberg, F. Ponten, B. Forsstrom, M. Uhlen, Gene-specific correlation of RNA and protein levels in human cells and tissues. *Mol Syst Biol* **12**, 883 (2016); published online EpubOct 20 (10.15252/msb.20167144).
13. F. Edfors, F. Danielsson, B. M. Hallström, L. Käll, E. Lundberg, F. Pontén, B. Forsström, M. Uhlén, Gene-specific correlation of RNA and protein levels in

human cells and tissues. *Mol Syst Biol* **12**, 883 (2016); published online EpubOct 20 (10.15252/msb.20167144).
14. K. A. Spriggs, M. Bushell, A. E. Willis, Translational regulation of gene expression during conditions of cell stress. *Mol Cell* **40**, 228-237 (2010); published online EpubOct 22 (10.1016/j.molcel.2010.09.028).
15. K. Van Der Kelen, R. Beyaert, D. Inzé, L. De Veylder, Translational control of eukaryotic gene expression. *Crit Rev Biochem Mol Biol* **44**, 143-168 (2009); published online EpubJul-Aug (10.1080/10409230902882090).
16. M. VanInsberghe, J. van den Berg, A. Andersson-Rolf, H. Clevers, A. van Oudenaarden, Single-cell Ribo-seq reveals cell cycle-dependent translational pausing. *Nature* **597**, 561-565 (2021); published online Epub2021/09/01 (10.1038/s41586-021-03887-4).
17. J. M. Barnes, L. Przybyla, V. M. Weaver, Tissue mechanics regulate brain development, homeostasis and disease. *J Cell Sci* **130**, 71-82 (2017); published online EpubJan 1 (10.1242/jcs.191742).
18. X. Wang, Y. Xuan, Y. Han, X. Ding, K. Ye, F. Yang, P. Gao, S. P. Goff, G. Gao, Regulation of HIV-1 Gag-Pol Expression by Shiftless, an Inhibitor of Programmed -1 Ribosomal Frameshifting. *Cell* **176**, 625-635.e614 (2019); published online EpubJan 24 (10.1016/j.cell.2018.12.030).
19. M. U. Ali, M. S. Ur Rahman, Z. Jia, C. Jiang, Eukaryotic translation initiation factors and cancer. *Tumour Biol* **39**, 1010428317709805 (2017); published online EpubJun (10.1177/1010428317709805).
20. A. G. Hinnebusch, I. P. Ivanov, N. Sonenberg, Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science* **352**, 1413-1416 (2016); published online EpubJun 17 (10.1126/science.aad9868).
21. R. F. Harvey, T. S. Smith, T. Mulroney, R. M. L. Queiroz, M. Pizzinga, V. Dezi, E. Villenueva, M. Ramakrishna, K. S. Lilley, A. E. Willis, Trans-acting translational regulatory RNA binding proteins. *Wiley Interdiscip Rev RNA* **9**, e1465 (2018); published online EpubMay (10.1002/wrna.1465).
22. S. Hombach, M. Kretz, Non-coding RNAs: Classification, Biology and Functioning. *Adv Exp Med Biol* **937**, 3-17 (2016)10.1007/978-3-319-42059-2_1).
23. H. Guo, Specialized ribosomes and the control of translation. *Biochem Soc Trans* **46**, 855-869 (2018); published online EpubAug 20 (10.1042/bst20160426).
24. B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, Global quantification of mammalian gene expression control. *Nature* **473**, 337-342 (2011); published online EpubMay 19 (10.1038/nature10098).
25. I. Kosti, N. Jain, D. Aran, A. J. Butte, M. Sirota, Cross-tissue Analysis of Gene and Protein Expression in Normal and Cancer Tissues. *Sci Rep* **6**, 24799 (2016); published online EpubMay 4 (10.1038/srep24799).
26. C. Vogel, E. M. Marcotte, Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat Rev Genet* **13**, 227-232 (2012); published online EpubMar 13 (10.1038/nrg3185).
27. K. Perl, K. Ushakov, Y. Pozniak, O. Yizhar-Barnea, Y. Bhonker, S. Shivatzki, T. Geiger, K. B. Avraham, R. Shamir, Reduced changes in protein compared to

mRNA levels across non-proliferating tissues. *BMC Genomics* **18**, 305 (2017); published online EpubApr 18 (10.1186/s12864-017-3683-9).
28. D. Shore, B. Albert, Ribosome biogenesis and the cellular energy economy. *Current Biology* **32**, R611-R617 (2022); published online Epub2022/06/20/ (https://doi.org/10.1016/j.cub.2022.04.083).
29. J. Pelletier, G. Thomas, S. Volarević, Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nature Reviews Cancer* **18**, 51-63 (2018); published online Epub2018/01/01 (10.1038/nrc.2017.104).
30. M. Fromont-Racine, B. Senger, C. Saveanu, F. Fasiolo, Ribosome assembly in eukaryotes. *Gene* **313**, 17-42 (2003); published online EpubAug 14 (10.1016/s0378-1119(03)00629-2).
31. Z. Shi, K. Fujii, K. M. Kovary, N. R. Genuth, H. L. Röst, M. N. Teruel, M. Barna, Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. *Mol Cell* **67**, 71-83.e77 (2017); published online EpubJul 6 (10.1016/j.molcel.2017.05.021).
32. D. Simsek, G. C. Tiu, R. A. Flynn, G. W. Byeon, K. Leppek, A. F. Xu, H. Y. Chang, M. Barna, The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* **169**, 1051-1065 e1018 (2017); published online EpubJun 1 (10.1016/j.cell.2017.05.022).
33. T. Hopes, K. Norris, M. Agapiou, C. G. P. McCarthy, P. A. Lewis, M. J. O'Connell, J. Fontana, J. L. Aspden, Ribosome heterogeneity in Drosophila melanogaster gonads through paralog-switching. *Nucleic Acids Res* **50**, 2240-2257 (2022); published online EpubFeb 28 (10.1093/nar/gkab606).
34. S. Xue, M. Barna, Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat Rev Mol Cell Biol* **13**, 355-369 (2012); published online EpubMay 23 (10.1038/nrm3359).
35. D. Simsek, M. Barna, An emerging role for the ribosome as a nexus for post-translational modifications. *Curr Opin Cell Biol* **45**, 92-101 (2017); published online EpubApr (10.1016/j.ceb.2017.02.010).
36. S. G. Clarke, Protein methylation at the surface and buried deep: thinking outside the histone box. *Trends Biochem Sci* **38**, 243-252 (2013); published online EpubMay (10.1016/j.tibs.2013.02.004).
37. A. Mahesh, M. I. K. Khan, G. Govindaraju, M. Verma, S. Awasthi, P. L. Chavali, S. Chavali, A. Rajavelu, A. Dhayalan, SET7/9 interacts and methylates the ribosomal protein, eL42 and regulates protein synthesis. *Biochim Biophys Acta Mol Cell Res* **1867**, 118611 (2020); published online EpubFeb (10.1016/j.bbamcr.2019.118611).
38. N. R. Genuth, M. Barna, The Discovery of Ribosome Heterogeneity and Its Implications for Gene Regulation and Organismal Life. *Mol Cell* **71**, 364-374 (2018); published online EpubAug 2 (10.1016/j.molcel.2018.07.018).
39. E. W. Mills, R. Green, Ribosomopathies: There's strength in numbers. *Science* **358**, (2017); published online EpubNov 3 (10.1126/science.aan2755).
40. H. F. Lodish, Model for the regulation of mRNA translation applied to haemoglobin synthesis. *Nature* **251**, 385-388 (1974); published online EpubOct 4 (10.1038/251385a0).

41. Z. Cheng, C. F. Mugler, A. Keskin, S. Hodapp, L. Y. Chan, K. Weis, P. Mertins, A. Regev, M. Jovanovic, G. A. Brar, Small and Large Ribosomal Subunit Deficiencies Lead to Distinct Gene Expression Signatures that Reflect Cellular Growth Rate. *Mol Cell* **73**, 36-47.e10 (2019); published online EpubJan 3 (10.1016/j.molcel.2018.10.032).
42. R. K. Khajuria, M. Munschauer, J. C. Ulirsch, C. Fiorini, L. S. Ludwig, S. K. McFarland, N. J. Abdulhay, H. Specht, H. Keshishian, D. R. Mani, M. Jovanovic, S. R. Ellis, C. P. Fulco, J. M. Engreitz, S. Schütz, J. Lian, K. W. Gripp, O. K. Weinberg, G. S. Pinkus, L. Gehrke, A. Regev, E. S. Lander, H. T. Gazda, W. Y. Lee, V. G. Panse, S. A. Carr, V. G. Sankaran, Ribosome Levels Selectively Regulate Translation and Lineage Commitment in Human Hematopoiesis. *Cell* **173**, 90-103.e119 (2018); published online EpubMar 22 (10.1016/j.cell.2018.02.036).
43. L. S. Ludwig, H. T. Gazda, J. C. Eng, S. W. Eichhorn, P. Thiru, R. Ghazvinian, T. I. George, J. R. Gotlib, A. H. Beggs, C. A. Sieff, H. F. Lodish, E. S. Lander, V. G. Sankaran, Altered translation of GATA1 in Diamond-Blackfan anemia. *Nat Med* **20**, 748-753 (2014); published online EpubJul (10.1038/nm.3557).
44. T. Chaillou, T. J. Kirby, J. J. McCarthy, Ribosome biogenesis: emerging evidence for a central role in the regulation of skeletal muscle mass. *J Cell Physiol* **229**, 1584-1594 (2014); published online EpubNov (10.1002/jcp.24604).
45. J.-J. Jia, R. M. Lahr, M. T. Solgaard, B. J. Moraes, R. Pointet, A.-D. Yang, G. Celucci, T. E. Graber, H.-D. Hoang, Marius R. Niklaus, I. A. Pena, A. K. Hollensen, E. M. Smith, M. Chaker-Margot, L. Anton, C. Dajadian, M. Livingstone, J. Hearnden, X.-D. Wang, Y. Yu, T. Maier, C. K. Damgaard, A. J. Berman, T. Alain, B. D. Fonseca, mTORC1 promotes TOP mRNA translation through site-specific phosphorylation of LARP1. *Nucleic Acids Research* **49**, 3461-3489 (2021)10.1093/nar/gkaa1239).
46. H. B. Jefferies, S. Fumagalli, P. B. Dennis, C. Reinhard, R. B. Pearson, G. Thomas, Rapamycin suppresses 5'TOP mRNA translation through inhibition of p70s6k. *Embo j* **16**, 3693-3704 (1997); published online EpubJun 16 (10.1093/emboj/16.12.3693).
47. C. C. Thoreen, L. Chantranupong, H. R. Keys, T. Wang, N. S. Gray, D. M. Sabatini, A unifying model for mTORC1-mediated regulation of mRNA translation. *Nature* **485**, 109-113 (2012); published online EpubMay 2 (10.1038/nature11083).
48. L. Philippe, A. M. G. van den Elzen, M. J. Watson, C. C. Thoreen, Global analysis of LARP1 translation targets reveals tunable and dynamic features of 5′ TOP motifs. *Proceedings of the National Academy of Sciences* **117**, 5319-5328 (2020)doi:10.1073/pnas.1912864117).
49. C. Mayer, J. Zhao, X. Yuan, I. Grummt, mTOR-dependent activation of the transcription factor TIF-IA links rRNA synthesis to nutrient availability. *Genes Dev* **18**, 423-434 (2004); published online EpubFeb 15 (10.1101/gad.285504).
50. K. M. Hannan, Y. Brandenburger, A. Jenkins, K. Sharkey, A. Cavanaugh, L. Rothblum, T. Moss, G. Poortinga, G. A. McArthur, R. B. Pearson, R. D. Hannan, mTOR-dependent regulation of ribosomal gene transcription requires S6K1 and is mediated by phosphorylation of the carboxy-terminal activation domain of the

nucleolar transcription factor UBF. *Mol Cell Biol* **23**, 8862-8877 (2003); published online EpubDec (10.1128/mcb.23.23.8862-8877.2003).
51. M. S. Dai, H. Lu, Crosstalk between c-Myc and ribosome in ribosomal biogenesis and cancer. *J Cell Biochem* **105**, 670-677 (2008); published online EpubOct 15 (10.1002/jcb.21895).
52. J. van Riggelen, A. Yetil, D. W. Felsher, MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer* **10**, 301-309 (2010); published online EpubApr (10.1038/nrc2819).
53. S. K. Das, B. A. Lewis, D. Levens, MYC: a complex problem. *Trends in Cell Biology* **33**, 235-246 (2023)10.1016/j.tcb.2022.07.006).
54. Z. Nie, C. Guo, S. K. Das, C. C. Chow, E. Batchelor, S. S. J. Simons, D. Levens, Dissecting transcriptional amplification by MYC. *eLife* **9**, e52483 (2020); published online Epub2020/07/27 (10.7554/eLife.52483).
55. M. Muhar, A. Ebert, T. Neumann, C. Umkehrer, J. Jude, C. Wieshofer, P. Rescheneder, J. J. Lipp, V. A. Herzog, B. Reichholf, D. A. Cisneros, T. Hoffmann, M. F. Schlapansky, P. Bhat, A. von Haeseler, T. Köcher, A. C. Obenauf, J. Popow, S. L. Ameres, J. Zuber, SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis. *Science* **360**, 800-805 (2018); published online EpubMay 18 (10.1126/science.aao2793).
56. C. Grandori, N. Gomez-Roman, Z. A. Felton-Edkins, C. Ngouenet, D. A. Galloway, R. N. Eisenman, R. J. White, c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nature Cell Biology* **7**, 311-318 (2005); published online Epub2005/03/01 (10.1038/ncb1224).
57. N. Gomez-Roman, C. Grandori, R. N. Eisenman, R. J. White, Direct activation of RNA polymerase III transcription by c-Myc. *Nature* **421**, 290-294 (2003); published online EpubJan 16 (10.1038/nature01327).
58. R. J. White, RNA polymerases I and III, growth control and cancer. *Nature Reviews Molecular Cell Biology* **6**, 69-78 (2005); published online Epub2005/01/01 (10.1038/nrm1551).
59. T. Oskarsson, A. Trumpp, The Myc trilogy: lord of RNA polymerases. *Nature Cell Biology* **7**, 215-217 (2005); published online Epub2005/03/01 (10.1038/ncb0305-215).
60. I. Schlosser, M. Hölzel, M. Mürnseer, H. Burtscher, U. H. Weidle, D. Eick, A role for c-Myc in the regulation of ribosomal RNA processing. *Nucleic Acids Res* **31**, 6148-6156 (2003); published online EpubNov 1 (10.1093/nar/gkg794).
61. C. Petibon, M. Malik Ghulam, M. Catala, S. Abou Elela, Regulation of ribosomal protein genes: An ordered anarchy. *WIREs RNA* **12**, e1632 (2021)https://doi.org/10.1002/wrna.1632).
62. L. Sun, Y. Yan, H. Lv, J. Li, Z. Wang, K. Wang, L. Wang, Y. Li, H. Jiang, Y. Zhang, Rapamycin targets STAT3 and impacts c-Myc to suppress tumor growth. *Cell Chem Biol* **29**, 373-385.e376 (2022); published online EpubMar 17 (10.1016/j.chembiol.2021.10.006).
63. Z. Liu, T. Sun, C. Piao, Z. Zhang, C. Kong, METTL13 inhibits progression of clear cell renal cell carcinoma with repression on PI3K/AKT/mTOR/HIF-1α pathway and c-Myc expression. *J Transl Med* **19**, 209 (2021); published online EpubMay 13 (10.1186/s12967-021-02879-2).

64. N. G. Norwitz, H. Querfurth, mTOR Mysteries: Nuances and Questions About the Mechanistic Target of Rapamycin in Neurodegeneration. *Front Neurosci* **14**, 775 (2020)10.3389/fnins.2020.00775).
65. D. Horiuchi, B. Anderton, A. Goga, Taking on challenging targets: making MYC druggable. *Am Soc Clin Oncol Educ Book*, e497-502 (2014)10.14694/EdBook_AM.2014.34.e497).
66. A. R. Elhamamsy, B. J. Metge, H. A. Alsheikh, L. A. Shevde, R. S. Samant, Ribosome Biogenesis: A Central Player in Cancer Metastasis and Therapeutic Resistance. *Cancer Res* **82**, 2344-2353 (2022); published online EpubJul 5 (10.1158/0008-5472.Can-21-4087).
67. N. Robichaud, N. Sonenberg, D. Ruggero, R. J. Schneider, Translational Control in Cancer. *Cold Spring Harb Perspect Biol* **11**, (2019); published online EpubJul 1 (10.1101/cshperspect.a032896).
68. T. Charitou, S. Srihari, M. A. Lynn, M. A. Jarboui, E. Fasterius, M. Moldovan, S. Shirasawa, T. Tsunoda, M. Ueffing, J. Xie, J. Xin, X. Wang, C. G. Proud, K. Boldt, C. Al-Khalili Szigyarto, W. Kolch, D. J. Lynn, Transcriptional and metabolic rewiring of colorectal cancer cells expressing the oncogenic KRAS(G13D) mutation. *Br J Cancer* **121**, 37-50 (2019); published online EpubJul (10.1038/s41416-019-0477-7).
69. M. Nedeljković, N. Tanić, T. Dramićanin, Z. Milovanović, S. Šušnjar, V. Milinković, I. Vujović, M. Prvanović, N. Tanić, Importance of Copy Number Alterations of FGFR1 and C-MYC Genes in Triple Negative Breast Cancer. *J Med Biochem* **38**, 63-70 (2019); published online EpubMar (10.2478/jomb-2018-0012).
70. G. Mendiratta, E. Ke, M. Aziz, D. Liarakos, M. Tong, E. C. Stites, Cancer gene mutation frequencies for the U.S. population. *Nature Communications* **12**, 5961 (2021); published online Epub2021/10/13 (10.1038/s41467-021-26213-y).
71. T. Maehama, M. Nishio, J. Otani, T. W. Mak, A. Suzuki, Nucleolar stress: Molecular mechanisms and related human diseases. *Cancer Sci*, (2023); published online EpubFeb 10 (10.1111/cas.15755).
72. P. Jaako, J. Flygare, K. Olsson, R. Quere, M. Ehinger, A. Henson, S. Ellis, A. Schambach, C. Baum, J. Richter, J. Larsson, D. Bryder, S. Karlsson, Mice with ribosomal protein S19 deficiency develop bone marrow failure and symptoms like patients with Diamond-Blackfan anemia. *Blood* **118**, 6087-6096 (2011); published online EpubDec 1 (10.1182/blood-2011-08-371963).
73. S. R. Ellis, Nucleolar stress in Diamond Blackfan anemia pathophysiology. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **1842**, 765-768 (2014); published online Epub2014/06/01/ (https://doi.org/10.1016/j.bbadis.2013.12.013).
74. S. Dutt, A. Narla, K. Lin, A. Mullally, N. Abayasekara, C. Megerdichian, F. H. Wilson, T. Currie, A. Khanna-Gupta, N. Berliner, J. L. Kutok, B. L. Ebert, Haploinsufficiency for ribosomal protein genes causes selective activation of p53 in human erythroid progenitor cells. *Blood* **117**, 2567-2576 (2011); published online EpubMar 3 (10.1182/blood-2010-07-295238).
75. G. C. Tiu, C. H. Kerr, C. M. Forester, P. S. Krishnarao, H. D. Rosenblatt, N. Raj, T. C. Lantz, O. Zhulyn, M. E. Bowen, L. Shokat, L. D. Attardi, D. Ruggero, M.

Barna, A p53-dependent translational program directs tissue-selective phenotypes in a model of ribosomopathies. *Developmental Cell* **56**, 2089-2102.e2011 (2021); published online Epub2021/07/26/ (https://doi.org/10.1016/j.devcel.2021.06.013).
76. P. Ulz, E. Heitzer, M. R. Speicher, Co-occurrence of MYC amplification and TP53 mutations in human cancer. *Nature Genetics* **48**, 104-106 (2016); published online Epub2016/02/01 (10.1038/ng.3468).
77. A. Stengel, S. Schnittger, S. Weissmann, S. Kuznia, W. Kern, A. Kohlmann, T. Haferlach, C. Haferlach, TP53 mutations occur in 15.7% of ALL and are associated with MYC-rearrangement, low hypodiploidy, and a poor prognosis. *Blood* **124**, 251-258 (2014)10.1182/blood-2014-02-558833).
78. A. Bastide, A. David, The ribosome, (slow) beating heart of cancer (stem) cell. *Oncogenesis* **7**, 34 (2018); published online Epub2018/04/20 (10.1038/s41389-018-0044-8).
79. A. Artero-Castro, H. Kondoh, P. J. Fernández-Marcos, M. Serrano, S. Ramón y Cajal, M. E. Lleonart, Rplp1 bypasses replicative senescence and contributes to transformation. *Exp Cell Res* **315**, 1372-1383 (2009); published online EpubMay 1 (10.1016/j.yexcr.2009.02.007).
80. J. H. Kim, K. R. You, I. H. Kim, B. H. Cho, C. Y. Kim, D. G. Kim, Over-expression of the ribosomal protein L36a gene is associated with cellular proliferation in hepatocellular carcinoma. *Hepatology* **39**, 129-138 (2004); published online EpubJan (10.1002/hep.20017).
81. Y. Shirakawa, T. Hide, M. Yamaoka, Y. Ito, N. Ito, K. Ohta, N. Shinojima, A. Mukasa, H. Saito, H. Jono, Ribosomal protein S6 promotes stem-like characters in glioma cells. *Cancer Sci* **111**, 2041-2051 (2020); published online EpubJun (10.1111/cas.14399).
82. R. Y. Ebright, S. Lee, B. S. Wittner, K. L. Niederhoffer, B. T. Nicholson, A. Bardia, S. Truesdell, D. F. Wiley, B. Wesley, S. Li, A. Mai, N. Aceto, N. Vincent-Jordan, A. Szabolcs, B. Chirn, J. Kreuzer, V. Comaills, M. Kalinich, W. Haas, D. T. Ting, M. Toner, S. Vasudevan, D. A. Haber, S. Maheswaran, D. S. Micalizzi, Deregulation of ribosomal protein expression and translation promotes breast cancer metastasis. *Science* **367**, 1468-1473 (2020); published online EpubMar 27 (10.1126/science.aay0939).
83. J. He, L. Shen, M. Wan, O. Taranova, H. Wu, Y. Zhang, Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nat Cell Biol* **15**, 373-384 (2013); published online EpubApr (10.1038/ncb2702).
84. G. Liang, J. He, Y. Zhang, Kdm2b promotes induced pluripotent stem cell generation by facilitating gene activation early in reprogramming. *Nat Cell Biol* **14**, 457-466 (2012); published online EpubApr 22 (10.1038/ncb2483).
85. F. Kottakis, P. Foltopoulou, I. Sanidas, P. Keller, A. Wronski, B. T. Dake, S. A. Ezell, Z. Shen, S. P. Naber, P. W. Hinds, E. McNiel, C. Kuperwasser, P. N. Tsichlis, NDY1/KDM2B functions as a master regulator of polycomb complexes and controls self-renewal of breast cancer stem cells. *Cancer Res* **74**, 3935-3946 (2014); published online EpubJul 15 (10.1158/0008-5472.Can-13-2733).
86. A. Tzatsos, P. Paskaleva, F. Ferrari, V. Deshpande, S. Stoykova, G. Contino, K. K. Wong, F. Lan, P. Trojer, P. J. Park, N. Bardeesy, KDM2B promotes pancreatic

cancer via Polycomb-dependent and -independent transcriptional programs. *J Clin Invest* **123**, 727-739 (2013); published online EpubFeb (10.1172/jci64535).
87. I. C. Kurt, I. Sur, E. Kaya, A. Cingoz, S. Kazancioglu, Z. Kahya, O. D. Toparlak, F. Senbabaoglu, Z. Kaya, E. Ozyerli, S. Karahuseyinoglu, N. A. Lack, Z. H. Gumus, T. T. Onder, T. Bagci-Onder, KDM2B, an H3K36-specific demethylase, regulates apoptotic response of GBM cells to TRAIL. *Cell Death Dis* **8**, e2897 (2017); published online EpubJun 29 (10.1038/cddis.2017.288).
88. M. Yan, X. Yang, H. Wang, Q. Shao, The critical role of histone lysine demethylase KDM2B in cancer. *Am J Transl Res* **10**, 2222-2233 (2018).
89. R. Pfau, A. Tzatsos, S. C. Kampranis, O. B. Serebrennikova, S. E. Bear, P. N. Tsichlis, Members of a family of JmjC domain-containing oncoproteins immortalize embryonic fibroblasts via a JmjC domain-dependent process. *Proc Natl Acad Sci U S A* **105**, 1907-1912 (2008); published online EpubFeb 12 (10.1073/pnas.0711865105).
90. J. He, E. M. Kallin, Y.-I. Tsukada, Y. Zhang, The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15(Ink4b). *Nature structural & molecular biology* **15**, 1169-1175 (2008)10.1038/nsmb.1499).
91. J. Y. Kang, J. Y. Kim, K. B. Kim, J. W. Park, H. Cho, J. Y. Hahm, Y. C. Chae, D. Kim, H. Kook, S. Rhee, N. C. Ha, S. B. Seo, KDM2B is a histone H3K79 demethylase and induces transcriptional repression via sirtuin-1-mediated chromatin silencing. *Faseb j* **32**, 5737-5750 (2018); published online EpubOct (10.1096/fj.201800242R).
92. Z. Zhou, X. Yang, J. He, J. Liu, F. Wu, S. Yu, Y. Liu, R. Lin, H. Liu, Y. Cui, C. Zhou, X. Wang, J. Wu, S. Cao, L. Guo, L. Lin, T. Wang, X. Peng, B. Qiang, A. P. Hutchins, D. Pei, J. Chen, Kdm2b Regulates Somatic Reprogramming through Variant PRC1 Complex-Dependent Function. *Cell Rep* **21**, 2160-2170 (2017); published online EpubNov 21 (10.1016/j.celrep.2017.10.091).
93. A. M. Farcas, N. P. Blackledge, I. Sudbery, H. K. Long, J. F. McGouran, N. R. Rose, S. Lee, D. Sims, A. Cerase, T. W. Sheahan, H. Koseki, N. Brockdorff, C. P. Ponting, B. M. Kessler, R. J. Klose, KDM2B links the Polycomb Repressive Complex 1 (PRC1) to recognition of CpG islands. *Elife* **1**, e00205 (2012); published online EpubDec 18 (10.7554/eLife.00205).
94. D. Huo, Z. Yu, R. Li, M. Gong, S. Sidoli, X. Lu, Y. Hou, Z. Dai, Y. Kong, G. Liu, O. N. Jensen, W. Xie, K. Helin, C. Xiong, G. Li, Y. Zhang, X. Wu, CpG island reconfiguration for the establishment and synchronization of polycomb functions upon exit from naive pluripotency. *Mol Cell* **82**, 1169-1185.e1167 (2022); published online EpubMar 17 (10.1016/j.molcel.2022.01.027).
95. M. J. Barrero, J. C. Izpisua Belmonte, Polycomb complex recruitment in pluripotent stem cells. *Nat Cell Biol* **15**, 348-350 (2013); published online EpubApr (10.1038/ncb2723).
96. T. Inagaki, S. Iwasaki, Y. Matsumura, T. Kawamura, T. Tanaka, Y. Abe, A. Yamasaki, Y. Tsurutani, A. Yoshida, Y. Chikaoka, K. Nakamura, K. Magoori, R. Nakaki, T. F. Osborne, K. Fukami, H. Aburatani, T. Kodama, J. Sakai, The FBXL10/KDM2B scaffolding protein associates with novel polycomb repressive

complex-1 to regulate adipogenesis. *J Biol Chem* **290**, 4163-4177 (2015); published online EpubFeb 13 (10.1074/jbc.M114.626929).
97. A. Banito, X. Li, A. N. Laporte, J. S. Roe, F. Sanchez-Vega, C. H. Huang, A. R. Dancsok, K. Hatzi, C. C. Chen, D. F. Tschaharganeh, R. Chandwani, N. Tasdemir, K. B. Jones, M. R. Capecchi, C. R. Vakoc, N. Schultz, M. Ladanyi, T. O. Nielsen, S. W. Lowe, The SS18-SSX Oncoprotein Hijacks KDM2B-PRC1.1 to Drive Synovial Sarcoma. *Cancer Cell* **33**, 527-541.e528 (2018); published online EpubMar 12 (10.1016/j.ccell.2018.01.018).
98. H. L. Chan, F. Beckedorff, Y. Zhang, J. Garcia-Huidobro, H. Jiang, A. Colaprico, D. Bilbao, M. E. Figueroa, J. LaCava, R. Shiekhattar, L. Morey, Polycomb complexes associate with enhancers and promote oncogenic transcriptional programs in cancer through multiple mechanisms. *Nature Communications* **9**, 3377 (2018); published online Epub2018/08/23 (10.1038/s41467-018-05728-x).
99. N. Giner-Laguarda, M. Vidal, Functions of Polycomb Proteins on Active Targets. *Epigenomes* **4**, (2020); published online EpubAug 17 (10.3390/epigenomes4030017).
100. V. Parreno, A.-M. Martinez, G. Cavalli, Mechanisms of Polycomb group protein function in cancer. *Cell Research* **32**, 231-253 (2022); published online Epub2022/03/01 (10.1038/s41422-021-00606-6).
101. Z. Geng, Z. Gao, Mammalian PRC1 Complexes: Compositional Complexity and Diverse Molecular Mechanisms. *Int J Mol Sci* **21**, (2020); published online EpubNov 14 (10.3390/ijms21228594).
102. V. van den Boom, H. Maat, M. Geugien, A. Rodriguez Lopez, A. M. Sotoca, J. Jaques, A. Z. Brouwers-Vos, F. Fusetti, R. W. Groen, H. Yuan, A. C. Martens, H. G. Stunnenberg, E. Vellenga, J. H. Martens, J. J. Schuringa, Non-canonical PRC1.1 Targets Active Genes Independent of H3K27me3 and Is Essential for Leukemogenesis. *Cell Rep* **14**, 332-346 (2016); published online EpubJan 12 (10.1016/j.celrep.2015.12.034).
103. M. Boulard, J. R. Edwards, T. H. Bestor, FBXL10 protects Polycomb-bound genes from hypermethylation. *Nat Genet* **47**, 479-485 (2015); published online EpubMay (10.1038/ng.3272).
104. R. H. van Jaarsveld, J. Reilly, M. C. Cornips, M. A. Hadders, E. Agolini, P. Ahimaz, K. Anyane-Yeboa, S. A. Bellanger, E. van Binsbergen, M. J. van den Boogaard, E. Brischoux-Boucher, R. C. Caylor, A. Ciolfi, T. A. J. van Essen, P. Fontana, S. Hopman, M. Iascone, M. M. Javier, E. J. Kamsteeg, J. Kerkhof, J. Kido, H. G. Kim, T. Kleefstra, F. Lonardo, A. Lai, D. Lev, M. A. Levy, M. E. S. Lewis, A. Lichty, M. Mannens, N. Matsumoto, I. Maya, H. McConkey, A. Megarbane, V. Michaud, E. Miele, M. Niceta, A. Novelli, R. Onesimo, R. Pfundt, B. Popp, E. Prijoles, R. Relator, S. Redon, D. Rots, K. Rouault, K. Saida, J. Schieving, M. Tartaglia, R. Tenconi, K. Uguen, N. Verbeek, C. A. Walsh, K. Yosovich, C. J. Yuskaitis, G. Zampino, B. Sadikovic, M. Alders, R. Oegema, Delineation of a KDM2B-related neurodevelopmental disorder and its associated DNA methylation signature. *Genet Med* **25**, 49-62 (2023); published online EpubJan (10.1016/j.gim.2022.09.006).

105. Q. Zhuang, T. Feng, M. L. Coleman, Modifying the maker: Oxygenases target ribosome biology. *Translation (Austin)* **3**, e1009331 (2015); published online EpubJan-Jun (10.1080/21690731.2015.1009331).
106. E. M. Cornett, L. Ferry, P. A. Defossez, S. B. Rothbart, Lysine Methylation Regulators Moonlighting outside the Epigenome. *Mol Cell* **75**, 1092-1101 (2019); published online EpubSep 19 (10.1016/j.molcel.2019.08.026).
107. S. B. Rothbart, B. M. Dickson, B. D. Strahl, From histones to ribosomes: a chromatin regulator tangoes with translation. *Cancer Discov* **5**, 228-230 (2015); published online EpubMar (10.1158/2159-8290.Cd-15-0073).
108. C. Van Rechem, J. C. Black, M. Boukhali, M. J. Aryee, S. Graslund, W. Haas, C. H. Benes, J. R. Whetstine, Lysine demethylase KDM4A associates with translation machinery and regulates protein synthesis. *Cancer Discov* **5**, 255-263 (2015); published online EpubMar (10.1158/2159-8290.Cd-14-1326).
109. T. Hamidi, A. K. Singh, N. Veland, V. Vemulapalli, J. Chen, S. Hardikar, J. Bao, C. J. Fry, V. Yang, K. A. Lee, A. Guo, C. H. Arrowsmith, M. T. Bedford, T. Chen, Identification of Rpl29 as a major substrate of the lysine methyltransferase Set7/9. *The Journal of biological chemistry* **293**, 12770-12780 (2018)10.1074/jbc.RA118.002890).
110. Y. Hori, C. Engel, T. Kobayashi, Regulation of ribosomal RNA gene copy number, transcription and nucleolus organization in eukaryotes. *Nat Rev Mol Cell Biol*, (2023); published online EpubFeb 2 (10.1038/s41580-022-00573-9).
111. L. Morera, M. Lubbert, M. Jung, Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. *Clin Epigenetics* **8**, 57 (2016)10.1186/s13148-016-0223-4).
112. D. Frescas, D. Guardavaccaro, F. Bassermann, R. Koyama-Nasu, M. Pagano, JHDM1B/FBXL10 is a nucleolar protein that represses transcription of ribosomal RNA genes. *Nature* **450**, 309-313 (2007); published online EpubNov 8 (10.1038/nature06255).
113. M. Penzo, L. Casoli, D. Pollutri, L. Sicuro, C. Ceccarelli, D. Santini, M. Taffurelli, M. Govoni, D. Brina, D. Trere, L. Montanaro, JHDM1B expression regulates ribosome biogenesis and cancer cell growth in a p53 dependent manner. *Int J Cancer* **136**, E272-281 (2015); published online EpubMar 1 (10.1002/ijc.29240).
114. A. Galbiati, M. Penzo, M. G. Bacalini, C. Onofrillo, A. N. Guerrieri, P. Garagnani, C. Franceschi, D. Trere, L. Montanaro, Epigenetic up-regulation of ribosome biogenesis and more aggressive phenotype triggered by the lack of the histone demethylase JHDM1B in mammary epithelial cells. *Oncotarget* **8**, 37091-37103 (2017); published online EpubJun 6 (10.18632/oncotarget.16181).
115. S. J. Wong, M. D. Gearhart, A. B. Taylor, D. R. Nanyes, D. J. Ha, A. K. Robinson, J. A. Artigas, O. J. Lee, B. Demeler, P. J. Hart, V. J. Bardwell, C. A. Kim, KDM2B Recruitment of the Polycomb Group Complex, PRC1.1, Requires Cooperation between PCGF1 and BCORL1. *Structure* **24**, 1795-1801 (2016); published online EpubOct 4 (10.1016/j.str.2016.07.011).
116. J. G. P. Sanches, B. Song, Q. Zhang, X. Cui, I. B. Yabasin, M. Ntim, X. Li, J. He, Y. Zhang, J. Mao, Y. Lu, L. Li, The Role of KDM2B and EZH2 in Regulating the

Stemness in Colorectal Cancer Through the PI3K/AKT Pathway. *Front Oncol* **11**, 637298 (2021)10.3389/fonc.2021.637298).
117. H. Shima, M. Pende, Y. Chen, S. Fumagalli, G. Thomas, S. C. Kozma, Disruption of the p70(s6k)/p85(s6k) gene reveals a small mouse phenotype and a new functional S6 kinase. *Embo j* **17**, 6649-6659 (1998); published online EpubNov 16 (10.1093/emboj/17.22.6649).
118. Q. Zheng, H. Fan, Z. Meng, L. Yuan, C. Liu, Y. Peng, W. Zhao, L. Wang, J. Li, J. Feng, Histone demethylase KDM2B promotes triple negative breast cancer proliferation by suppressing p15INK4B, p16INK4A, and p57KIP2 transcription. *Acta Biochim Biophys Sin (Shanghai)* **50**, 897-904 (2018); published online EpubSep 1 (10.1093/abbs/gmy084).
119. Y. Wang, J. Zang, D. Zhang, Z. Sun, B. Qiu, X. Wang, KDM2B overexpression correlates with poor prognosis and regulates glioma cell growth. *Onco Targets Ther* **11**, 201-209 (2018)10.2147/ott.S149833).
120. S. Wanna-Udom, M. Terashima, K. Suphakhong, A. Ishimura, T. Takino, T. Suzuki, KDM2B is involved in the epigenetic regulation of TGF-β-induced epithelial-mesenchymal transition in lung and pancreatic cancer cell lines. *J Biol Chem* **296**, 100213 (2021); published online EpubJan-Jun (10.1074/jbc.RA120.015502).
121. A. Tsherniak, F. Vazquez, P. G. Montgomery, B. A. Weir, G. Kryukov, G. S. Cowley, S. Gill, W. F. Harrington, S. Pantel, J. M. Krill-Burger, R. M. Meyers, L. Ali, A. Goodale, Y. Lee, G. Jiang, J. Hsiao, W. F. J. Gerath, S. Howell, E. Merkel, M. Ghandi, L. A. Garraway, D. E. Root, T. R. Golub, J. S. Boehm, W. C. Hahn, Defining a Cancer Dependency Map. *Cell* **170**, 564-576.e516 (2017); published online EpubJul 27 (10.1016/j.cell.2017.06.010).
122. K. Shimada, J. A. Bachman, J. L. Muhlich, T. J. Mitchison, shinyDepMap, a tool to identify targetable cancer genes and their functional connections from Cancer Dependency Map data. *eLife* **10**, e57116 (2021); published online Epub2021/02/08 (10.7554/eLife.57116).
123. J. He, E. M. Kallin, Y. Tsukada, Y. Zhang, The H3K36 demethylase Jhdm1b/Kdm2b regulates cell proliferation and senescence through p15(Ink4b). *Nat Struct Mol Biol* **15**, 1169-1175 (2008); published online EpubNov (10.1038/nsmb.1499).
124. J. Andricovich, Y. Kai, W. Peng, A. Foudi, A. Tzatsos, Histone demethylase KDM2B regulates lineage commitment in normal and malignant hematopoiesis. *J Clin Invest* **126**, 905-920 (2016); published online EpubMar 1 (10.1172/jci84014).
125. Y. Kuang, H. Xu, F. Lu, J. Meng, Y. Yi, H. Yang, H. Hou, H. Wei, S. Su, Inhibition of microRNA let-7b expression by KDM2B promotes cancer progression by targeting EZH2 in ovarian cancer. *Cancer Sci* **112**, 231-242 (2021); published online EpubJan (10.1111/cas.14708).
126. A. Tzatsos, P. Paskaleva, S. Lymperi, G. Contino, S. Stoykova, Z. Chen, K. K. Wong, N. Bardeesy, Lysine-specific demethylase 2B (KDM2B)-let-7-enhancer of zester homolog 2 (EZH2) pathway regulates cell cycle progression and senescence in primary cells. *J Biol Chem* **286**, 33061-33069 (2011); published online EpubSep 23 (10.1074/jbc.M111.257667).

127. Y. Isshiki, Y. Nakajima-Takagi, M. Oshima, K. Aoyama, M. Rizk, S. Kurosawa, A. Saraya, T. Kondo, E. Sakaida, C. Nakaseko, K. Yokote, H. Koseki, A. Iwama, KDM2B in polycomb repressive complex 1.1 functions as a tumor suppressor in the initiation of T-cell leukemogenesis. *Blood Adv* **3**, 2537-2549 (2019); published online EpubSep 10 (10.1182/bloodadvances.2018028522).
128. J. M. Lebert, R. Lester, E. Powell, M. Seal, J. McCarthy, Advances in the systemic treatment of triple-negative breast cancer. *Curr Oncol* **25**, S142-s150 (2018); published online EpubJun (10.3747/co.25.3954).
129. C. Vaklavas, S. W. Blume, W. E. Grizzle, Translational Dysregulation in Cancer: Molecular Insights and Potential Clinical Applications in Biomarker Development. *Front Oncol* **7**, 158 (2017)10.3389/fonc.2017.00158).
130. B. Liu, S. B. Qian, Translational reprogramming in cellular stress response. *Wiley Interdiscip Rev RNA* **5**, 301-315 (2014); published online EpubMay-Jun (10.1002/wrna.1212).
131. M. Sauert, H. Temmel, I. Moll, Heterogeneity of the translational machinery: Variations on a common theme. *Biochimie* **114**, 39-47 (2015); published online EpubJul (10.1016/j.biochi.2014.12.011).
132. V. K. Rajasekhar, A. Viale, N. D. Socci, M. Wiedmann, X. Hu, E. C. Holland, Oncogenic Ras and Akt signaling contribute to glioblastoma formation by differential recruitment of existing mRNAs to polysomes. *Mol Cell* **12**, 889-901 (2003); published online EpubOct (10.1016/s1097-2765(03)00395-2).
133. J. R. Masters, J. A. Thomson, B. Daly-Burns, Y. A. Reid, W. G. Dirks, P. Packer, L. H. Toji, T. Ohno, H. Tanabe, C. F. Arlett, L. R. Kelland, M. Harrison, A. Virmani, T. H. Ward, K. L. Ayres, P. G. Debenham, Short tandem repeat profiling provides an international reference standard for human cell lines. *Proc Natl Acad Sci U S A* **98**, 8012-8017 (2001); published online EpubJul 3 (10.1073/pnas.121616198).
134. J. L. Almeida, C. T. Korch, in *Assay Guidance Manual,* S. Markossian, A. Grossman, K. Brimacombe, M. Arkin, D. Auld, C. Austin, J. Baell, T. D. Y. Chung, N. P. Coussens, J. L. Dahlin, V. Devanarayan, T. L. Foley, M. Glicksman, K. Gorshkov, J. V. Haas, M. D. Hall, S. Hoare, J. Inglese, P. W. Iversen, S. C. Kales, M. Lal-Nag, Z. Li, J. McGee, O. McManus, T. Riss, P. Saradjian, G. S. Sittampalam, M. Tarselli, O. J. Trask, Jr., Y. Wang, J. R. Weidner, M. J. Wildey, K. Wilson, M. Xia, X. Xu, Eds. (Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), 2004).
135. S. Andrews, in *FastQC: a quality control tool for high throughput sequence data.,* B. Bioinformatics, Ed. (Babraham Bioinformatics, 2010).
136. P. Ewels, M. Magnusson, S. Lundin, M. Käller, MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047-3048 (2016)10.1093/bioinformatics/btw354).
137. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, 3 (2011); published online Epub2011-05-02 (10.14806/ej.17.1.200).
138. A. La Ferlita, S. Alaimo, S. Di Bella, E. Martorana, G. I. Laliotis, F. Bertoni, L. Cascione, P. N. Tsichlis, A. Ferro, R. Bosotti, A. Pulvirenti, RNAdetector: a free user-friendly stand-alone and cloud-based system for RNA-Seq data analysis.

BMC Bioinformatics **22**, 298 (2021); published online Epub2021/06/03 (10.1186/s12859-021-04211-7).
139. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2012)10.1093/bioinformatics/bts635).
140. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2013)10.1093/bioinformatics/btt656).
141. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014); published online Epub2014/12/05 (10.1186/s13059-014-0550-8).
142. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2009)10.1093/bioinformatics/btp616).
143. M. E. Ritchie, B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi, G. K. Smyth, limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47-e47 (2015)10.1093/nar/gkv007).
144. P. Moulos, P. Hatzis, Systematic integration of RNA-Seq statistical algorithms for accurate detection of differential gene expression patterns. *Nucleic Acids Research* **43**, e25-e25 (2014)10.1093/nar/gku1273).
145. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* **102**, 15545-15550 (2005)doi:10.1073/pnas.0506580102).
146. V. K. Mootha, C. M. Lindgren, K.-F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstråle, E. Laurila, N. Houstis, M. J. Daly, N. Patterson, J. P. Mesirov, T. R. Golub, P. Tamayo, B. Spiegelman, E. S. Lander, J. N. Hirschhorn, D. Altshuler, L. C. Groop, PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics* **34**, 267-273 (2003); published online Epub2003/07/01 (10.1038/ng1180).
147. M. A. Harris, J. Clark, A. Ireland, J. Lomax, M. Ashburner, R. Foulger, K. Eilbeck, S. Lewis, B. Marshall, C. Mungall, J. Richter, G. M. Rubin, J. A. Blake, C. Bult, M. Dolan, H. Drabkin, J. T. Eppig, D. P. Hill, L. Ni, M. Ringwald, R. Balakrishnan, J. M. Cherry, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. Engel, D. G. Fisk, J. E. Hirschman, E. L. Hong, R. S. Nash, A. Sethuraman, C. L. Theesfeld, D. Botstein, K. Dolinski, B. Feierbach, T. Berardini, S. Mundodi, S. Y. Rhee, R. Apweiler, D. Barrell, E. Camon, E. Dimmer, V. Lee, R. Chisholm, P. Gaudet, W. Kibbe, R. Kishore, E. M. Schwarz, P. Sternberg, M. Gwinn, L. Hannick, J. Wortman, M. Berriman, V. Wood, N. de la Cruz, P. Tonellato, P. Jaiswal, T. Seigfried, R. White, The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* **32**, D258-261 (2004); published online EpubJan 1 (10.1093/nar/gkh036).

148. D. Gaidatzis, L. Burger, M. Florescu, M. B. Stadler, Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. *Nature Biotechnology* **33**, 722-729 (2015); published online Epub2015/07/01 (10.1038/nbt.3269).
149. K. A. Pillman, K. G. Scheer, E. Hackett-Jones, K. Saunders, A. G. Bert, J. Toubia, H. J. Whitfield, S. Sapkota, L. Sourdin, H. Pham, T. D. Le, J. Cursons, M. J. Davis, P. A. Gregory, G. J. Goodall, C. P. Bracken, Extensive transcriptional responses are co-ordinated by microRNAs as revealed by Exon–Intron Split Analysis (EISA). *Nucleic Acids Research* **47**, 8606-8619 (2019)10.1093/nar/gkz664).
150. J. Sirén, N. Välimäki, V. Mäkinen, Indexing graphs for path queries with applications in genome research. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* **11**, 375–388 (2014)10.1109/tcbb.2013.2297101).
151. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012); published online EpubMar 4 (10.1038/nmeth.1923).
152. J. Feng, T. Liu, B. Qin, Y. Zhang, X. S. Liu, Identifying ChIP-seq enrichment using MACS. *Nat Protoc* **7**, 1728-1740 (2012); published online EpubSep (10.1038/nprot.2012.101).
153. L. J. Zhu, C. Gazin, N. D. Lawson, H. Pagès, S. M. Lin, D. S. Lapointe, M. R. Green, ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data. *BMC Bioinformatics* **11**, 237 (2010); published online Epub2010/05/11 (10.1186/1471-2105-11-237).
154. G. Yu, L.-G. Wang, Q.-Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382-2383 (2015)10.1093/bioinformatics/btv145).
155. G. Georgiou, S. J. van Heeringen, fluff: exploratory analysis and visualization of high-throughput sequencing data. *PeerJ* **4**, e2209 (2016)10.7717/peerj.2209).
156. R. Stark, G. Brown. (Bioconductor, 2011), vol. 2022.
157. K. Chang, C. J. Creighton, C. Davis, L. Donehower, J. Drummond, D. Wheeler, A. Ally, M. Balasundaram, I. Birol, Y. S. N. Butterfield, A. Chu, E. Chuah, H.-J. E. Chun, N. Dhalla, R. Guin, M. Hirst, C. Hirst, R. A. Holt, S. J. M. Jones, D. Lee, H. I. Li, M. A. Marra, M. Mayo, R. A. Moore, A. J. Mungall, A. G. Robertson, J. E. Schein, P. Sipahimalani, A. Tam, N. Thiessen, R. J. Varhol, R. Beroukhim, A. S. Bhatt, A. N. Brooks, A. D. Cherniack, S. S. Freeman, S. B. Gabriel, E. Helman, J. Jung, M. Meyerson, A. I. Ojesina, C. S. Pedamallu, G. Saksena, S. E. Schumacher, B. Tabak, T. Zack, E. S. Lander, C. A. Bristow, A. Hadjipanayis, P. Haseley, R. Kucherlapati, S. Lee, E. Lee, L. J. Luquette, H. S. Mahadeshwar, A. Pantazi, M. Parfenov, P. J. Park, A. Protopopov, X. Ren, N. Santoso, J. Seidman, S. Seth, X. Song, J. Tang, R. Xi, A. W. Xu, L. Yang, D. Zeng, J. T. Auman, S. Balu, E. Buda, C. Fan, K. A. Hoadley, C. D. Jones, S. Meng, P. A. Mieczkowski, J. S. Parker, C. M. Perou, J. Roach, Y. Shi, G. O. Silva, D. Tan, U. Veluvolu, S. Waring, M. D. Wilkerson, J. Wu, W. Zhao, T. Bodenheimer, D. N. Hayes, A. P. Hoyle, S. R. Jeffreys, L. E. Mose, J. V. Simons, M. G. Soloway, S. B. Baylin, B. P. Berman, M. S. Bootwalla, L. Danilova, J. G. Herman, T. Hinoue, P. W. Laird, S. K. Rhie, H. Shen, T. Triche, D. J. Weisenberger, S. L. Carter, K. Cibulskis, L. Chin, J. Zhang, G. Getz, C. Sougnez,

M. Wang, G. Saksena, S. L. Carter, K. Cibulskis, L. Chin, J. Zhang, G. Getz, H. Dinh, H. V. Doddapaneni, R. Gibbs, P. Gunaratne, Y. Han, D. Kalra, C. Kovar, L. Lewis, M. Morgan, D. Morton, D. Muzny, J. Reid, L. Xi, J. Cho, D. DiCara, S. Frazer, N. Gehlenborg, D. I. Heiman, J. Kim, M. S. Lawrence, P. Lin, Y. Liu, M. S. Noble, P. Stojanov, D. Voet, H. Zhang, L. Zou, C. Stewart, B. Bernard, R. Bressler, A. Eakin, L. Iype, T. Knijnenburg, R. Kramer, R. Kreisberg, K. Leinonen, J. Lin, Y. Liu, M. Miller, S. M. Reynolds, H. Rovira, I. Shmulevich, V. Thorsson, D. Yang, W. Zhang, S. Amin, C.-J. Wu, C.-C. Wu, R. Akbani, K. Aldape, K. A. Baggerly, B. Broom, T. D. Casasent, J. Cleland, C. Creighton, D. Dodda, M. Edgerton, L. Han, S. M. Herbrich, Z. Ju, H. Kim, S. Lerner, J. Li, H. Liang, W. Liu, P. L. Lorenzi, Y. Lu, J. Melott, G. B. Mills, L. Nguyen, X. Su, R. Verhaak, W. Wang, J. N. Weinstein, A. Wong, Y. Yang, J. Yao, R. Yao, K. Yoshihara, Y. Yuan, A. K. Yung, N. Zhang, S. Zheng, M. Ryan, D. W. Kane, B. A. Aksoy, G. Ciriello, G. Dresdner, J. Gao, B. Gross, A. Jacobsen, A. Kahles, M. Ladanyi, W. Lee, K.-V. Lehmann, M. L. Miller, R. Ramirez, G. Rätsch, B. Reva, C. Sander, N. Schultz, Y. Senbabaoglu, R. Shen, R. Sinha, S. O. Sumer, Y. Sun, B. S. Taylor, N. Weinhold, S. Fei, P. Spellman, C. Benz, D. Carlin, M. Cline, B. Craft, K. Ellrott, M. Goldman, D. Haussler, S. Ma, S. Ng, E. Paull, A. Radenbaugh, S. Salama, A. Sokolov, J. M. Stuart, T. Swatloski, V. Uzunangelov, P. Waltman, C. Yau, J. Zhu, S. R. Hamilton, G. Getz, C. Sougnez, S. Abbott, R. Abbott, N. D. Dees, K. Delehaunty, L. Ding, D. J. Dooling, J. M. Eldred, C. C. Fronick, R. Fulton, L. L. Fulton, J. Kalicki-Veizer, K.-L. Kanchi, C. Kandoth, D. C. Koboldt, D. E. Larson, T. J. Ley, L. Lin, C. Lu, V. J. Magrini, E. R. Mardis, M. D. McLellan, J. F. McMichael, C. A. Miller, M. O'Laughlin, C. Pohl, H. Schmidt, S. M. Smith, J. Walker, J. W. Wallis, M. C. Wendl, R. K. Wilson, T. Wylie, Q. Zhang, R. Burton, M. A. Jensen, A. Kahn, T. Pihl, D. Pot, Y. Wan, D. A. Levine, A. D. Black, J. Bowen, N. The Cancer Genome Atlas Research, C. Genome Characterization, C. Genome Data Analysis, C. Sequencing, C. Data Coordinating, S. Tissue Source, C. Biospecimen Core Resource, The Cancer Genome Atlas Pan-Cancer analysis project. *Nature Genetics* **45**, 1113-1120 (2013); published online Epub2013/10/01 (10.1038/ng.2764).

158. J. Lonsdale, J. Thomas, M. Salvatore, R. Phillips, E. Lo, S. Shad, R. Hasz, G. Walters, F. Garcia, N. Young, B. Foster, M. Moser, E. Karasik, B. Gillard, K. Ramsey, S. Sullivan, J. Bridge, H. Magazine, J. Syron, J. Fleming, L. Siminoff, H. Traino, M. Mosavel, L. Barker, S. Jewell, D. Rohrer, D. Maxim, D. Filkins, P. Harbach, E. Cortadillo, B. Berghuis, L. Turner, E. Hudson, K. Feenstra, L. Sobin, J. Robb, P. Branton, G. Korzeniewski, C. Shive, D. Tabor, L. Qi, K. Groch, S. Nampally, S. Buia, A. Zimmerman, A. Smith, R. Burges, K. Robinson, K. Valentino, D. Bradbury, M. Cosentino, N. Diaz-Mayoral, M. Kennedy, T. Engel, P. Williams, K. Erickson, K. Ardlie, W. Winckler, G. Getz, D. DeLuca, D. MacArthur, M. Kellis, A. Thomson, T. Young, E. Gelfand, M. Donovan, Y. Meng, G. Grant, D. Mash, Y. Marcus, M. Basile, J. Liu, J. Zhu, Z. Tu, N. J. Cox, D. L. Nicolae, E. R. Gamazon, H. K. Im, A. Konkashbaev, J. Pritchard, M. Stevens, T. Flutre, X. Wen, E. T. Dermitzakis, T. Lappalainen, R. Guigo, J. Monlong, M. Sammeth, D. Koller, A. Battle, S. Mostafavi, M. McCarthy, M. Rivas, J. Maller, I. Rusyn, A. Nobel, F. Wright, A. Shabalin, M. Feolo, N.

Sharopova, A. Sturcke, J. Paschal, J. M. Anderson, E. L. Wilder, L. K. Derr, E. D. Green, J. P. Struewing, G. Temple, S. Volpi, J. T. Boyer, E. J. Thomson, M. S. Guyer, C. Ng, A. Abdallah, D. Colantuoni, T. R. Insel, S. E. Koester, A. R. Little, P. K. Bender, T. Lehner, Y. Yao, C. C. Compton, J. B. Vaught, S. Sawyer, N. C. Lockhart, J. Demchok, H. F. Moore, The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**, 580-585 (2013); published online Epub2013/06/01 (10.1038/ng.2653).

159. M. J. Goldman, B. Craft, M. Hastie, K. Repečka, F. McDade, A. Kamath, A. Banerjee, Y. Luo, D. Rogers, A. N. Brooks, J. Zhu, D. Haussler, Visualizing and interpreting cancer genomics data via the Xena platform. *Nature Biotechnology* **38**, 675-678 (2020); published online Epub2020/06/01 (10.1038/s41587-020-0546-8).

160. J. H. Bullard, E. Purdom, K. D. Hansen, S. Dudoit, Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. *BMC Bioinformatics* **11**, 94 (2010); published online Epub2010/02/18 (10.1186/1471-2105-11-94).

161. J. Zyprych-Walczak, A. Szabelska, L. Handschuh, K. Górczak, K. Klamecka, M. Figlerowicz, I. Siatkowski, The Impact of Normalization Methods on RNA-Seq Data Analysis. *Biomed Res Int* **2015**, 621690 (2015)10.1155/2015/621690).

162. D. C. Koboldt, R. S. Fulton, M. D. McLellan, H. Schmidt, J. Kalicki-Veizer, J. F. McMichael, L. L. Fulton, D. J. Dooling, L. Ding, E. R. Mardis, R. K. Wilson, A. Ally, M. Balasundaram, Y. S. N. Butterfield, R. Carlsen, C. Carter, A. Chu, E. Chuah, H.-J. E. Chun, R. J. N. Coope, N. Dhalla, R. Guin, C. Hirst, M. Hirst, R. A. Holt, D. Lee, H. I. Li, M. Mayo, R. A. Moore, A. J. Mungall, E. Pleasance, A. Gordon Robertson, J. E. Schein, A. Shafiei, P. Sipahimalani, J. R. Slobodan, D. Stoll, A. Tam, N. Thiessen, R. J. Varhol, N. Wye, T. Zeng, Y. Zhao, I. Birol, S. J. M. Jones, M. A. Marra, A. D. Cherniack, G. Saksena, R. C. Onofrio, N. H. Pho, S. L. Carter, S. E. Schumacher, B. Tabak, B. Hernandez, J. Gentry, H. Nguyen, A. Crenshaw, K. Ardlie, R. Beroukhim, W. Winckler, G. Getz, S. B. Gabriel, M. Meyerson, L. Chin, P. J. Park, R. Kucherlapati, K. A. Hoadley, J. Todd Auman, C. Fan, Y. J. Turman, Y. Shi, L. Li, M. D. Topal, X. He, H.-H. Chao, A. Prat, G. O. Silva, M. D. Iglesia, W. Zhao, J. Usary, J. S. Berg, M. Adams, J. Booker, J. Wu, A. Gulabani, T. Bodenheimer, A. P. Hoyle, J. V. Simons, M. G. Soloway, L. E. Mose, S. R. Jefferys, S. Balu, J. S. Parker, D. Neil Hayes, C. M. Perou, S. Malik, S. Mahurkar, H. Shen, D. J. Weisenberger, T. Triche Jr, P. H. Lai, M. S. Bootwalla, D. T. Maglinte, B. P. Berman, D. J. Van Den Berg, S. B. Baylin, P. W. Laird, C. J. Creighton, L. A. Donehower, G. Getz, M. Noble, D. Voet, G. Saksena, N. Gehlenborg, D. DiCara, J. Zhang, H. Zhang, C.-J. Wu, S. Yingchun Liu, M. S. Lawrence, L. Zou, A. Sivachenko, P. Lin, P. Stojanov, R. Jing, J. Cho, R. Sinha, R. W. Park, M.-D. Nazaire, J. Robinson, H. Thorvaldsdottir, J. Mesirov, P. J. Park, L. Chin, S. Reynolds, R. B. Kreisberg, B. Bernard, R. Bressler, T. Erkkila, J. Lin, V. Thorsson, W. Zhang, I. Shmulevich, G. Ciriello, N. Weinhold, N. Schultz, J. Gao, E. Cerami, B. Gross, A. Jacobsen, R. Sinha, B. Arman Aksoy, Y. Antipin, B. Reva, R. Shen, B. S. Taylor, M. Ladanyi, C. Sander, P. Anur, P. T. Spellman, Y. Lu, W. Liu, R. R. G. Verhaak, G. B. Mills, R. Akbani, N. Zhang, B. M. Broom, T. D. Casasent, C. Wakefield, A. K. Unruh, K. Baggerly, K.

Coombes, J. N. Weinstein, D. Haussler, C. C. Benz, J. M. Stuart, S. C. Benz, J. Zhu, C. C. Szeto, G. K. Scott, C. Yau, E. O. Paull, D. Carlin, C. Wong, A. Sokolov, J. Thusberg, S. Mooney, S. Ng, T. C. Goldstein, K. Ellrott, M. Grifford, C. Wilks, S. Ma, B. Craft, C. Yan, Y. Hu, D. Meerzaman, J. M. Gastier-Foster, J. Bowen, N. C. Ramirez, A. D. Black, R. E. Pyatt, P. White, E. J. Zmuda, J. Frick, T. M. Lichtenberg, R. Brookens, M. M. George, M. A. Gerken, H. A. Harper, K. M. Leraas, L. J. Wise, T. R. Tabler, C. McAllister, T. Barr, M. Hart-Kothari, K. Tarvin, C. Saller, G. Sandusky, C. Mitchell, M. V. Iacocca, J. Brown, B. Rabeno, C. Czerwinski, N. Petrelli, O. Dolzhansky, M. Abramov, O. Voronina, O. Potapova, J. R. Marks, W. M. Suchorska, D. Murawa, W. Kycler, M. Ibbs, K. Korski, A. Spychała, P. Murawa, J. J. Brzeziński, H. Perz, R. Łaźniak, M. Teresiak, H. Tatka, E. Leporowska, M. Bogusz-Czerniewicz, J. Malicki, A. Mackiewicz, M. Wiznerowicz, X. Van Le, B. Kohl, N. Viet Tien, R. Thorp, N. Van Bang, H. Sussman, B. Duc Phu, R. Hajek, N. Phi Hung, T. Viet The Phuong, H. Quyet Thang, K. Zaki Khan, R. Penny, D. Mallery, E. Curley, C. Shelton, P. Yena, J. N. Ingle, F. J. Couch, W. L. Lingle, T. A. King, A. Maria Gonzalez-Angulo, G. B. Mills, M. D. Dyer, S. Liu, X. Meng, M. Patangan, N. The Cancer Genome Atlas, L. Genome sequencing centres: Washington University in St, B. C. C. A. Genome characterization centres, I. Broad, Brigham, H. Women's, S. Harvard Medical, C. H. University of North Carolina, H. University of Southern California/Johns, M. Genome data analysis: Baylor College of, B. Institute for Systems, C. Memorial Sloan-Kettering Cancer, H. Oregon, U. Science, M. D. A. C. C. The University of Texas, S. C. B. I. University of California, Nci, R. Biospecimen core resource: Nationwide Children's Hospital Biospecimen Core, A.-I. Tissue source sites, Christiana, Cureline, C. Duke University Medical, C. The Greater Poland Cancer, Ilsbio, C. International Genomics, C. Mayo, Mskcc, M. D. A. C. Center, Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012); published online Epub2012/10/01 (10.1038/nature11412).

163. K. Krug, E. J. Jaehnig, S. Satpathy, L. Blumenberg, A. Karpova, M. Anurag, G. Miles, P. Mertins, Y. Geffen, L. C. Tang, D. I. Heiman, S. Cao, Y. E. Maruvka, J. T. Lei, C. Huang, R. B. Kothadia, A. Colaprico, C. Birger, J. Wang, Y. Dou, B. Wen, Z. Shi, Y. Liao, M. Wiznerowicz, M. A. Wyczalkowski, X. S. Chen, J. J. Kennedy, A. G. Paulovich, M. Thiagarajan, C. R. Kinsinger, T. Hiltke, E. S. Boja, M. Mesri, A. I. Robles, H. Rodriguez, T. F. Westbrook, L. Ding, G. Getz, K. R. Clauser, D. Fenyö, K. V. Ruggles, B. Zhang, D. R. Mani, S. A. Carr, M. J. Ellis, M. A. Gillette, Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. *Cell* **183**, 1436-1456.e1431 (2020); published online EpubNov 25 (10.1016/j.cell.2020.10.036).
164. E. Cerami, J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer, B. A. Aksoy, A. Jacobsen, C. J. Byrne, M. L. Heuer, E. Larsson, Y. Antipin, B. Reva, A. P. Goldberg, C. Sander, N. Schultz, The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* **2**, 401-404 (2012)10.1158/2159-8290.Cd-12-0095).
165. J. Gao, B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, N. Schultz, Integrative

analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* **6**, pl1 (2013); published online EpubApr 2 (10.1126/scisignal.2004088).
166. Y.-Z. Jiang, D. Ma, C. Suo, J. Shi, M. Xue, X. Hu, Y. Xiao, K.-D. Yu, Y.-R. Liu, Y. Yu, Y. Zheng, X. Li, C. Zhang, P. Hu, J. Zhang, Q. Hua, J. Zhang, W. Hou, L. Ren, D. Bao, B. Li, J. Yang, L. Yao, W.-J. Zuo, S. Zhao, Y. Gong, Y.-X. Ren, Y.-X. Zhao, Y.-S. Yang, Z. Niu, Z.-G. Cao, D. G. Stover, C. Verschraegen, V. Kaklamani, A. Daemen, J. R. Benson, K. Takabe, F. Bai, D.-Q. Li, P. Wang, L. Shi, W. Huang, Z.-M. Shao, Genomic and Transcriptomic Landscape of Triple-Negative Breast Cancers: Subtypes and Treatment Strategies. *Cancer Cell* **35**, 428-440.e425 (2019); published online Epub2019/03/18/ (https://doi.org/10.1016/j.ccell.2019.02.001).
167. H. An, A. Ordureau, M. Körner, J. A. Paulo, J. W. Harper, Systematic quantitative analysis of ribosome inventory during nutrient stress. *Nature* **583**, 303-309 (2020); published online EpubJul (10.1038/s41586-020-2446-y).
168. F. Zhang, in *Genome Engineering,* Addgene, Ed. (Addgene, Zhang Lab CRISPR Plasmids, 2014), vol. 2023.
169. F. A. Ran, P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, F. Zhang, Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281-2308 (2013); published online EpubNov (10.1038/nprot.2013.143).
170. D. Matsumoto, H. Tamamura, W. Nomura, A cell cycle-dependent CRISPR-Cas9 activation system based on an anti-CRISPR protein shows improved genome editing accuracy. *Communications Biology* **3**, 601 (2020); published online Epub2020/10/23 (10.1038/s42003-020-01340-2).
171. J. B. Grimm, L. Xie, J. C. Casler, R. Patel, A. N. Tkachuk, N. Falco, H. Choi, J. Lippincott-Schwartz, T. A. Brown, B. S. Glick, Z. Liu, L. D. Lavis, A General Method to Improve Fluorophores Using Deuterated Auxochromes. *JACS Au* **1**, 690-696 (2021); published online EpubMay 24 (10.1021/jacsau.1c00006).
172. L. D. Lavis, What if we just give everything away? *eLife* **10**, e74981 (2021); published online Epub2021/11/05 (10.7554/eLife.74981).
173. N. H. Cho, K. C. Cheveralls, A.-D. Brunner, K. Kim, A. C. Michaelis, P. Raghavan, H. Kobayashi, L. Savy, J. Y. Li, H. Canaj, J. Y. S. Kim, E. M. Stewart, C. Gnann, F. McCarthy, J. P. Cabrera, R. M. Brunetti, B. B. Chhun, G. Dingle, M. Y. Hein, B. Huang, S. B. Mehta, J. S. Weissman, R. Gómez-Sjöberg, D. N. Itzhak, L. A. Royer, M. Mann, M. D. Leonetti, OpenCell: Endogenous tagging for the cartography of human cellular organization. *Science* **375**, eabi6983 (2022)doi:10.1126/science.abi6983).
174. T. Kremmer, I. Pälyi, D. Daubner, M. Boldizsár, B. Vincze, E. Paulik, J. Sugár, E. Pokorny, E. Túry, Comparative studies on the polyamine metabolism and DFMO treatment of MCF-7 and MDA-MB-231 breast cancer cell lines and xenografts. *Anticancer Res* **11**, 1807-1813 (1991); published online EpubSep-Oct (
175. E. S. Pringle, C. McCormick, Z. Cheng, Polysome Profiling Analysis of mRNA and Associated Proteins Engaged in Translation. *Current Protocols in Molecular Biology* **125**, e79 (2019)https://doi.org/10.1002/cpmb.79).

176. L. Breuza, S. Poux, A. Estreicher, M. L. Famiglietti, M. Magrane, M. Tognolli, A. Bridge, D. Baratin, N. Redaschi, The UniProtKB guide to the human proteome. *Database (Oxford)* **2016**, (2016)10.1093/database/bav120).
177. M. V. Ivanov, L. I. Levitsky, M. V. Gorshkov, Adaptation of Decoy Fusion Strategy for Existing Multi-Stage Search Workflows. *Journal of The American Society for Mass Spectrometry* **27**, 1579-1582 (2016); published online Epub2016/09/01 (10.1007/s13361-016-1436-7).
178. S. Ito, S. Horikawa, T. Suzuki, H. Kawauchi, Y. Tanaka, T. Suzuki, T. Suzuki, Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). *J Biol Chem* **289**, 35724-35730 (2014); published online EpubDec 26 (10.1074/jbc.C114.602698).
179. J. Zhang, J. Wang, Y. M. Lee, T. K. Lim, Q. Lin, H. M. Shen, Proteomic Profiling of De Novo Protein Synthesis in Starvation-Induced Autophagy Using Bioorthogonal Noncanonical Amino Acid Tagging. *Methods Enzymol* **588**, 41-59 (2017)10.1016/bs.mie.2016.09.075).
180. N. T. Ingolia, G. A. Brar, S. Rouskin, A. M. McGeachy, J. S. Weissman, The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nature Protocols* **7**, 1534-1550 (2012); published online Epub2012/08/01 (10.1038/nprot.2012.086).
181. N. J. McGlincy, N. T. Ingolia, Transcriptome-wide measurement of translation by ribosome profiling. *Methods* **126**, 112-129 (2017); published online EpubAug 15 (10.1016/j.ymeth.2017.05.028).
182. F. Aeschimann, J. Xiong, A. Arnold, C. Dieterich, H. Großhans, Transcriptome-wide measurement of ribosomal occupancy by ribosome profiling. *Methods* **85**, 75-89 (2015); published online EpubSep 1 (10.1016/j.ymeth.2015.06.013).
183. J. C. Halpin, R. Jangi, T. O. Street, Multimapping confounds ribosome profiling analysis: A case-study of the Hsp90 molecular chaperone. *Proteins* **88**, 57-68 (2020); published online EpubJan (10.1002/prot.25766).
184. G. Ramadori, R. M. Ioris, Z. Villanyi, R. Firnkes, O. O. Panasenko, G. Allen, G. Konstantinidou, E. Aras, X. Brenachot, T. Biscotti, A. Charollais, M. Luchetti, F. Bezrukov, A. Santinelli, M. Samad, P. Baldi, M. A. Collart, R. Coppari, FKBP10 Regulates Protein Translation to Sustain Lung Cancer Growth. *Cell Rep* **30**, 3851-3863.e3856 (2020); published online EpubMar 17 (10.1016/j.celrep.2020.02.082).
185. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017); published online EpubApr (10.1038/nmeth.4197).
186. F. Lauria, T. Tebaldi, P. Bernabò, E. J. N. Groen, T. H. Gillingwater, G. Viero, riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comput Biol* **14**, e1006169 (2018); published online EpubAug (10.1371/journal.pcbi.1006169).
187. S. Chothani, E. Adami, J. F. Ouyang, S. Viswanathan, N. Hubner, S. A. Cook, S. Schafer, O. J. L. Rackham, deltaTE: Detection of Translationally Regulated Genes by Integrative Analysis of Ribo-seq and RNA-seq Data. *Current Protocols in Molecular Biology* **129**, e108 (2019)https://doi.org/10.1002/cpmb.108).
188. Y. Zhou, B. Zhou, L. Pache, M. Chang, A. H. Khodabakhshi, O. Tanaseichuk, C. Benner, S. K. Chanda, Metascape provides a biologist-oriented resource for the

analysis of systems-level datasets. *Nat Commun* **10**, 1523 (2019); published online EpubApr 3 (10.1038/s41467-019-09234-6).
189. F. Alkan, O. G. Wilkins, S. Hernández-Pérez, S. Ramalho, J. Silva, J. Ule, W. J. Faller, Identifying ribosome heterogeneity using ribosome profiling. *Nucleic Acids Res* **50**, e95 (2022); published online EpubSep 9 (10.1093/nar/gkac484).
190. P. Calamita, G. Gatti, A. Miluzio, A. Scagliola, S. Biffo, Translating the Game: Ribosomes as Active Players. *Frontiers in Genetics* **9**, (2018); published online Epub2018-November-15 (10.3389/fgene.2018.00533).
191. P. Zhou, Y. Zhang, Q. Ma, F. Gu, D. S. Day, A. He, B. Zhou, J. Li, S. M. Stevens, D. Romo, W. T. Pu, Interrogating translational efficiency and lineage-specific transcriptomes using ribosome affinity purification. *Proceedings of the National Academy of Sciences* **110**, 15395-15400 (2013)doi:10.1073/pnas.1304124110).
192. A. C. Hsieh, Y. Liu, M. P. Edlind, N. T. Ingolia, M. R. Janes, A. Sher, E. Y. Shi, C. R. Stumpf, C. Christensen, M. J. Bonham, S. Wang, P. Ren, M. Martin, K. Jessen, M. E. Feldman, J. S. Weissman, K. M. Shokat, C. Rommel, D. Ruggero, The translational landscape of mTOR signalling steers cancer initiation and metastasis. *Nature* **485**, 55-61 (2012); published online EpubFeb 22 (10.1038/nature10912).
193. Y. Xu, M. Poggio, H. Y. Jin, Z. Shi, C. M. Forester, Y. Wang, C. R. Stumpf, L. Xue, E. Devericks, L. So, H. G. Nguyen, A. Griselin, J. D. Gordan, S. E. Umetsu, S. H. Reich, S. T. Worland, S. Asthana, M. Barna, K. R. Webster, J. T. Cunningham, D. Ruggero, Translation control of the immune checkpoint in cancer and its therapeutic targeting. *Nat Med* **25**, 301-311 (2019); published online EpubFeb (10.1038/s41591-018-0321-2).
194. A. Tzatsos, R. Pfau, S. C. Kampranis, P. N. Tsichlis, Ndy1/KDM2B immortalizes mouse embryonic fibroblasts by repressing the Ink4a/Arf locus. *Proc Natl Acad Sci U S A* **106**, 2641-2646 (2009); published online EpubFeb 24 (10.1073/pnas.0813139106).
195. W. Yuan, Z. Yao, V. Veerapandian, X. Yang, X. Wang, D. Chen, L. Ma, C. Li, Y. Zheng, F. Luo, X. Y. Zhao, The histone demethylase KDM2B regulates human primordial germ cell-like cells specification. *Int J Biol Sci* **17**, 527-538 (2021)10.7150/ijbs.55873).
196. H. Y. Lam, S. Arumugam, H. G. Bae, C. C. Wang, S. Jung, A. L. St John, W. Hong, W. Han, V. Tergaonkar, ELKS1 controls mast cell degranulation by regulating the transcription of Stxbp2 and Syntaxin 4 via Kdm2b stabilization. *Sci Adv* **6**, (2020); published online EpubJul (10.1126/sciadv.abb2497).
197. Q. Zhou, Y. Zhang, B. Wang, W. Zhou, Y. Bi, W. Huai, X. Chen, Y. Chen, Z. Liu, X. Liu, Z. Zhan, KDM2B promotes IL-6 production and inflammatory responses through Brg1-mediated chromatin remodeling. *Cell Mol Immunol* **17**, 834-842 (2020); published online EpubAug (10.1038/s41423-019-0251-z).
198. G. Jiang, S. Zhang, A. Yazdanparast, M. Li, A. V. Pawar, Y. Liu, S. M. Inavolu, L. Cheng, Comprehensive comparison of molecular portraits between cell lines and tumors in breast cancer. *BMC Genomics* **17**, 525 (2016); published online Epub2016/08/22 (10.1186/s12864-016-2911-z).
199. X. Jin, Z. Demere, K. Nair, A. Ali, G. B. Ferraro, T. Natoli, A. Deik, L. Petronio, A. A. Tang, C. Zhu, L. Wang, D. Rosenberg, V. Mangena, J. Roth, K. Chung, R.

K. Jain, C. B. Clish, M. G. Vander Heiden, T. R. Golub, A metastasis map of human cancer cell lines. *Nature* **588**, 331-336 (2020); published online EpubDec (10.1038/s41586-020-2969-2).
200. A. Prat, J. S. Parker, C. Fan, C. M. Perou, PAM50 assay and the three-gene model for identifying the major and clinically relevant molecular subtypes of breast cancer. *Breast Cancer Res Treat* **135**, 301-306 (2012); published online EpubAug (10.1007/s10549-012-2143-0).
201. L. Matthews, G. Gopinath, M. Gillespie, M. Caudy, D. Croft, B. de Bono, P. Garapati, J. Hemish, H. Hermjakob, B. Jassal, A. Kanapin, S. Lewis, S. Mahajan, B. May, E. Schmidt, I. Vastrik, G. Wu, E. Birney, L. Stein, P. D'Eustachio, Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Research* **37**, D619-D622 (2008)10.1093/nar/gkn863).
202. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012); published online EpubOct 4 (10.1038/nature11412).
203. C. A. Hirsch, H. H. Hiatt, Turnover of liver ribosomes in fed and in fasted rats. *J Biol Chem* **241**, 5936-5940 (1966); published online EpubDec 25 (
204. J. Defoiche, Y. Zhang, L. Lagneaux, R. Pettengell, A. Hegedus, L. Willems, D. C. Macallan, Measurement of ribosomal RNA turnover in vivo by use of deuterium-labeled glucose. *Clin Chem* **55**, 1824-1833 (2009); published online EpubOct (10.1373/clinchem.2008.119446).
205. E. N. Nikolov, M. D. Dabeva, T. K. Nikolov, Turnover of ribosomes in regenerating rat liver. *Int J Biochem* **15**, 1255-1260 (1983)10.1016/0020-711x(83)90215-x).
206. F. Sinturel, A. Gerber, D. Mauvoisin, J. Wang, D. Gatfield, J. J. Stubblefield, C. B. Green, F. Gachon, U. Schibler, Diurnal Oscillations in Liver Mass and Cell Size Accompany Ribosome Assembly Cycles. *Cell* **169**, 651-663.e614 (2017); published online EpubMay 4 (10.1016/j.cell.2017.04.015).
207. C. C. Wu, A. Peterson, B. Zinshteyn, S. Regot, R. Green, Ribosome Collisions Trigger General Stress Responses to Regulate Cell Fate. *Cell* **182**, 404-416.e414 (2020); published online EpubJul 23 (10.1016/j.cell.2020.06.006).
208. E. Chavdoula, V. Anastas, A. La Ferlita, J. Aldana, G. Carota, M. Spampinato, B. Soysal, I. Cosentini, S. Parashar, A. Sircar, G. Nigita, L. Sehgal, M. A. Freitas, P. N. Tsichlis, Transcriptional regulation of amino acid metabolism by KDM2B, in the context of ncPRC1.1 and in concert with MYC and ATF4. *Metabolism* **150**, 155719 (2024); published online EpubJan (10.1016/j.metabol.2023.155719).
209. C. de Haro, R. Mendez, J. Santoyo, The eIF-2alpha kinases and the control of protein synthesis. *Faseb j* **10**, 1378-1387 (1996); published online EpubOct (10.1096/fasebj.10.12.8903508).
210. B. R. Williams, PKR; a sentinel kinase for cellular stress. *Oncogene* **18**, 6112-6120 (1999); published online EpubNov 1 (10.1038/sj.onc.1203127).
211. A. G. Hinnebusch, Translational regulation of GCN4 and the general amino acid control of yeast. *Annu Rev Microbiol* **59**, 407-450 (2005)10.1146/annurev.micro.59.031805.133833).
212. N. T. Ingolia, S. Ghaemmaghami, J. R. Newman, J. S. Weissman, Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.

Science **324**, 218-223 (2009); published online EpubApr 10 (10.1126/science.1168978).
213. K. Leppek, R. Das, M. Barna, Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* **19**, 158-174 (2018); published online EpubMar (10.1038/nrm.2017.103).
214. F. Bertucci, P. Finetti, D. Birnbaum, Basal breast cancer: a complex and deadly molecular subtype. *Curr Mol Med* **12**, 96-110 (2012); published online EpubJan (10.2174/156652412798376134).
215. G. Botti, M. Cantile, F. Collina, M. Cerrone, S. Sarno, A. Anniciello, M. Di Bonito, Morphological and pathological features of basal-like breast cancer. *Transl Cancer Res* **8**, S503-s509 (2019); published online EpubOct (10.21037/tcr.2019.06.50).
216. E. A. Rakha, M. E. El-Sayed, J. Reis-Filho, I. O. Ellis, Patho-biological aspects of basal-like breast cancer. *Breast Cancer Res Treat* **113**, 411-422 (2009); published online EpubFeb (10.1007/s10549-008-9952-1).
217. J. Y. Kang, J. W. Park, J. Y. Hahm, H. Jung, S. B. Seo, Histone H3K79 demethylation by KDM2B facilitates proper DNA replication through PCNA dissociation from chromatin. *Cell Prolif* **53**, e12920 (2020); published online EpubNov (10.1111/cpr.12920).
218. I. C. Kurt, I. Sur, E. Kaya, A. Cingoz, S. Kazancioglu, Z. Kahya, O. D. Toparlak, F. Senbabaoglu, Z. Kaya, E. Ozyerli, S. Karahüseyinoglu, N. A. Lack, Z. H. Gümüs, T. T. Onder, T. Bagci-Onder, KDM2B, an H3K36-specific demethylase, regulates apoptotic response of GBM cells to TRAIL. *Cell Death Dis* **8**, e2897 (2017); published online EpubJun 29 (10.1038/cddis.2017.288).
219. C. J. Spangler, A. Skrajna, C. A. Foley, A. Nguyen, G. R. Budziszewski, D. N. Azzam, E. C. Arteaga, H. C. Simmons, C. B. Smith, N. A. Wesley, E. M. Wilkerson, J. E. McPherson, D. Kireev, L. I. James, S. V. Frye, D. Goldfarb, R. K. McGinty, Structural basis of paralog-specific KDM2A/B nucleosome recognition. *Nat Chem Biol*, (2023); published online EpubFeb 16 (10.1038/s41589-023-01256-y).
220. M. Ozawa, T. Fukuda, R. Sakamoto, H. Honda, N. Yoshida, The Histone Demethylase FBXL10 Regulates the Proliferation of Spermatogonia and Ensures Long-Term Sustainable Spermatogenesis in Mice1. *Biology of Reproduction* **94**, (2016)10.1095/biolreprod.115.135988).
221. B. J. E. Martin, E. F. Ablondi, C. Goglia, K. Adelman, Global identification of direct SWI/SNF targets reveals compensation by EP400. *bioRxiv*, 2023.2003.2007.531379 (2023)10.1101/2023.03.07.531379).
222. L. Zhang, Z. Dai, J. Yu, M. Xiao, CpG-island-based annotation and analysis of human housekeeping genes. *Brief Bioinform* **22**, 515-525 (2021); published online EpubJan 18 (10.1093/bib/bbz134).
223. E. Y. Lee, W. J. Muller, Oncogenes and tumor suppressor genes. *Cold Spring Harb Perspect Biol* **2**, a003236 (2010); published online EpubOct (10.1101/cshperspect.a003236).
224. Pan-cancer analysis of whole genomes. *Nature* **578**, 82-93 (2020); published online EpubFeb (10.1038/s41586-020-1969-6).

225. S. Singh, A. Vanden Broeck, L. Miller, M. Chaker-Margot, S. Klinge, Nucleolar maturation of the human small subunit processome. *Science* **373**, eabj5338 (2021); published online EpubSep 10 (10.1126/science.abj5338).
226. M. Chaker-Margot, J. Barandun, M. Hunziker, S. Klinge, Architecture of the yeast small subunit processome. *Science* **355**, eaal1880 (2017)doi:10.1126/science.aal1880).
227. L. Tafforeau, C. Zorbas, J.-L. Langhendries, S.-T. Mullineux, V. Stamatopoulou, R. Mullier, L. Wacheul, Denis L. J. Lafontaine, The Complexity of Human Ribosome Biogenesis Revealed by Systematic Nucleolar Screening of Pre-rRNA Processing Factors. *Molecular Cell* **51**, 539-551 (2013); published online Epub2013/08/22/ (https://doi.org/10.1016/j.molcel.2013.08.011).
228. Z. Liu, Y. Lv, N. Zhao, G. Guan, J. Wang, Protein kinase R-like ER kinase and its role in endoplasmic reticulum stress-decided cell fate. *Cell Death Dis* **6**, e1822 (2015); published online EpubJul 30 (10.1038/cddis.2015.183).
229. P. Herr, J. Boström, E. Rullman, S. G. Rudd, M. Vesterlund, J. Lehtiö, T. Helleday, G. Maddalo, M. Altun, Cell Cycle Profiling Reveals Protein Oscillation, Phosphorylation, and Localization Dynamics. *Mol Cell Proteomics* **19**, 608-623 (2020); published online EpubApr (10.1074/mcp.RA120.001938).
230. K. A. Bernstein, S. J. Baserga, The small subunit processome is required for cell cycle progression at G1. *Mol Biol Cell* **15**, 5038-5046 (2004); published online EpubNov (10.1091/mbc.e04-06-0515).
231. Y. Luan, N. Tang, J. Yang, S. Liu, C. Cheng, Y. Wang, C. Chen, Y.-n. Guo, H. Wang, W. Zhao, Q. Zhao, W. Li, M. Xiang, R. Ju, Z. Xie, Deficiency of ribosomal proteins reshapes the transcriptional and translational landscape in human cells. *Nucleic Acids Research* **50**, 6601-6617 (2022)10.1093/nar/gkac053).
232. M. Gérus, M. Caizergues-Ferrer, Y. Henry, A. Henras, in *The Nucleolus,* M. O. J. Olson, Ed. (Springer New York, New York, NY, 2011), pp. 157-184.
233. C. R. Stumpf, M. V. Moreno, A. B. Olshen, B. S. Taylor, D. Ruggero, The translational landscape of the mammalian cell cycle. *Mol Cell* **52**, 574-582 (2013); published online EpubNov 21 (10.1016/j.molcel.2013.09.018).
234. Q. Vicens, J. S. Kieft, O. S. Rissland, Revisiting the Closed-Loop Model and the Nature of mRNA 5'-3' Communication. *Mol Cell* **72**, 805-812 (2018); published online EpubDec 6 (10.1016/j.molcel.2018.10.047).
235. D. W. Rogers, M. A. Böttcher, A. Traulsen, D. Greig, Ribosome reinitiation can explain length-dependent translation of messenger RNA. *PLoS Comput Biol* **13**, e1005592 (2017); published online EpubJun (10.1371/journal.pcbi.1005592).
236. M. K. Thompson, M. F. Rojas-Duran, P. Gangaramani, W. V. Gilbert, The ribosomal protein Asc1/RACK1 is required for efficient translation of short mRNAs. *Elife* **5**, (2016); published online EpubApr 27 (10.7554/eLife.11154).
237. M. K. Thompson, W. V. Gilbert, mRNA length-sensing in eukaryotic translation: reconsidering the "closed loop" and its implications for translational control. *Curr Genet* **63**, 613-620 (2017); published online EpubAug (10.1007/s00294-016-0674-3).
238. J. Costello, L. M. Castelli, W. Rowe, C. J. Kershaw, D. Talavera, S. S. Mohammad-Qureshi, P. F. G. Sims, C. M. Grant, G. D. Pavitt, S. J. Hubbard, M. P. Ashe, Global mRNA selection mechanisms for translation initiation. *Genome*

Biology **16**, 10 (2015); published online Epub2015/01/05 (10.1186/s13059-014-0559-z).
239. C. de la Parra, A. Ernlund, A. Alard, K. Ruggles, B. Ueberheide, R. J. Schneider, A widespread alternate form of cap-dependent mRNA translation initiation. *Nat Commun* **9**, 3068 (2018); published online EpubAug 3 (10.1038/s41467-018-05539-0).
240. Y. Liu, J. Cui, A. R. Hoffman, J.-F. Hu, Eukaryotic translation initiation factor eIF4G2 opens novel paths for protein synthesis in development, apoptosis and cell differentiation. *Cell Proliferation* **56**, e13367 (2023)https://doi.org/10.1111/cpr.13367).
241. A. S. Lee, P. J. Kranzusch, J. H. Cate, eIF3 targets cell-proliferation messenger RNAs for translational activation or repression. *Nature* **522**, 111-114 (2015); published online EpubJun 4 (10.1038/nature14267).
242. T. R. Hartman, S. Qian, C. Bolinger, S. Fernandez, D. R. Schoenberg, K. Boris-Lawrie, RNA helicase A is necessary for translation of selected messenger RNAs. *Nat Struct Mol Biol* **13**, 509-516 (2006); published online EpubJun (10.1038/nsmb1092).
243. K. M. Kim, H. Cho, K. Choi, J. Kim, B. W. Kim, Y. G. Ko, S. K. Jang, Y. K. Kim, A new MIF4G domain-containing protein, CTIF, directs nuclear cap-binding protein CBP80/20-dependent translation. *Genes Dev* **23**, 2033-2045 (2009); published online EpubSep 1 (10.1101/gad.1823409).
244. S. I. A. Bukhari, S. S. Truesdell, S. Lee, S. Kollu, A. Classon, M. Boukhali, E. Jain, R. D. Mortensen, A. Yanagiya, R. I. Sadreyev, W. Haas, S. Vasudevan, A Specialized Mechanism of Translation Mediated by FXR1a-Associated MicroRNP in Cellular Quiescence. *Mol Cell* **61**, 760-773 (2016); published online EpubMar 3 (10.1016/j.molcel.2016.02.013).
245. S. J. Jeong, S. Park, L. T. Nguyen, J. Hwang, E. Y. Lee, H. K. Giong, J. S. Lee, I. Yoon, J. H. Lee, J. H. Kim, H. K. Kim, D. Kim, W. S. Yang, S. Y. Kim, C. Y. Lee, K. Yu, N. Sonenberg, M. H. Kim, S. Kim, A threonyl-tRNA synthetase-mediated translation initiation machinery. *Nat Commun* **10**, 1357 (2019); published online EpubMar 22 (10.1038/s41467-019-09086-0).
246. G. Singh, S. E. Fritz, B. Seufzer, K. Boris-Lawrie, The mRNA encoding the JUND tumor suppressor detains nuclear RNA-binding proteins to assemble polysomes that are unaffected by mTOR. *J Biol Chem* **295**, 7763-7773 (2020); published online EpubMay 29 (10.1074/jbc.RA119.012005).